doi:10.1093/brain/awx011 BRAIN 2017: 140; 1041–1052 | 1041

BRAIN
A JOURNAL OF NEUROLOGY

Better than sham? A double-blind placebo-controlled neurofeedback study in primary insomnia

Manuel Schabus,[1,2] Hermann Griessenberger,[1] Maria-Teresa Gnjezda,[1] Dominik P. J. Heib,[1,2] Malgorzata Wislowska[1,2] and Kerstin Hoedlmoser[1,2]

See Thibault *et al.* (doi:10.1093/awx033) for a scientific commentary on this article.

Neurofeedback training builds upon the simple concept of instrumental conditioning, i.e. behaviour that is rewarded is more likely to reoccur, an effect Thorndike referred to as the 'law of effect'. In the case of neurofeedback, information about specific electroencephalographic activity is fed back to the participant who is rewarded whenever the desired electroencephalography pattern is generated. If some kind of hyperarousal needs to be addressed, the neurofeedback community considers sensorimotor rhythm neurofeedback as the gold standard. Earlier treatment approaches using sensorimotor-rhythm neurofeedback indicated that training to increase 12–15 Hz sensorimotor rhythm over the sensorimotor cortex during wakefulness could reduce attention-deficit/hyperactivity disorder and epilepsy symptoms and even improve sleep quality by enhancing sleep spindle activity (lying in the same frequency range). In the present study we sought to critically test whether earlier findings on the positive effect of sensorimotor rhythm neurofeedback on sleep quality and memory could also be replicated in a double-blind placebo-controlled study on 25 patients with insomnia. Patients spent nine polysomnography nights and 12 sessions of neurofeedback and 12 sessions of placebo-feedback training (sham) in our laboratory. Crucially, we found both neurofeedback and placebo feedback to be equally effective as reflected in subjective measures of sleep complaints suggesting that the observed improvements were due to unspecific factors such as experiencing trust and receiving care and empathy from experimenters. In addition, these improvements were not reflected in objective electroencephalographic-derived measures of sleep quality. Furthermore, objective electroencephalographic measures that potentially reflected mechanisms underlying the efficacy of neurofeedback such as spectral electroencephalographic measures and sleep spindle parameters remained unchanged following 12 training sessions. A stratification into 'true' insomnia patients and 'insomnia misperceivers' (subjective, but no objective sleep problems) did not alter the results. Based on this comprehensive and well-controlled study, we conclude that for the treatment of primary insomnia, neurofeedback does not have a specific efficacy beyond unspecific placebo effects. Importantly, we do not find an advantage of neurofeedback over placebo feedback, therefore it cannot be recommended as an alternative to cognitive behavioural therapy for insomnia, the current (non-pharmacological) standard-of-care treatment. In addition, our study may foster a critical discussion that generally questions the effectiveness of neurofeedback, and emphasizes the importance of demonstrating neurofeedback efficacy in other study samples and disorders using truly placebo and double-blind controlled trials.

1 Laboratory for Sleep, Cognition and Consciousness Research, Department of Psychology, University of Salzburg, Austria
2 Centre for Cognitive Neuroscience Salzburg (CCNS), University of Salzburg, Austria

Correspondence to: Manuel Schabus, PhD,
University of Salzburg,
Hellbrunnerstr. 34, 5020 Salzburg,
Austria
E-mail: Manuel.Schabus@sbg.ac.at

Received July 22, 2016. Revised December 1, 2016. Accepted December 6, 2016. Advance Access publication February 23, 2017

Keywords: neurofeedback; SMR; insomnia; placebo; neurotherapy

Abbreviations: CBT-I = cognitive behavioural therapy for insomnia; NFT = neurofeedback training; PFT = placebo feedback training (or randomized frequency sham protocol); PSQI = Pittsburgh Sleep Quality Index Questionnaire; SMR = 12–15 Hz sensorimotor rhythm

