Time course of clinical change following neurofeedback.

Mariela Rance[1], **Christopher Walsh**[1], **Denis G. Sukhodolsky**[2], **Brian Pittman**[3], **Maolin Qiu**[1], **Stephen A. Kichuk**[3], **Suzanne Wasylink**[3], **William N. Koller**[1], **Michael Bloch**[3], **Patricia Gruner**[3], **Dustin Scheinost**[1,2], **Christopher Pittenger**[2,3,4], and **Michelle Hampson**[1,2,3,*]

[1]Department of Radiology and Biomedical Imaging, Yale University School of Medicine, New Haven, CT 06520, USA

[2]Child Study Center, Yale University School of Medicine, New Haven, CT 06519, USA

[3]Department of Psychiatry, Yale University School of Medicine, New Haven, CT 06511, United States

[4]Department of Psychology, Yale University, New Haven, CT 06520, USA;

Abstract

Neurofeedback – learning to modulate brain function through real-time monitoring of current brain state – is both a powerful method to perturb and probe brain function and an exciting potential clinical tool. For neurofeedback effects to be useful clinically, they must persist. Here we examine the time course of symptom change following neurofeedback in two clinical populations, combining data from two ongoing neurofeedback studies. This analysis reveals a shared pattern of symptom change, in which symptoms continue to improve for weeks after neurofeedback. This time course has several implications for future neurofeedback studies. Most neurofeedback studies are not designed to test an intervention with this temporal pattern of response. We recommend that new studies incorporate regular follow-up of subjects for weeks or months after the intervention to ensure that the time point of greatest effect is sampled. Furthermore, this time course of continuing clinical change has implications for crossover designs, which may attribute long-term, ongoing effects of real neurofeedback to the control intervention that follows. Finally, interleaving neurofeedback sessions with assessments and examining when clinical improvement peaks may not be an appropriate approach to determine the optimal number of sessions for an application.

Introduction

Real-time functional magnetic resonance imaging (rt-fMRI) neurofeedback is a novel, noninvasive approach to altering human brain function (Sitaram et al., 2017; Weiskopf et al., 2003). rt-fMRI neurofeedback can induce alterations in many different aspects of mental

*Corresponding author: Michelle Hamspon, Magnetic Resonance Research Center (MRRC), Department of Radiology and Biomedical Imaging, Yale University School of Medicine, The Anlyan Center, N121, 300 Cedar Street, New Haven, CT 06520-8043, USA; michelle.hampson@yale.edu.

function (Caria et al., 2010; Cortese et al., 2016; deBettencourt et al., 2015; Grone et al., 2015; Koush et al., 2017; Rota et al., 2009; Scharnowski et al., 2015; Zhang et al., 2013a) and shows promise for improving a variety of different neuropsychiatric symptoms (deCharms et al., 2005; Gerin et al., 2016; Linden et al., 2012; Paret et al., 2016; Scheinost et al., 2013; Subramanian et al., 2011; Young et al., 2017). Furthermore, changes in brain function and behavior induced by rt-fMRI neurofeedback have been reported to persist for weeks (Subramanian et al., 2011; Yoo et al., 2007; Yoo et al., 2008) or even months to years (Amano et al., 2016; Megumi et al., 2015; Ramot et al., 2017; Robineau et al., 2017) after the intervention. This is consistent with the long-lasting effects of neurofeedback that have been reported in the EEG literature (Engelbregt et al., 2016; Kotchoubey et al., 1997; Surmeli and Ertem, 2011; Surmeli et al., 2012). Despite these data supporting the persistence of learning effects induced by neurofeedback, the time course of changes in brain function and behavior following neurofeedback are not well characterized. Understanding the time course of clinical change following neurofeedback is important not only to inform patients' expectations, but also to optimize the design of neurofeedback experiments.

In this manuscript, we present an analysis designed to explore the temporal pattern of symptom response to neurofeedback. As we are interested in patterns that are potentially shared across neurofeedback applications, the analysis combines data from two different real-time fMRI neurofeedback studies: a study training control over the ventral frontal cortex in adults with obsessive-compulsive disorder (OCD; NCT02206945), and a second study training control over the supplementary motor area in adolescents with Tourette Syndrome (TS; NCT01702077). Importantly, the purpose of this manuscript is not to present interim analyses of the outcome data from either of these ongoing clinical trials, and therefore the data from the trials are not examined individually. Instead, the results of an omnibus analysis combining data from both studies are presented to characterize temporal effects of neurofeedback that may be shared across applications.

In particular, in the patients who appeared to respond to neurofeedback in the two studies, we noticed a surprising pattern of symptom change. Although we anticipated symptom improvement during and shortly after neurofeedback that would subsequently persist or gradually return to baseline, we instead noted *continuing* symptom improvement over the weeks following completion of the neurofeedback interventions. Therefore, our analysis was designed to address the question of whether there was a statistically significant improvement in symptoms in the weeks after the intervention that was shared across studies. If such a pattern is relatively common in neurofeedback studies, it is critical that neurofeedback researchers are aware of the possibility that their application will show such effects, so they can design their studies with this in mind, for example, by including long-term follow-up of patients to ensure studies are fully powered.

Materials and Methods

All study procedures were approved by the Human Investigation Committee of the Yale School of Medicine. Written informed consent was obtained from all subjects.

OCD study (NCT02206945)

Subjects: A total of 17 subjects with OCD were recruited through the Yale OCD Research Clinic (ocd.yale.edu). Subjects were between 18 and 60 years old and had a primary diagnosis of OCD according to Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV). All subjects had a score of between 16 and 32, corresponding to moderate-to-severe symptom severity, on the Yale-Brown Obsessive Compulsive Scale (Y-BOCS; Goodman et al., 1989a; Goodman et al., 1989b), with symptoms primarily in the checking or contamination domain (Bloch et al., 2008). Subjects were free of psychotropic medication or stably medicated with selective serotonin reuptake inhibitors (SSRIs) or clomipramine (stable dose 8 weeks and throughout the study period); low-dose hypnotic or anxiolytic medications taken occasionally on an as-needed basis were permitted. Subjects were excluded if they had initiated cognitive behavioral therapy within 3 months of study enrollment. Subjects with bipolar I or a psychotic disorder, autism spectrum disorder, significant neurological abnormalities, or current or recent (6 months) substance use disorder were excluded.

