Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance

Benedikt Zoefel a, René J. Huster b, Christoph S. Herrmann b,

a Otto-Von-Guericke University, Institute for Biology, Magdeburg, Germany
b Carl-Von-Ossietzky University Oldenburg, Department of Psychology, Experimental Psychology Lab, Oldenburg, Germany

A R T I C L E   I N F O
Article history:
Received 16 February 2010
Revised 24 August 2010
Accepted 31 August 2010
Available online xxxx

Keywords:
Neurofeedback
EEG
Upper alpha
Trainability
Cognitive performance
Mental rotation

A B S T R A C T
In this study, the individually determined upper alpha frequency band in EEG (electroencephalogram) was investigated as a neurofeedback parameter. Fourteen subjects were trained on five sessions within 1 week by means of feedback dependent on the current upper alpha amplitude. On the first and fifth session, cognitive ability was tested by a mental rotation test. As a result, eleven of the fourteen subjects showed significant behavioral improvements in the neurofeedback group compared to the control group who did not receive feedback. Thus, enhanced cognitive control went along with an increased upper alpha amplitude that was found in the neurofeedback group only.

Introduction

Recently, the significant role of oscillations for brain functions and behavior as well as for psychiatric diseases became increasingly obvious (Basar and Güntekin, 2008; Basar et al., 2000; Herrmann and Knight, 2001; Herrmann et al., 2004; Strüber and Herrmann, 2002; Uhlhaas and Singer, 2006). Neurofeedback training (NFT) as an operant conditioning method to control oneself's brain activity has been shown to be an appropriate way to control or change these oscillations. In addition to clinical treatment for attention-deficit hyperactivity disorder (Lubar, 2003), epilepsy (Seman, 2000) and other disorders (Saxby and Peniston, 1995; Gruzelier et al., 1999; Hardt and Kamiya, 1978), NFT is applied to train locked-in patients to communicate (Birbaumer et al., 1999). It even seems to have effects on certain performance measures like semantic working memory (Vernon et al., 2003) and mental rotation ability (Hanslmayr et al., 2005; see Vernon, 2005, for an overview).

But the application of NFT has also received criticism concerning the reliability of its effects. Egner et al. (2004) addressed one main problem that—despite of significant behavioral impact—sometimes no spectral effects can be found after NFT. In that study, when effects with regard to the EEG were observed at all, they often were unreliable or did not meet expectations concerning the frequency spectrum or the topography (see also the tables in Vernon, 2005).

We propose three criteria for the validation of a neurofeedback parameter: There should be spectral effects within the trained frequency band caused by the training (trainability). These spectral changes should not affect other frequency bands (independence). Finally, it is reasonable to choose a frequency band that is associated with certain cognitive functions to increase the probability of reliable behavioral effects as well as applicability (interpretability).

Many of the existing studies on NFT parameters do not satisfy all three criteria. The alpha band (8–12.5 Hz), for example, has been shown to be increased after NFT (Bauer, 1976), but for the frequency range tested there, no cognitive changes could be demonstrated. In a recently published article, Cho et al. (2008) reported an increased amplitude in the same frequency band, but possible effects on cognitive performance were not examined. Other studies concerning the impact of non-clinical NFT of SMR (sensory-motor rhythm, 12–15 Hz) on EEG and performance either found no effects or found effects but at unexpected frequencies or electrodes (e.g., Vernon, 2005; Egner et al., 2004).

Constraining alpha to the individually determined upper alpha band (ranging from the individual alpha frequency, IAF, to IAF+2 Hz) might cause an improvement of trainability (Hanslmayr et al., 2005). However, this was shown only within 1 day of NFT—a possible long-term effect was not examined. There is evidence of at least two independent (lower and upper) alpha subbands (Angelakis and Lubar, 2002; Klimesch et al., 1997; Michels et al., 2008). Finally, upper alpha is widely shown to be correlated with cognitive performance (for a review, see Klimesch, 1999), indicating interpretability. For instance, prestimulus activity in the UA band has been shown to be positively related to performance in mental rotation tasks (Hanslmayr et al., 2005; Klimesch et al., 2003).

* Corresponding author. Fax: +49 441 798 3865.
E-mail address: christoph.herrmann@uni-oldenburg.de (C.S. Herrmann).
Despite these promising results, the usability of upper alpha (UA) for NFT and its spectral as well as cognitive impact for longer periods of time are hardly investigated yet. Thus, in the current study the trainability of the UA band and its impact on cognitive abilities were examined. Our hypotheses were (i) an increase of the UA amplitude, (ii) being independent of other frequency bands, and (iii) being related to improved performance.