Introduction

Neurofeedback builds upon the simple principle of instrumental conditioning, that is, that behaviour which is rewarded is more likely to reoccur in the future, an effect also referred to as the 'law of effect' (Thorndike, 1905). Specifically, in the case of neurofeedback information about neural processes that are beyond wilful control—such as the generation of specific EEG activity patterns—is fed back to the participant and rewarded in the form of auditory or visual cues or tokens. Various neurofeedback training (NFT) protocols have been proposed and studied for the treatment of a wide range of disorders (for review see Hammond, 2011). Many of these protocols focus on sensorimotor rhythm (SMR) training, where 12–15 Hz at EEG sites above the sensorimotor cortex (i.e. electrode positions C3 or C4) is rewarded. The rationale is that enhanced SMR power will go hand-in-hand with increased relaxation or inhibition of (motor) activity (Sterman, 2000), which should counteract the hyperarousal linked to some prevalent disorders. Specifically, SMR-NFT has been readily used for the treatment of epilepsy (Tan *et al.*, 2009) and ADHD (Arns *et al.*, 2014) with usually clear beneficial outcomes. Besides this, Sterman *et al.* (1970) was able to show in his pioneering work that when cats are trained to increase EEG power in the (12–14 Hz) SMR frequency band during wakefulness, they also presented with more sleep spindles and enhanced sleep quality (i.e. less fragmentation) during subsequent sleep. This is intriguing not only because it suggests that NFT training effects may translate to other vigilance states (i.e. from wakefulness to sleep), but it importantly also suggests a possible treatment for insomnia, a burden that is estimated to affect between 10 and 35% of the general population worldwide (Morin *et al.*, 2006*b*). One mechanism that has been proposed to explain Sterman's findings is that the SMR frequency band significantly overlaps with the 12–15 Hz frequency range in which sleep spindles occur during non-REM sleep. Eventually, an increase in spindle activity may account for the improvements in sleep quality (i.e. shorter sleep onset latency and less sleep fragmentation) observed. However, despite the burden insomnia depicts for our society, only few studies followed-up on Sterman's findings (Hauri, 1981; Hauri *et al.*, 1982; Cortoos *et al.*, 2010) with all of them attesting neurofeedback beneficial effects on sleep.

Despite these promising findings, solid research on neurofeedback almost came to a standstill (for critical discussion see Thibault and Raz, 2016), and much of the knowledge we have about NFT today is derived from classic studies conducted in the 1980s. In the present study we aimed to resume and extend beyond those studies, and specifically build on our earlier findings (for review see Hoedlmoser *et al.*, 2011), which indicated positive effects of SMR-NFT in young healthy individuals (Hoedlmoser *et al.*, 2008) and young patients with subclinical insomnia (Schabus *et al.*, 2014). In particular, results indicated a beneficial effect of only 10 sessions of 12–15 Hz SMR-NFT training on sleep quality and memory performance (Hoedlmoser *et al.*, 2008) or even overnight memory consolidation (Schabus *et al.*, 2014). The latter finding is especially relevant from a clinical perspective as insomniacs frequently complain about problems related to attention and memory. However, the external validity of these findings and eventually the translation into clinical practice may be hampered by limitations such as a single blind study design and the lack of follow-up assessments. Besides this, earlier results obtained in a subclinical insomnia sample indicated a placebo effect with participants feeling better from visit to visit independent of whether they received NFT or placebo feedback training (PFT) (*cf.* Fig. 6 in Schabus *et al.*, 2014), a finding that is in line with conclusions drawn by Thibault *et al.* (2015) in a recent review. Importantly, in our PFT condition participants also received real EEG feedback; however, not about SMR, but varying frequency bands, i.e. each training session they received feedback on a different frequency band.

To circumvent the above mentioned limitations and to obtain a reliable and valid evaluation of the usefulness of SMR-NFT for the treatment of primary insomnia, we here adopted a counterbalanced double-blind cross-over design in a sample of patients with primary insomnia. As we here find that NFT positively affects subjective measures, yet is ineffective in changing objective parameters such as spectral EEG measures, sleep architecture or memory performance, we *post hoc* added an additional healthy NFT control group (undergoing identical SMR-NFT training) to replicate the effectiveness of our NFT protocol that had previously been established (Hoedlmoser *et al.*, 2008). This young healthy control group underwent the same procedure as did the clinical sample, but exclusively received NFT. The protocol for all patients comprised multiple meetings for psychometric testing, nine nights in the laboratory with polysomnography recording (i.e. one adaptation/screening as well as two pre- and two post-treatment nights flanking the PFT and NFT blocks), 12 PFT and 12 NFT sessions.

More specifically, the NFT training protocol was akin to the one used in previous studies from our group (Hoedlmoser *et al.*, 2008; Schabus *et al.*, 2014) with feedback consisting of information about SMR EEG power

above a central brain area (electrode C3) and no simultaneous block/inhibit filters to ensure that patients actually would be able to learn increasing their SMR rhythm within 12 sessions. We chose to only include 12 sessions of NFT as non-pharmacological treatment alternatives such as cognitive behavioural therapy for insomnia (CBT-I, Morin *et al.*, 2006*a*; Riemann *et al.*, 2015) are known to be successful with as few as four to eight sessions.

In summary, the aim was to replicate previous results and verify the clinical efficacy of SMR-NFT in a sample of primary insomnia patients adopting a double-blind study design.

Materials and methods

Ethics

The present study was conducted in accordance with the Declaration of Helsinki, approved by the local ethics committee ('Ethikkommission Paris Lodron-Universität Salzburg'), and registered at the German Clinical Trials Register (DRKS00003265). Participants gave written informed consent. Some of the data presented here (i.e. the data concerning sleep-dependent memory consolidation) have partly been published in Griessenberger *et al.* (2013).