Study protocol: Subjects underwent two sessions of 6 rt-fMRI neurofeedback training runs, 12 in total, approximately half a week apart, following a protocol described previously (Hampson et al., 2012; Scheinost et al., 2013). During feedback runs of 4 minutes and 20 seconds (total training time: 52 minutes), current activity of an individually localized symptom-responsive region within the orbitofrontal/frontal polar cortex was provided at the bottom of the screen in the form of a line graph, which was updated as each brain volume was collected (Figure 1A). Subjects were cued by an arrow on the left side of the screen as to their current task: rest (white arrow pointing to the right); increase activity in target region (red arrow pointing up); or decrease activity (blue arrow pointing down). During the increase and decrease blocks, images were shown that were designed to provoke OCD symptoms (contamination or checking). During the resting blocks, neutral images were shown. A detailed description of scan parameters, the localization of individual regions, and the feedback computation can be found in the Supplementary Material.

Subjects were randomized to receive either real or yoked sham neurofeedback. Both subjects and clinical raters were blind to intervention assignment. For yoked sham feedback, the line graphs of brain activity presented to the subject were taken not from their own brain patterns, but from a matched real neurofeedback subject's brain pattern. The blocks were time-locked across subjects, so the degree to which a neurofeedback subject succeeded in increasing/decreasing activity during the correct blocks was captured in their time-course, and the matched sham subject was led to believe they were having the same level of success. Thus, this control is designed to match the perception of success during neurofeedback, which may influence clinical response. Subjects were informed that they would be randomly assigned to receive either an experimental feedback intervention that we hoped would help symptoms or a control feedback intervention that we did not believe would help symptoms.

Clinical status was assessed at baseline (prior to starting feedback), post-feedback (typically half a week after completing 2 sessions of sham or neurofeedback), at 2 weeks post-intervention, and 1 month post-intervention. After observing the temporal pattern of change

that motivated the current report, we added two more clinical follow-ups at 6 weeks and 2 months post-intervention; a subset of recently completed subjects (n = 5) had these additional follow-up assessments.

TS study (NCT01702077)

Subjects: Twenty adolescents with Tourette Syndrome with at least moderate level of current tic severity, defined by a total score of 13 or higher on the Yale Global Tic Severity Scale (YGTSS; Leckman et al., 1989), were recruited from the Yale Child Study Center Tic Disorder Specialty Clinic and the greater New Haven area. A DSM-IV Diagnosis of TS was assigned based a on structured TS module developed at the Yale Child Study Center that inquired about tic onset, developmental history, and current and past motor and vocal tics. Stable medication was allowed, if unchanged in the past month.

Study Protocol: Subjects underwent a randomized, crossover trial involving four feedback training visits (two for each arm of the trial). The experimental arm involved neurofeedback from the individually localized supplementary motor area (SMA). The control arm involved yoked sham feedback (with the time courses for each subject taken from the real time courses of the preceding subject's neurofeedback scans, as described for the OCD study above). Each feedback visit consisted of 6 rt-fMRI neurofeedback training runs, 12 in total, each with a duration of 2 minutes and 46 seconds (total training time: 32 minutes and 48 seconds). The 2 feedback visits were approximately half a week apart. Subjects saw the activation in their brain area represented by a line graph, similar to that in OCD protocol. They were cued by a red arrow pointing up to cause the line to climb up and a blue down arrow to cause the line to go down. Subjects' clinical status was assessed using the YGTSS at baseline (typically half a week prior-to the start of each arm) and post-intervention (typically half a week after completion of each arm) for each of the two arms. The MR data acquisition, localization of the individual region, and feedback computation is described in detail in the Supplementary Material. Both the subject and the clinical rater were blind to which intervention was received first. An example screen shot from the end of a neurofeedback run is shown in Figure 1B.

Approximately half of the subjects moved directly from the first arm of this crossover design into the second, such that the post-assessment for first arm was used as the pre-assessment of the second arm. However, due to scheduling constraints and other practical considerations, the other half of the subjects had the two arms of the intervention separated in time; the delay between the two arms was variable across subjects, although always less than one month. For subjects with the two arms separated in time, an assessment conducted half a week after the completion of the first arm served as the post-assessment for that arm, and a second assessment was conducted half a week before the initiation of feedback in the second arm. For these subjects, the pre-assessment scan for the second arm of the crossover design may be considered a delayed follow-up assessment for the first arm. This separation of the two arms in half of our subjects thus had the fortuitous benefit of allowing us to examine the trajectory of neurofeedback effects over time in the present analysis.

It is important to note that ongoing symptom change after neurofeedback would lead to carryover effects in this crossover design. Indeed, if symptoms continue to improve during the weeks following neurofeedback, then subjects who received true neurofeedback first and sham feedback second might show symptom improvement during the second/sham arm that is in fact attributable to real neurofeedback delivered during the first arm. To avoid any contamination caused by such carryover effects, sham neurofeedback data from subjects who received sham feedback in the second arm were not included for analysis as part of the sham neurofeedback intervention type. Instead, these data were treated as a late follow-up for real neurofeedback and were included in the real neurofeedback intervention type in the primary analysis. However, realizing that this may conflate potential effects caused by the sham intervention (e.g., a placebo response) with ongoing effects of the neurofeedback intervention, a follow-up analysis was performed that excluded these data (see Supplementary Figure S1 for details).

Data analyses

For each assessment collected post-intervention, we coded latency as the number of days that had elapsed between the start of neurofeedback (i.e., the start of the applicable arm of intervention in the case of the crossover TS study) and the day that the assessment was collected.

Transformation of clinical ratings to shared space: In order to analyze the data from these two different studies (TS and OCD) together, the clinical data were normalized in such a way that data from both studies were in a comparable space before conducting a general linear model analysis (GLM) examining time course effects. To ensure results of the analysis were not an artifact of the preprocessing adopted, we used two different approaches to preprocessing the data and confirmed that they yielded similar results. These are:

1. Z-transformation: For the OCD study, we computed the mean and the standard deviation of baseline Y-BOCS measurements across all subjects (including both the neurofeedback and sham subjects). For every Y-BOCS measure, we then subtracted this mean and divided by the standard deviation. This results in a mean of zero and a standard deviation of 1 for baseline scores, with all subsequent data points in the same space. A similar z- transformation was performed on the YGTSS measures from the TS study, using the mean and standard deviation of baseline YGTSS (taken before the first neurofeedback arm).