Methods

Design

For each subject, the experiment consisted of five sessions within the same week from Monday to Friday with one training session each day. Within one session, the structure was the following (Fig. 1A):

At first, a short EOG (electrooculogram) calibration measurement (not depicted) was conducted for subsequent detection of eye-movement artefacts. Then, a base rate of 5 min was recorded. Afterwards, five training blocks of 5 min each, followed by a second base rate measurement, were acquired. On Monday and Friday, a mental rotation test was given after the first base rate to test cognitive abilities. To evaluate the influence of practice on this performance, a control group was tested with the same design on equivalent days and times, except for the training blocks. Control participants did not have to show up on days two to four, but were asked not to expose themselves to exceptional stress.

Participants

A total of 24 students, fourteen in the NFT group and ten in the control group, took part in the experiment. They gave written informed consent and received monetary reward for participation. Because of non-compliant behavior, two subjects of the NFT group had to be excluded from further statistical analyses. These subjects admittedly hastened through the second cognitive performance assessment leading to strongly deteriorated scores outside of the 95% confidence interval of the NFT group. The final sample consisted of 12 subjects in the NFT (nine females and five males, 23.7 ± 2.3 years) and ten in the control group (five females and five males, 22.1 ± 3.8 years).

EEG recordings

During the experiment, the participants sat in an electrically shielded and sound attenuated room. All devices inside the cabin were battery operated to avoid line frequency interference (50 Hz in Germany). EEG was measured from 32 Ag/AgCl electrodes, placed in an elastic cap (EasyCap, Falk Minow Services, Munich) according to the international 10-20 system, with P3, Pz, P4, O1 and O2 used for feedback. The signals were amplified by a BrainAmp 32-channel system (BrainAmp, Brain Products GmbH, Munich), analog filtered between 0.01 and 250 Hz and digitally stored at a sampling frequency of 1000 Hz. A nose reference was used, the ground electrode was located at FCz and EOG was recorded by an electrode below the right eye. Impedances of feedback electrodes were kept below 5 kΩ.

The EEG signals were read out from the EEG and further processed using an in-home software programmed in C++. EEG power was calculated by means of a sliding FFT algorithm including a hamming window (1024 sampling points), updated each second during the base rates and every 100 ms during the training blocks. This calculation of frequency-specific EEG-power was used in training blocks to provide a fast and reliable feedback, whereas during the base rates it was recorded for offline statistical analyses.

Stimuli and experimental procedure

In each session, the IAF was calculated as the peak frequency of the alpha band during the first base rate and UA was defined as the frequency band from IAF to IAF + 2 Hz.

During the training blocks, the feedback was given by means of a colored square. The subjects received feedback according to the average of the UA amplitudes measured at the five feedback sites. The saturation of the feedback color indicated the UA amplitude in relation to the base rate. Red and blue values symbolized an amplitude above and below the base rate, respectively. The saturation scale covered 95% of the amplitude range. Values above 97.5% or below 2.5% were indicated by maximal red or blue saturation, respectively. If no difference from baseline was present, the square was grey.

An amplitude value was declared invalid, if the spectral power of eye movements leaked into the UA band. This, for example, is the case when the UA power of the EOG channel exceeds the initially determined change of UA power for eye-movement artefacts. To assure the comparability of the base rate and the training measurements, a square randomly changing its color from grey to red or blue was presented during recording of base rates. The same color saturations as in the training blocks were used. In order to challenge subjects cognitively, they were asked to count the red gradients to avoid drowsiness and confirm comparability. Additionally, this task could assure that the subjects did not close their eyes during the base rates. With the training blocks, they were informed that eye closure is not a valid strategy for UA regulation. Closing eyes obviously would have prevented the subjects to benefit from visually presented feedback, making this strategy meaningless.

Cognitive performance was assessed by the mental rotation test A3DW (Adaptiver dreidimensionaler Wuerfeltest; Gittler, 2007), based on computerized adaptive testing (the number of items is chosen individually for each subject) and the Rasch-model (Rasch, 1980). In each trial, a reference cube and six other cubes were shown, with at most one of them being convertible into the reference cube. If this...
cube existed, it had to be selected. An extended form was used in order to obtain reliable measurements.

Statistical analyses

For each subject, the average UA amplitude was calculated for the base rates and training blocks of each session. Paired, one-tailed t-tests were used for the following statistical analyses.

“Responding” to UA NFT was defined as showing a mean UA amplitude in the training blocks of the last session, that was significantly higher than the first base rate of session one. As a result, three subjects of the NFT group did not respond to the UA training and were excluded from further analyses.