Participants

With an *a priori* power analysis using G*Power we determined a sample size of at least 21 subjects to be sufficient to reach a power of 0.80 for our predefined primary endpoints, that is, increase in sleep spindle activity and decrease in sleep onset latency. We had found large effect sizes (Cohen's *d*, 0.86 to 1.19) in a previous NFT study of our group (Hoedlmoser *et al.*, 2008) for these primary endpoints. To be able to also reliably detect these changes in a clinical sample of patients with insomnia, we aimed for a sample size of 30 subjects.

We thus recruited 30 patients with primary insomnia [mean age = 38.59, standard deviation (SD) = 11.18, 19 females], who underwent NFT and PFT. An additional control group underwent NFT only (neurofeedback control, mean age = 26.67, SD = 4.46, six females) but, importantly, did not sleep in the lab and did not undergo PFT. This group was included to obtain normal NFT learning curves in the absence of any insomnia complaints and verified the efficacy of our NFT protocol (Hoedlmoser *et al.*, 2008).

Besides these two groups, we also included 31 age- and sex-matched healthy sleep controls (mean age = 35.52, SD = 10.63, 19 females) in the study. Importantly, this group did not undergo NFT or PFT, but only completed 'Visit 1' (Fig. 1) to obtain age- and sex-matched standard values for sleep and memory parameters. Insomnia patients were diagnosed according to the research criteria of Edinger *et al.* (2004) and had been free of medication for at least 2 weeks prior to study onset. Following an eligibility assessment [medical history, Pittsburgh Sleep Quality Index Questionnaire (PSQI), depression and anxiety questionnaires via email and phone] patients attended the lab where they were screened overnight more thoroughly for psychiatric disorders according to DSM-IV (American Psychiatric Association, 1995) using the Structured Clinical Interview for DSM disorders. They also completed additional questionnaires such as intelligence and personality questionnaires. Sleep controls did not have any history of past or current psychiatric disorders, which was verified through clinical interviews and the PSQI questionnaire (PSQI score $\leqslant 5$; Buysse *et al.*, 1989). Additionally, sleep controls could only take part in the study if their sleep efficiency during the screening night was < 1 SD below the average sleep efficiency in age- and sex-matched healthy sleepers (according to a European database: mean = 88.97; SD = 6.7; Anderer *et al.*, 2005). Patients received a remuneration of €500 after study completion (€150 for block 1, and €350 for block 2) whereas the control group received €100. Recruitment was through announcements in local newspapers, radio, flyers at general practitioners', and via our laboratory homepage.

For statistical analyses, we excluded five healthy controls (sleep controls) with abnormal sleep values (mean sleep efficiency $< 80\%$, mean wake after sleep onset > 50 min in two polysomnography nights). We also excluded five with insomnia as they did not complete both study parts (i.e. NFT and PFT) or missed multiple polysomnography nights. Importantly, we noticed that some patients with insomnia did not present with objective sleep problems (i.e. decreased sleep efficiency or increased waking after sleep onset) although all their subjective measures qualified them as primary insomnia patients. This subgroup of nine patients with insomnia exceeded a sleep efficiency of 90% and neither showed a sleep onset latency or wake after sleep onset exceeding 30 min on more than half of the laboratory nights (i.e. four of eight nights following the screening night), therefore we considered these patients not as patients with insomnia, but 'sleep state misperception' patients (misperception insomniacs). This subgroup was handled as an additional group in the statistical analyses. In conclusion, we included up to four different groups in our statistical analyses, i.e. patients with insomnia ($n = 16$), misperception insomniacs ($n = 9$), sleep control subjects ($n = 26$), and neurofeedback control subjects ($n = 12$).

Experimental design

After the eligibility assessment, patients with insomnia, misperception insomniacs and sleep control subjects slept in the laboratory for a screening night to exclude sleep disorders other than insomnia. Thereafter, two experimental nights followed (referred to as 'Visit 1' in the following). Each of these two nights was accompanied by either a procedural (finger-tapping task) or a declarative (word-pair association) memory task. As this study focuses on neurofeedback effects on EEG, sleep and quality of life in insomnia patients, we refrain from discussing the results from these memory tasks in further detail (Supplementary material). Before and after each night, subjects completed questionnaires regarding subjective sleepiness (Stanford Sleepiness Scale; Hoddes *et al.*, 1972) and mood [Multidimensional Mood Questionnaire (MDBF); Steyer *et al.*, 1997] and performed the psychomotor vigilance task (Dinges and Powell, 1985).

Patients with insomnia underwent this procedure four times (Visits 1–4) whereas sleep control subjects only completed the first visit (see above). In between the first and second, as well as the third and fourth of these visits, patients with insomnia completed 12 sessions of NFT and 12 sessions of PFT, i.e.