2. Percent change: For each assessment collected post-intervention, we computed the percent change from individual baseline.

Main analyses: Data were analyzed using a mixed model with intervention type (a binary variable indicating neurofeedback or sham), latency (continuous variable indicating for each assessment how many days had elapsed since the intervention began) as fixed effects, and normalized clinical scores as the dependent variable, together with all interaction terms. Random subject effects were also included to account for the correlation between multiple observations within the same subject. All assessments, including the baseline assessments

(latency being coded as zero for all baseline assessments), were included in the z-transformed data analysis. For percent change measures, only post-assessment time-points were included. Significant effects were interpreted by estimating slopes for each group from the model post-hoc. To explore whether findings could be driven exclusively by one of the studies (and thus did not represent a shared pattern across the TS and OCD studies), we repeated these analyses including the variable study (TS/OCD) as an additional fixed effect, together with all its interaction terms.

Analysis excluding baseline time point (to examine post-intervention improvement): In order to examine the extent to which symptom improvement occurred after the completion of the real or sham neurofeedback intervention, the z-transformed data were reanalyzed discarding the baseline time point. Note that the baseline time point was never included in the percent signal change analysis, so this analysis was only relevant for the z-transformed data.

Analyses discarding final assessment for TS subjects who received sham as the second arm: For those subjects receiving sham in the second arm of the TS crossover study, symptoms may be improving during this sham arm due to the prior neurofeedback arm. Therefore, in the main analyses, their final assessment was treated as a long-term follow-up for the preceding neurofeedback arm. However, if there is some response to the sham intervention (e.g., a placebo response), it could bias us to find significant symptom improvement after neurofeedback. Therefore, all analyses were repeated, discarding this final time point to remove any potential for such bias.

Results

Main analyses:

Mixed model analysis of z-transformed data revealed no main effect of intervention type; there was a significant effect of latency ($F(104)=21.88$; $p<0.0001$) and a significant interaction between latency and intervention type ($F(104)=12.11$; $p=0.0007$). Post-hoc analyses revealed a significant negative slope of symptom improvement for the real ($t(104)= -6.76$; $p<0.0001$) but not for the sham ($t(104)= -0.75$; $p=0.45$) intervention (Figure 2). There were no main effects or interaction effects of study when it was included in the model, and its inclusion did not alter the results.

A similar analysis of percent change data showed a significant interaction between latency and intervention type ($F(58)=6.71$; $p=0.012$). Post-hoc analyses again revealed a significant negative slope of symptom improvement for the real ($t(58)= -3.64$; $p =0.0006$) but not for the sham intervention ($t(58)=0.45$; $p =0.65$; Figure 3). There were no main effects or interaction effects of study when it was included in the model, and its inclusion did not alter the results.

Analysis excluding baseline time point (to examine post-intervention improvement)

Effects that occurred after the intervention (as opposed to during the intervention) can be isolated in z-transformed data by discarding the first time-point. This analysis yielded a

significant interaction between intervention type and latency (F(58)=5.56; p=0.02). Post-hoc analyses revealed this effect to be driven by a significant slope for the neurofeedback group (t(58)= −3.52; p=0.0009) but not for the sham intervention (t(58)=0.25; p=0.80). These results suggest ongoing symptom improvement in the neurofeedback group *after* the intervention was completed, above and beyond any symptom change that occurred during the intervention itself.

Analyses discarding final assessment for TS subjects who received sham as the second arm:

When the main analysis was repeated discarding the final assessment for subjects who received sham in the second arm of the intervention, the results remained the same for both the z-score and percent change analyses.

Discussion

Neurofeedback shows promise as a noninvasive strategy to modulate dysregulated brain circuitry in neuropsychiatric disease. Before initiating our studies of neurofeedback in OCD and TS, we anticipated that any clinical benefit would be maximal immediately after the neurofeedback sessions, and then fade with time. Contrary to this expectation, across two different applications of neurofeedback, we find that clinical improvements grew in the weeks following completion of the intervention. This consistent temporal pattern of symptom change is particularly remarkable given that the two studies targeted different brain areas in distinct patient populations and used different clinical instruments to assess different types of symptoms. The shared pattern across these two very different studies suggests that this pattern may represent a characteristic temporal pattern of response to neurofeedback training.

A reëxamination of published studies reveals hints of a similar pattern of change in previous neurofeedback experiments. For example, when subjects were trained to associate line orientation with color, the perceptual shifts induced after neurofeedback were more pronounced at a 3–5 month follow-up than immediately after the intervention (compare Figures 3b and S2 in Amano et al., 2016). The follow-up group in that study was a subset of the original sample, so we cannot rule out the possibility that this is due to nonrandom dropouts, but the observed pattern is consistent with a growing effect over time. In a separate study of neurofeedback for depression that followed subjects for four weeks post-neurofeedback there was a pattern of continuing symptom improvement over the follow-up period (see Figure 2 of Schnyer et al., 2015).

The extent to which these temporal effects are limited to clinical/behavioral changes is unclear, but there is data suggesting that measurable changes in brain function after neurofeedback may show a similar pattern. In a study of resting state functional connectivity changes induced by neurofeedback, connectivity changes were more pronounced a day after neurofeedback than they were immediately post-neurofeedback (Harmelech et al., 2013). In a study examining the maintenance of learning effects at 6 and 14 month post-intervention, measures of control over the brain area were numerically larger (albeit not significantly so) at follow-up than they were during training (Robineau et al., 2017). Finally, in a

neurofeedback study in depression, resting state connectivity changes induced by the intervention were shown to grow over a period of weeks following the intervention (Figure 5 of Yuan et al., 2014).

A similar pattern of symptom improvements and brain changes that grow over time has been occasionally reported after behavioral interventions. For example, in a randomized trial of family cognitive behavioral therapy for pediatric OCD, symptoms at the one and six month follow-up assessments were significantly better than immediately post-intervention (Piacentini et al., 2011). Similarly, trials of addiction treatments have reported delayed effects of cognitive-behavioral therapy emerging months after the intervention (Carroll et al., 1994; Goldstein et al., 1989). Also, a study of neurofeedback for spider phobia that involved exposure to spider imagery and cognitive reappraisal for both the neurofeedback and control groups showed symptom improvements that grew over the months following the intervention in both groups, possibly due to long term effects of exposure with cognitive reappraisal (Zilverstand et al., 2015). Behavioral interventions aim to produce lasting changes in coping skills and thereby to reduce psychiatric symptoms. After an individual learns a coping skill, he or she may continue to use it, possibly becoming more proficient and integrating the skill into their life habits, such that benefits accrue and symptoms continue to decrease. The same explanatory framework may be applicable to neurofeedback interventions, in which a subject learns to control neural activity. It is possible that, having learned to control neural activity during neurofeedback, subjects continue to practice this newly acquired skill, which may result in continuing improvement of both symptoms and neural reorganization.