In order to evaluate trainability, the first base rate of the first session and the first base rate of the last session were tested for differences. In order to evaluate independence, the frequency spectrum before the experiment and the spectrum during the very last training block were compared (first base rate of the first session versus training block five of the last session, n = 11). Amplitude differences between the spectra were tested for UA as well as for the frequency range below (lower alpha, tested for IAF-3 to IAF-1 Hz) and above (lower beta, tested for IAF+3 to IAF+5 Hz) UA. A range of 1 Hz below and above UA could not be tested due to frequency smearing (cf. Fig. 2).

For comparison of the cognitive performance, the raw scores of the mental rotation test (x = 100, SD = 10) were utilized, using a one-tailed t-test for independent samples. Both the NFT group and the control group would probably show increased performance due to the positive slope of the UA amplitudes over the course of the training. The performance measures of the control group for the two mental rotation tests were 108.8 ± 11.38 and 114.7 ± 19.08. The performance measures of the NFT group were 116.9 ± 11.50 and 129.7 ± 11.63. The mean increase of mental rotation test performance in the control group was 5.9 ± 11.48. In the NFT group the mean increase was 12.8 ± 7.98. The performance was significantly increased for the NFT group (t(16) = 2.21, p = 0.029), but not the control group (n.s.).

For the control group, the UA before the second cognitive test was not significantly different from the UA before the first cognitive test. For the NFT group, as it was already mentioned, there was a significantly higher UA amplitude in the first base rate of the last session (before the second cognitive test) than in the first base rate of the first session (before the first cognitive test). Thus, the NFT group, who did show UA enhancement, also showed better cognitive performance, while the control group lacked both UA enhancement and pronounced differences in cognitive performance.

Results

Trainability

The average IAF was 9.2 ± 0.91 Hz during the first base rate and 9.16 ± 0.76 Hz during the second base rate. No significant changes were observed in the course of the training.

In Fig. 1B, the average UA amplitudes across all responders are shown as a function of the temporal course of the study. Trainability is reflected by the positive slope of the UA amplitudes over the course of time. The gradients of a fitted regression line for each subject (n = 11) are significantly larger than zero (t(10) = 3.18, p = 0.010).

The expectation of an increase of UA was confirmed: In the last session, the UA amplitude during the first base rate showed a significantly higher amplitude than the first base rate of the first session (t(10) = 3.59, p = 0.003). This effect was not found in the control group (n.s.). The increase of UA is also visible in the raw EEG data, as shown in Fig. 3. For a selected subject, five representative seconds of the EEG during the first and last session are depicted, showing a clear change towards a pronounced EEG alpharhythm. In addition, maps averaged across NFT responders demonstrate the topographical stability of the training-induced UA enhancements (Fig. 3).

Independence

To examine whether UA was trainable independently of a change in other frequency bands, the frequency spectrum before the experiment (first base rate of the first session) and during the very last training block is plotted in Fig. 2. The just mentioned increase of UA amplitude is visible as the difference between the spectra. The effect was significant in the trained UA range between IAF and IAF + 2 Hz (t(10) = 2.39, p = 0.019). The other frequency bands were not affected significantly. These results reflect a regulation of UA, that was independent of other frequency bands.

Cognitive performance

On average, 18 ± 2.8 items were presented in the control group with an average presentation time for a single item of 56.6 ± 33.7. In the NFT group, 16.7 ± 3.5 items were presented, each for approximately 66.9 ± 23.1 s. The performance measures of the control group for the two mental rotation tests were 108.8 ± 11.38 and 114.7 ± 19.08. The performance measures of the NFT group were 116.9 ± 11.50 and 129.7 ± 11.63. The mean increase of mental rotation test performance in the control group was 5.9 ± 11.48. In the NFT group the mean increase was 12.8 ± 7.98. The performance was significantly increased for the NFT group (t(16) = 2.21, p = 0.029), but not the control group (n.s.).

For the control group, the UA before the second cognitive test was not significantly different from the UA before the first cognitive test. For the NFT group, as it was already mentioned, there was a significantly higher UA amplitude in the first base rate of the last session (before the second cognitive test) than in the first base rate of the first session (before the first cognitive test). Thus, the NFT group, who did show UA enhancement, also showed better cognitive performance, while the control group lacked both UA enhancement and pronounced differences in cognitive performance.

Discussion

In line with our hypotheses, the trainability of the UA band (hypothesis i) has been confirmed: All but three of the fourteen subjects of the NFT group showed significant training success. This is remarkable, since in therapy usually a ten-fold higher number of sessions is used (Lubar, 2003).