Alternatively, the phenomenon of behavioral/symptom change that accrues over time may be better explained at the neural level (though of course neural and psychological explanations need not be mutually exclusive). Neurofeedback is a form of learning, and it is well established that learning undergoes a series of consolidation and reconsolidation processes over time (Dudai, 2012; Kandel et al., 2014); gradual improvement over the weeks following a neurofeedback intervention may reflect such slow consolidation processes, which continue irrespective of how much the skills are practiced. At the network level, we and others have found neurofeedback to alter the correlational structure of network brain activity post-intervention (Hampson et al., 2011; Harmelech et al., 2013; Megumi et al., 2015; Scheinost et al., 2013; Yuan et al., 2014; Zhang et al., 2013b). Following the Hebbian principle that structures that 'fire together wire together' (Hebb, 1949; Lowel and Singer, 1992), these changes in correlational structure may self-reinforce over time: that is, brain regions whose firing becomes more correlated after neurofeedback may, by way of this correlation, become increasingly tightly coupled over time, while structures that are desynchronized by neurofeedback may similarly become increasingly disconnected.

Such mechanistic speculations have been discussed previously and are questions for future study (Gevensleben et al., 2014). Critically, however, they all are likely to generalize to other applications of neurofeedback, which brings us to our initial observation: that the appearance of a similar pattern of changes that increase over time in two distinct datasets suggests that this effect is likely to generalize to other neurofeedback applications. This conclusion leads to three major implications.

First, any study which does not follow up subjects for at least several weeks may not be sampling the time point of greatest effect, and may therefore be underpowered, leading to false negative results. Our data suggest that regularly assessing subjects following neurofeedback over a period of months will not only allow better characterization of the time courses of neurofeedback responses across applications, but more critically, it will ensure that studies are fully powered by sampling at the time point of greatest effect. As clinical applications of fMRI neurofeedback are relatively new, it may be helpful to consider follow up durations from clinical trials of other learning-based interventions, such as psychotherapy. Such trials often include a follow-up period of several months (Freeman et al., 2014; Piacentini et al., 2011; Piacentini et al., 2010; Wilhelm et al., 2012). In recent years, a growing number of rt-fMRI neurofeedback studies have included clinical follow-ups months after the intervention (Amano et al., 2016; Cox et al., 2016; Megumi et al., 2015; Ramot et al., 2017; Robineau et al., 2017). Our data indicate this promising development could be further enhanced by more frequent sampling of clinical changes over the first month or two post-intervention.

Second, cross-over studies (such as our study of TS) may be contaminated by substantial carry-over effects. As the time course of changes have not yet been characterized across neurofeedback applications, there will generally be no assurance that symptoms will have stabilized from the active intervention delivered in the first arm, even when the arms are spaced weeks apart. Even if the arms are spaced far enough apart to allow for symptom changes to stabilize, symptom severity level between the two groups may differ going into the second arm of the intervention if subjects who received neurofeedback in the first arm stabilize in a less symptomatic state. This may or may not be a concern for studies of basic mechanism, but is for clinical trials. This needs to be taken into consideration in the analysis of such studies.

Finally, when optimizing the number of neurofeedback sessions used in a particular application, a popular approach is to run repeated neurofeedback sessions with assessments in between, to establish when symptom improvement peaks. This approach assumes that symptom improvement between adjacent assessments is driven by the neurofeedback session between those assessments, rather than the delayed effects of earlier sessions. The pattern of change shown here undercuts this assumption, and suggests that this approach may substantially overestimate the optimal number of sessions in a particular application. A better approach for optimizing the number of sessions is a multiarm intervention study, although the expensive and time-consuming nature of such studies has limited their adoption.

In light of traditional EEG neurofeedback protocols that involve large numbers of sessions of neurofeedback, the changes seen here with two sessions of fMRI neurofeedback may seem surprisingly large. However, two sessions is a typical amount of training for clinical applications of rt-fMRI neurofeedback, and seems in many cases to be sufficient (Scheinost et al., 2013; Subramanian et al., 2011; Young et al., 2017). Perhaps a small number of sessions would be significant for some EEG protocols as well, if subjects were followed up for a month or two post-intervention to allow the full effects to manifest themselves. In general, more work is needed to determine the optimal number of sessions of neurofeedback

across different applications and modalities. Unfortunately, as discussed in the previous paragraph, proper optimization requires expensive, time-consuming, multiarm trials.

This study has several limitations. It is possible that the effects of neurofeedback are convolved in the active group with spontaneous symptom improvement. This is more likely in TS, in which symptom fluctuation over the course of days and weeks is not uncommon and which also often shows remission during adolescence (Leckman et al., 1998), than in OCD, which tends to be more chronic (Fineberg et al., 2013). In both cases, spontaneous symptom fluctuation should be adequately captured by the control group, and the significantly greater rate of improvement in the neurofeedback group provides evidence of a specific effect of the intervention. However, the sample size in this analysis remains modest, and replication is needed.

We emphasize that this study is reporting an omnibus analysis combining data across two ongoing clinical trials. These results should not be taken as evidence of efficacy of the intervention in either of the two ongoing clinical studies included in the analysis. Such a conclusion requires statistical examination of the effects separately in each intervention; these analyses will be reported once data collection is completed in both studies. Rather, our purpose in reporting these analyses on interim data combined across studies is to draw attention to a potentially common temporal pattern of response to neurofeedback. As discussed above, it is important to take this pattern of response into account in the design and interpretation of other neurofeedback studies.

Importantly, we do not assert that every neurofeedback intervention will induce a pattern of behavioral or perceptual change or of symptom improvement that increases over time after the intervention. There are published studies in the literature that do not exhibit this effect. For example, in a cross-over neurofeedback study that trained increases and decreases in perceptual confidence, confidence levels at the end of the first arm of the study showed little change a week later, at the start of second arm (Cortese et al., 2017). More work is needed to inform our understanding of the different temporal patterns of behavioural and symptom changes induced by different neurofeedback interventions. However, we maintain that the possibility of increasing behavioral change or symptom improvement after a neurofeedback intervention needs to be considered when designing neurofeedback protocols or interpreting data from neurofeedback studies.

Conclusion

Our data from two different clinical trials reveal a shared temporal pattern of clinical change after neurofeedback training. Symptoms neither regressed to baseline nor remained stable over a follow-up period of many weeks; instead, symptoms continued to improve over time. This result held even when the individual study was controlled for; thus, it appears to be shared across neurofeedback interventions that differ in design, have different target brain areas, and investigate different clinical populations. It is therefore likely that this pattern will be seen in other neurofeedback applications. This conclusion has important implications for the design, analysis, and interpretation of neurofeedback studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

Amano K, Shibata K, Kawato M, Sasaki Y, Watanabe T, 2016 Learning to Associate Orientation with Color in Early Visual Areas by Associative Decoded fMRI Neurofeedback. Curr Biol 26, 1861–1866. [PubMed: 27374335]

Bloch MH, Landeros-Weisenberger A, Rosario MC, Pittenger C, Leckman JF, 2008 Meta-analysis of the symptom structure of obsessive-compulsive disorder. Am J Psychiatry 165, 1532–1542. [PubMed: 18923068]

Caria A, Sitaram R, Veit R, Begliomini C, Birbaumer N, 2010 Volitional control of anterior insula activity modulates the response to aversive stimuli. A real-time functional magnetic resonance imaging study. Biological psychiatry 68, 425–432. [PubMed: 20570245]

Carroll KM, Rounsaville BJ, Gordon LT, Nich C, Jatlow P, Bisighini RM, Gawin FH, 1994 Psychotherapy and pharmacotherapy for ambulatory cocaine abusers. Arch Gen Psychiatry 51, 177–187. [PubMed: 8122955]

Cortese A, Amano K, Koizumi A, Kawato M, Lau H, 2016 Multivoxel neurofeedback selectively modulates confidence without changing perceptual performance. Nat Commun 7, 13669. [PubMed: 27976739]

Cortese A, Amano K, Koizumi A, Lau H, Kawato M, 2017 Decoded fMRI neurofeedback can induce bidirectional confidence changes within single participants. Neuroimage 149, 323–337. [PubMed: 28163140]

Cox WM, Subramanian L, Linden DE, Luhrs M, McNamara R, Playle R, Hood K, Watson G, Whittaker JR, Sakhuja R, Ihssen N, 2016 Neurofeedback training for alcohol dependence versus treatment as usual: study protocol for a randomized controlled trial. Trials 17, 480. [PubMed: 27716290]

deBettencourt MT, Cohen JD, Lee RF, Norman KA, Turk-Browne NB, 2015 Closed-loop training of attention with real-time brain imaging. Nature neuroscience 18, 470–475. [PubMed: 25664913]

deCharms RC, Maeda F, Glover GH, Ludlow D, Pauly JM, Soneji D, Gabrieli JDE, Mackey SC, 2005 Control over brain activation and pain learned by using real-time functional MRI. Proceedings of the National Academy of Sciences 102, 18626–18631.

Dudai Y, 2012 The restless engram: consolidations never end. Annu Rev Neurosci 35, 227–247. [PubMed: 22443508]

Engelbregt HJ, Keeser D, van Eijk L, Suiker EM, Eichhorn D, Karch S, Deijen JB, Pogarell O, 2016 Short and long-term effects of sham-controlled prefrontal EEG-neurofeedback training in healthy subjects. Clin Neurophysiol 127, 1931–1937. [PubMed: 26971473]

Fineberg NA, Hengartner MP, Bergbaum C, Gale T, Rossler W, Angst J, 2013 Remission of obsessive-compulsive disorders and syndromes; evidence from a prospective community cohort study over 30 years. Int J Psychiatry Clin Pract 17, 179–187. [PubMed: 23428237]

Freeman J, Garcia A, Frank H, Benito K, Conelea C, Walther M, Edmunds J, 2014 Evidence base update for psychosocial treatments for pediatric obsessive-compulsive disorder. J Clin Child Adolesc Psychol 43, 7–26. [PubMed: 23746138]

Gerin MI, Fichtenholtz H, Roy A, Walsh CJ, Krystal JH, Southwick S, Hampson M, 2016 Real-Time fMRI Neurofeedback with War Veterans with Chronic PTSD: A Feasibility Study. Front Psychiatry 7, 111. [PubMed: 27445868]

Gevensleben H, Moll GH, Rothenberger A, Heinrich H, 2014 Neurofeedback in attention-deficit/ hyperactivity disorder - different models, different ways of application. Front Hum Neurosci 8, 846. [PubMed: 25374528]

Goldstein MG, Niaura R, Follick MJ, Abrams DB, 1989 Effects of behavioral skills training and schedule of nicotine gum administration on smoking cessation. Am J Psychiatry 146, 56–60. [PubMed: 2643357]

Goodman WK, Price LH, Rasmussen SA, Mazure C, Delgado P, Heninger GR, Charney DS, 1989a The Yale-Brown Obsessive Compulsive Scale. II. Validity. Archives of general psychiatry 46, 1012–1016. [PubMed: 2510699]

Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, Heninger GR, Charney DS, 1989b The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. Archives of general psychiatry 46, 1006–1011. [PubMed: 2684084]

Grone M, Dyck M, Koush Y, Bergert S, Mathiak KA, Alawi EM, Elliott M, Mathiak K, 2015 Upregulation of the rostral anterior cingulate cortex can alter the perception of emotions: fMRI-based neurofeedback at 3 and 7 T. Brain Topogr 28, 197–207. [PubMed: 25087073]

Hampson M, Scheinost D, Qiu M, Bhawnani J, Lacadie CM, Leckman JF, Constable RT, Papademetris X, 2011 Biofeedback of real-time functional magnetic resonance imaging data from the supplementary motor area reduces functional connectivity to subcortical regions. Brain Connectivity 1, 91–98. [PubMed: 22432958]

Harmelech T, Preminger S, Wertman E, Malach R, 2013 The day-after effect: long term, Hebbian-like restructuring of resting-state fMRI patterns induced by a single epoch of cortical activation. J Neurosci 33, 9488–9497. [PubMed: 23719815]

Hebb DO, 1949 The organization of behavior; a neuropsychological theory. Wiley, New York,.