In the course of the week, a linear increase of the UA amplitude was visible (Fig. 1B). Also in between sessions the training effects remained present over an extended period of time. Even before training, during the first base rate, the UA amplitude in Session 5 was higher than in Session 1. This result is remarkable, because it suggests that each training session builds upon the training experiences of the previous days.

Furthermore, our data demonstrate that it is possible to control UA independently of other frequency bands (hypothesis ii; cf. Fig. 2). The frequency spectrum has been affected by UA NFT only in the trained UA band and was not significantly altered in the frequency range below and above UA.

Please cite this article as: Zoefel, B., et al., Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.08.078
These findings thus directly address the mentioned critique of Egner et al. (2004), stating that NFT is often unspecific with regard to affected frequency bands. In our study, NFT effects were restricted to the trained UA band, reflecting independence.

Given that the UA amplitudes were significantly augmented at the last when compared to the first session with the initial base rate measurements, our data thus confirm the results of trainability of UA from Hanslmayr et al. (2005) and extend them to a long-term aspect. Also, the findings of a long-term alpha enhancement reported from Cho et al. (2008) could be confirmed for UA with the applied NFT.

The expectation of an enhancement of cognitive performance was confirmed as well (hypothesis iii): the increase in scores of mental rotation was significantly larger for the NFT group than for the control group.

In addition, the UA of the NFT group was significantly higher before the second cognitive test than before the first. This was not the case for the control group, supporting a relationship of UA increase with cognitive performance. This finding is in line with previous studies (Doppelmayr et al., 2005; Fink et al., 2005; Klimesch et al., 1997; Klimesch, 1999) and supports the important interpretability of the NFT parameter.

Because of the frequently criticized lack of a control group in NFT studies (e.g., Gevensleben et al., 2009; Gruzelier and Egner, 2005) which makes it impossible to evaluate the impact of practice or social effects, a control group was assessed in this study. However, although comparability has been tried to ensure for most factors, the subjects of the control group had to show up just on 2 days. Three days of difference concerning the spectrum of the impact. These findings could also explain the large varieties of training success across studies. Most of the less successful responders in our study did not show a clear alpha peak, but rather a 1/f-distribution of the frequency spectrum. Gevensleben et al. (2009) as well as Gruzelier and Egner (2005) outlined the question of individual features predicting or at least influencing success of NFT. Personality factors, the alpha distribution or the individual impact of cognitions on the EEG frequency spectrum could be such factors for UA. Examining these dependencies and establishing some kind of individual instance just differing in the frequency band used for training, would be desirable for future studies.

The presented findings support UA as a promising NFT parameter that is worth exploring and investigating further. One design that is worthwhile considering would be a lower/upper alpha-feedback with the aim of minimizing the lower/upper alpha ratio. While theta/alpha, theta/beta or theta/SMR-feedback is already established in therapy (Gevensleben et al., 2009; Lubar, 1991; Moore, 2000; Sterman, 2000), the lower/upper alpha ratio is not investigated yet. It has been found that the effects on cognitive performance of lower alpha may be contrary to UA (Klimesch, 1999; Klimesch et al., 1997). Thus, the additional suppression of lower alpha may at the same time enhance independence and cognitive performance.

The precise functionality of NFT still remains unclear. Lubar (1997) proposes an explanation of the NFT mechanism as some kind of “dynamic resonant loop,” effecting other cortical areas. He also states “optimal coupling levels” (seen as oscillation frequency) for “any particular state.” UA may reflect the resonance frequency of such a loop for the state of an optimal cognitive performance. Most subjects reported “evoking emotions” as the best strategy to increase UA. Most interestingly, Fink (2005) as well as Güntekin and Basar (2007) showed an impact on UA after the presentation of emotional faces, with the former reporting a modulation by intro- and extraversion and the latter of intra- and interindividual differences concerning the spectrum of the impact. These findings could also explain the large varieties of training success across studies. Most of the less successful responders in our study did not show a clear alpha peak, but rather a 1/f-distribution of the frequency spectrum. Gevensleben et al. (2009) as well as Gruzelier and Egner (2005) outlined the question of individual features predicting or at least influencing success of NFT. Personality factors, the alpha distribution or the individual impact of cognitions on the EEG frequency spectrum could be such factors for UA. Examining these dependencies and establishing some kind of individual
protocol for the respective NFT parameters would be worth thinking about for future studies.

Conclusions

Our study revealed promising results for the usage of individually determined UA as a NFT frequency parameter, fulfilling the criteria of trainability, independence, and interpretability, thus inspiring further examinations of the interactions and relations between UA, NFT and cognitive performance.

References