Kandel ER, Dudai Y, Mayford MR, 2014 The molecular and systems biology of memory. Cell 157, 163–186. [PubMed: 24679534]

Kotchoubey B, Blankenhorn V, Froscher W, Strehl U, Birbaumer N, 1997 Stability of cortical self-regulation in epilepsy patients. Neuroreport 8, 1867–1870. [PubMed: 9223067]

Koush Y, Meskaldji DE, Pichon S, Rey G, Rieger SW, Linden DE, Van De Ville D, Vuilleumier P, Scharnowski F, 2017 Learning Control Over Emotion Networks Through Connectivity-Based Neurofeedback. Cereb Cortex 27, 1193–1202. [PubMed: 26679192]

Leckman JF, Riddle MA, Hardin MT, Ort SI, Swartz KL, Stevenson J, Cohen DJ, 1989 The Yale Global Tic Severity Scale: initial testing of a clinician-rated scale of tic severity. J Am Acad Child Adolesc Psychiatry 28, 566–573. [PubMed: 2768151]

Leckman JF, Zhang H, Vitale A, Lahnin F, Lynch K, Bondi C, Kim YS, Peterson BS, 1998 Course of tic severity in Tourette syndrome: the first two decades. Pediatrics 102, 14–19. [PubMed: 9651407]

Linden DE, Habes I, Johnston SJ, Linden S, Tatineni R, Subramanian L, Sorger B, Healy D, Goebel R, 2012 Real-time self-regulation of emotion networks in patients with depression. PLoS One 7, e38115. [PubMed: 22675513]

Lowel S, Singer W, 1992 Selection of intrinsic horizontal connections in the visual cortex by correlated neuronal activity. Science 255, 209–212. [PubMed: 1372754]

Megumi F, Yamashita A, Kawato M, Imamizu H, 2015 Functional MRI neurofeedback training on connectivity between two regions induces long-lasting changes in intrinsic functional network. Front Hum Neurosci 9, 160. [PubMed: 25870552]

Paret C, Kluetsch R, Zaehringer J, Ruf M, Demirakca T, Bohus M, Ende G, Schmahl C, 2016 Alterations of amygdala-prefrontal connectivity with real-time fMRI neurofeedback in BPD patients. Soc Cogn Affect Neurosci 11, 952–960. [PubMed: 26833918]

Piacentini J, Bergman RL, Chang S, Langley A, Peris T, Wood JJ, McCracken J, 2011 Controlled comparison of family cognitive behavioral therapy and psychoeducation/relaxation training for child obsessive-compulsive disorder. J Am Acad Child Adolesc Psychiatry 50, 1149–1161. [PubMed: 22024003]

Piacentini J, Woods DW, Scahill L, Wilhelm S, Peterson AL, Chang S, Ginsburg GS, Deckersbach T, Dziura J, Levi-Pearl S, Walkup JT, 2010 Behavior therapy for children with Tourette disorder: a randomized controlled trial. JAMA 303, 1929–1937. [PubMed: 20483969]

Ramot M, Kimmich S, Gonzalez-Castillo J, Roopchansingh V, Popal H, White E, Gotts SJ, Martin A, 2017 Direct modulation of aberrant brain network connectivity through real-time NeuroFeedback. Elife 6.

Robineau F, Meskaldji DE, Koush Y, Rieger SW, Mermoud C, Morgenthaler S, Van De Ville D, Vuilleumier P, Scharnowski F, 2017 Maintenance of Voluntary Self-regulation Learned through Real-Time fMRI Neurofeedback. Front Hum Neurosci 11, 131. [PubMed: 28386224]

Rota G, Sitaram R, Veit R, Erb M, Weiskopf N, Dogil G, Birbaumer N, 2009 Self-regulation of regional cortical activity using real-time fMRI: the right inferior frontal gyrus and linguistic processing. Hum Brain Mapp 30, 1605–1614. [PubMed: 18661503]

Scharnowski F, Veit R, Zopf R, Studer P, Bock S, Diedrichsen J, Goebel R, Mathiak K, Birbaumer N, Weiskopf N, 2015 Manipulating motor performance and memory through real-time fMRI neurofeedback. Biol Psychol 108, 85–97. [PubMed: 25796342]

Scheinost D, Stoica T, Saksa J, Papademetris X, Constable RT, Pittenger C, Hampson M, 2013 Orbitofrontal cortex neurofeedback produces lasting changes in contamination anxiety and resting-state connectivity. Translational psychiatry 3, e250. [PubMed: 23632454]

Schnyer DM, Beevers CG, deBettencourt MT, Sherman SM, Cohen JD, Norman KA, Turk-Browne NB, 2015 Neurocognitive therapeutics: from concept to application in the treatment of negative attention bias. Biol Mood Anxiety Disord 5, 1. [PubMed: 25905002]

Sitaram R, Ros T, Stoeckel L, Haller S, Scharnowski F, Lewis-Peacock J, Weiskopf N, Blefari ML, Rana M, Oblak E, Birbaumer N, Sulzer J, 2017 Closed-loop brain training: the science of neurofeedback. Nat Rev Neurosci 18, 86–100. [PubMed: 28003656]

Subramanian L, Hindle JV, Johnston S, Roberts MV, Husain M, Goebel R, Linden D, 2011 Real-time functional magnetic resonance imaging neurofeedback for treatment of Parkinson's disease. The Journal of neuroscience : the official journal of the Society for Neuroscience 31, 16309–16317.

Surmeli T, Ertem A, 2011 Obsessive compulsive disorder and the efficacy of qEEG-guided neurofeedback treatment: a case series. Clin EEG Neurosci 42, 195–201. [PubMed: 21870473]

Surmeli T, Ertem A, Eralp E, Kos IH, 2012 Schizophrenia and the efficacy of qEEG-guided neurofeedback treatment: a clinical case series. Clin EEG Neurosci 43, 133–144. [PubMed: 22715481]

Weiskopf N, Veit R, Erb M, Mathiak K, Grodd W, Goebel R, Birbaumer N, 2003 Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging (fMRI): methodology and exemplary data. NeuroImage 19, 577–586. [PubMed: 12880789]

Wilhelm S, Peterson AL, Piacentini J, Woods DW, Deckersbach T, Sukhodolsky DG, Chang S, Liu H, Dziura J, Walkup JT, Scahill L, 2012 Randomized trial of behavior therapy for adults with Tourette syndrome. Arch Gen Psychiatry 69, 795–803. [PubMed: 22868933]

Yoo SS, Lee JH, O'Leary H, Lee V, Choo SE, Jolesz FA, 2007 Functional magnetic resonance imaging-mediated learning of increased activity in auditory areas. Neuroreport 18, 1915–1920. [PubMed: 18007186]

Yoo SS, Lee JH, O'Leary H, Panych LP, Jolesz FA, 2008 Neurofeedback fMRI-mediated learning and consolidation of regional brain activation during motor imagery. Int J Imaging Syst Technol 18, 69–78. [PubMed: 19526048]

Young KD, Siegle GJ, Zotev V, Phillips R, Misaki M, Yuan H, Drevets WC, Bodurka J, 2017 Randomized Clinical Trial of Real-Time fMRI Amygdala Neurofeedback for Major Depressive Disorder: Effects on Symptoms and Autobiographical Memory Recall. Am J Psychiatry, appiajp201716060637.

Yuan H, Young KD, Phillips R, Zotev V, Misaki M, Bodurka J, 2014 Resting-state functional connectivity modulation and sustained changes after real-time functional magnetic resonance imaging neurofeedback training in depression. Brain Connect 4, 690–701. [PubMed: 25329241]

Zhang G, Yao L, Zhang H, Long Z, Zhao X, 2013a Improved working memory performance through self-regulation of dorsal lateral prefrontal cortex activation using real-time fMRI. PLoS One 8, e73735. [PubMed: 24013130]

Zhang G, Zhang H, Li X, Zhao X, Yao L, Long Z, 2013b Functional alteration of the DMN by learned regulation of the PCC using real-time fMRI. IEEE Trans Neural Syst Rehabil Eng 21, 595–606. [PubMed: 23070370]

Zilverstand A, Sorger B, Sarkheil P, Goebel R, 2015 fMRI neurofeedback facilitates anxiety regulation in females with spider phobia. Front Behav Neurosci 9, 148. [PubMed: 26106309]

Highlights

- Temporal pattern of symptom change following neurofeedback shared across studies

- Symptoms continued to improve for weeks after neurofeedback

- Neurofeedback studies should follow up subjects to maximize power

- Crossover designs may be contaminated by significant carryover effects

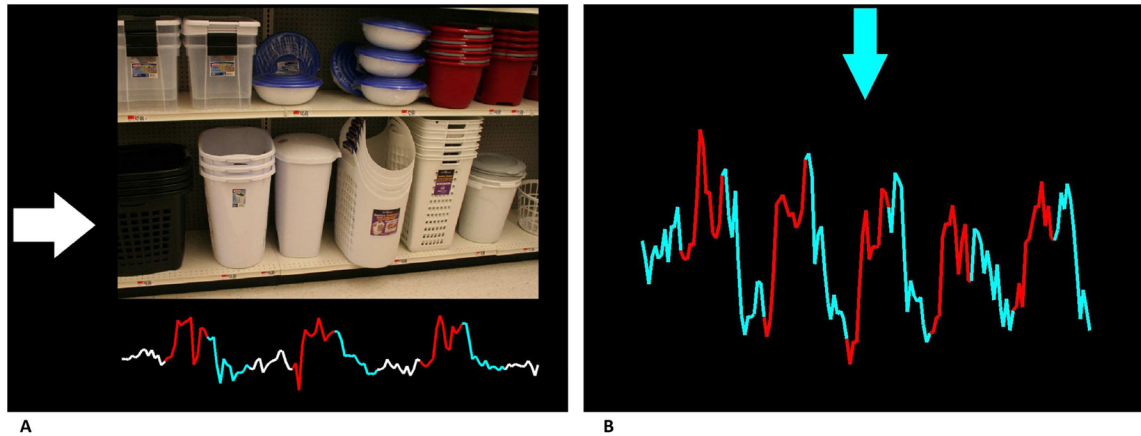- Optimizing number of sessions by embedding assessments not recommended

Figure 1.
Panel (A) Example feedback screen in the OCD study. The arrow pointing to the right (white) indicates a resting block with a neutral picture shown and corresponds with white segments of the feedback time course (below). As pictures changed to provocative, a red arrow pointing up indicated an upregulation phase and a blue arrow pointing down a downregulation phase. The feedback time course underneath the picture changed colors accordingly. Panel (B) Example feedback screen in TS study. The blue arrow at the top of the screen indicated a downregulation phase and would alternate with a red arrow pointing up during upregulation phases. The colors of the feedback time course corresponded with the regulation phase color.
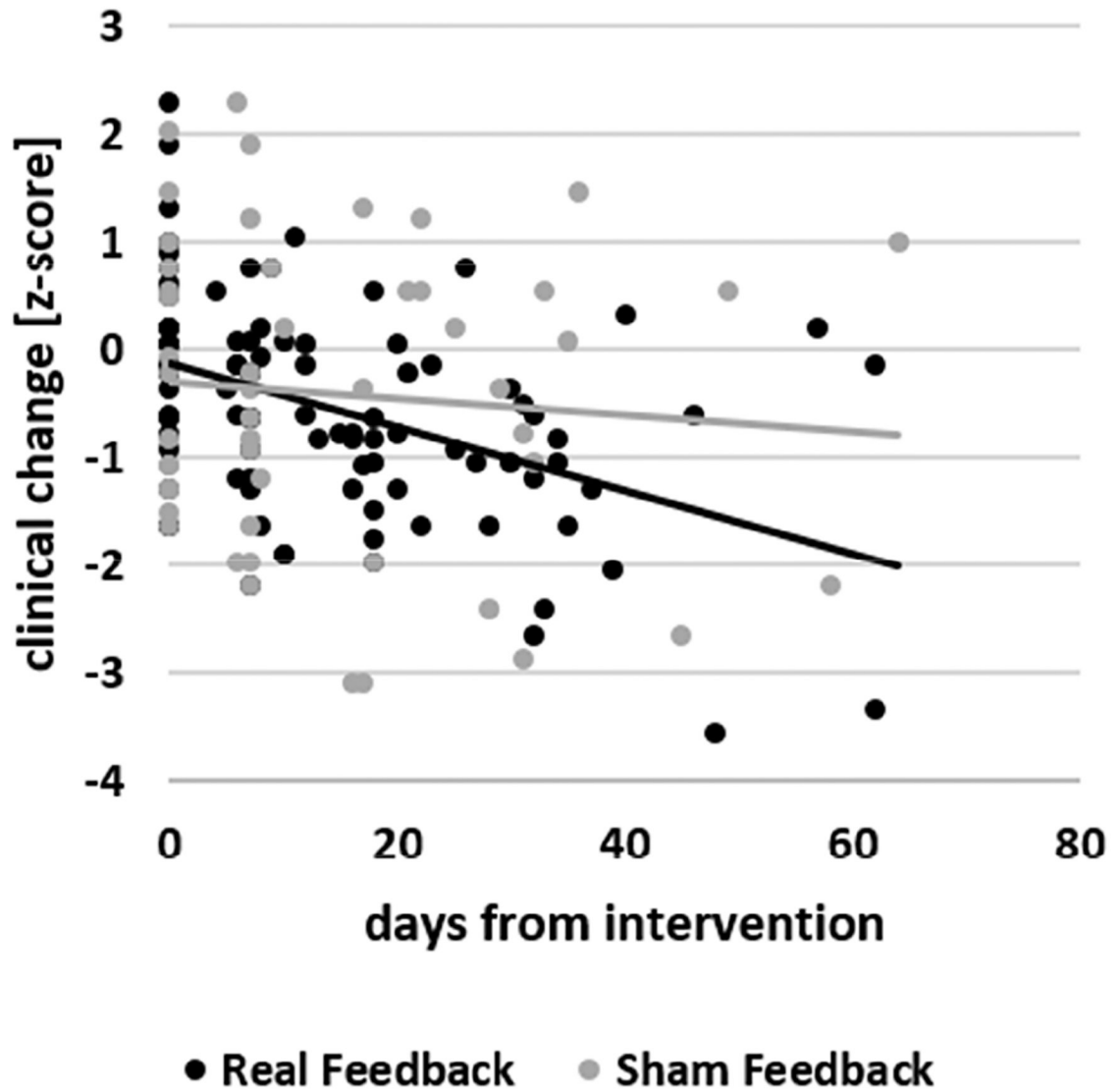
Figure 2.
Scatterplot of clinical ratings (z-standardized) against latency (days after the intervention) of the neurofeedback (black) and sham (gray) intervention groups and the corresponding best fit line for each group.
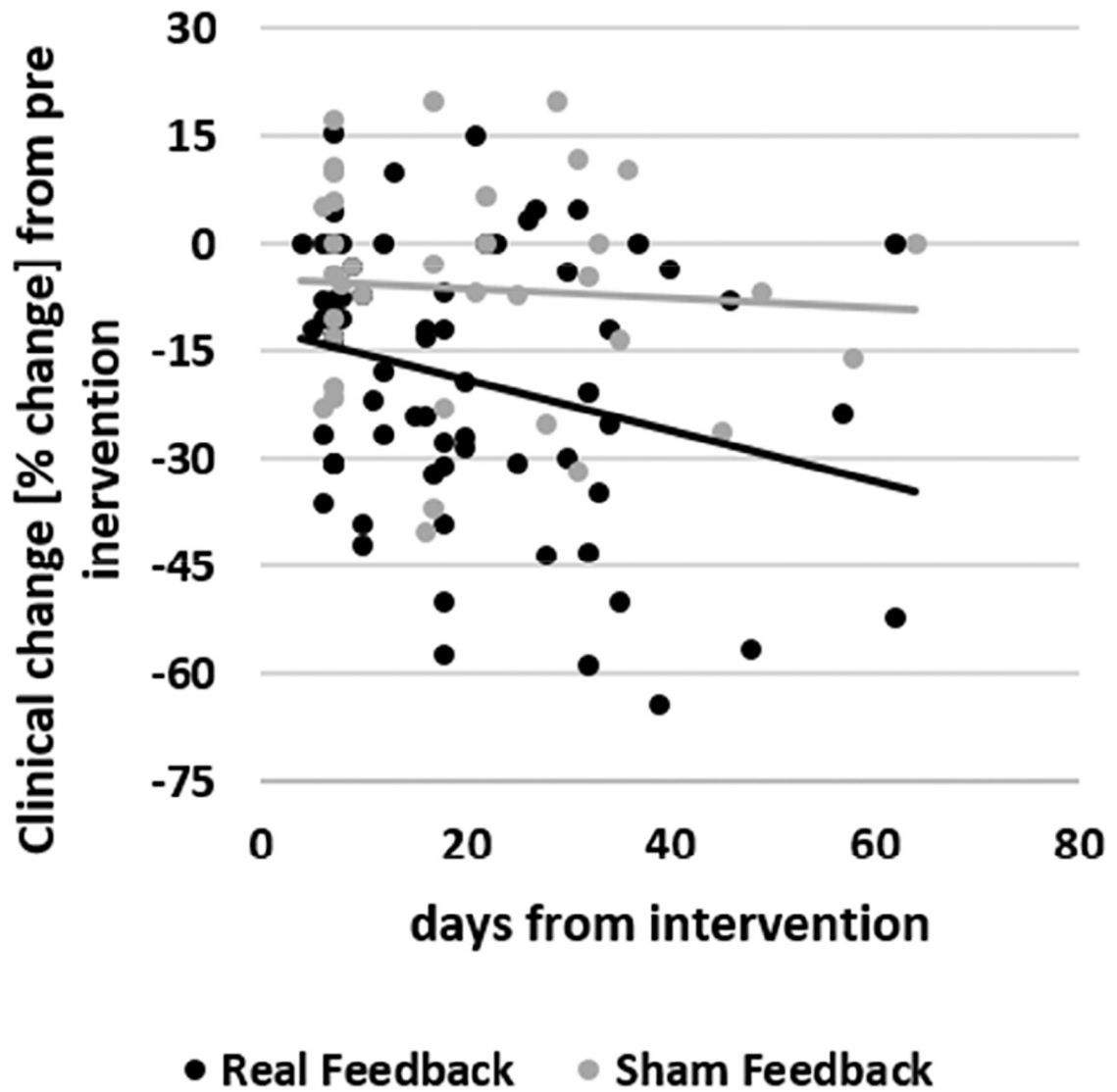
Figure 3.
Scatterplot of the clinical ratings (%-change from pre-intervention) against latency (days from intervention) of the neurofeedback (black) and sham (gray) intervention groups and the corresponding best fit line for each group.