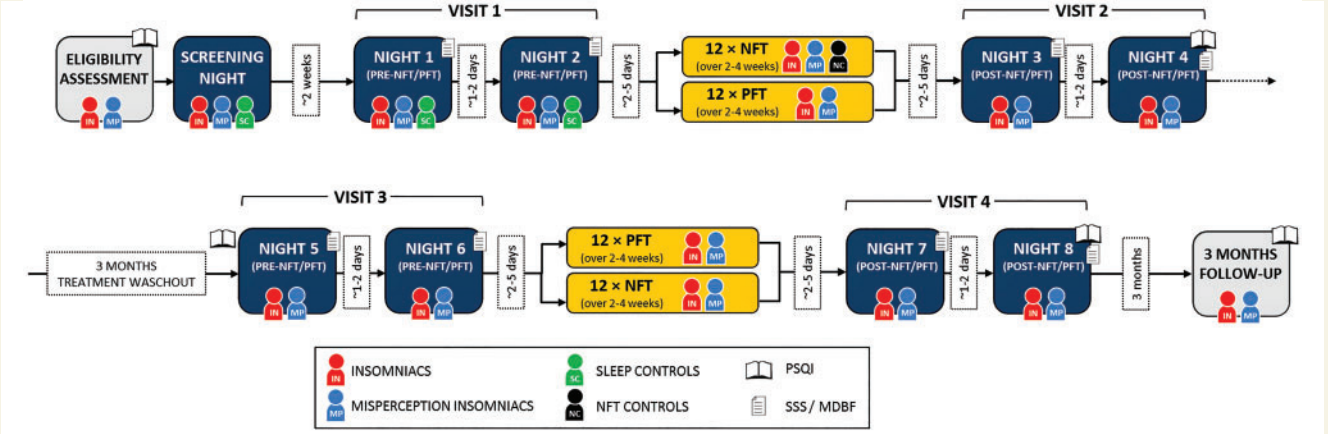

Figure 1 Study design. After participants had undergone an eligibility assessment (via email and phone) they came to the laboratory for a screening night to exclude sleep disorders besides insomnia. Then subjects completed four visits each, each one comprising two experimental nights with polysomnography (either preceded by a declarative or procedural learning task). In between Visits 1 and 2, and 3 and 4 all patients (patients with insomnia, misperception insomniacs) completed 12 sessions of NFT and PFT. Between Visits 2 and 3, a 3-month washout period was introduced and a final follow-up after 3 months was conducted. The order of NFT and PFT was counterbalanced (half of the patients receiving NFT first, and half of the patients receiving PFT first) and the protocol was kept double-blind until study completion.

a placebo or sham condition (with real EEG feedback, yet on varying frequency bands) (Fig. 1). The order of trainings (NFT or PFT) was counterbalanced across subjects and the 12 sessions were completed within ~2–4 weeks. Besides this, participants' sleep-wake cycle was assessed using sleep diaries and actigraphy (Cambridge Neurotechnology Actiwatch©) over the course of the whole protocol.

Neurofeedback and placebo-feedback methodology

In the 12 NFT or PFT training sessions, subjects learned to enhance EEG amplitudes within a specific frequency range while visual feedback was given online by the Eldith THERA PRAX (neuroConn) system. Each neurofeedback or instrumental conditioning session (NFT and PFT) consisted of eight 5-min training blocks (with ~13–25 trials within each block). Beyond this, the NFT protocol also included two 'transfer conditions' in which no immediate online feedback on the performance was given, but simply the reward or next trial sign at the end of the training block. The transfer blocks were included to better enable patients to apply the NFT technique at home, i.e. in the absence of NFT machinery and thus allow for a better transfer to the real world. For later NFT/PFT analyses only the six 5-min blocks with feedback were considered. Before the start and at the end of each training session, resting EEG activity was recorded during 2 min with eyes closed and 2 min with eyes open. Only the resting EEG recordings obtained during the eyes open period are of interest for the analyses presented here, therefore we refer to these recordings when mentioning resting state EEG hereafter.

The NFT/PFT training protocol used followed earlier studies from our group (Hoedlmoser *et al.*, 2008; Schabus *et al.*, 2014). Specifically, one trial consisted of a 3 s baseline measurement followed by a continuous feedback interval lasting until the EEG signal exceeded (for at least 250 ms) the predefined reward threshold established during the baseline. During the feedback interval, participants had to mentally 'move' a compass needle as far to the left as possible reaching the previously fixed threshold represented by a green dot (Supplementary Fig. 1). Subjects were instructed to test appropriate strategies and find their personally most successful approach to mastering the task. They were told that, for example, a combination of relaxation techniques and positive thought might help them exceed the threshold. After the appearance of the reward signal the next trial started with a new 3 s baseline measurement. Every 5 min there was a pause of ~1 min before the training continued with the next training block. To prevent rewarding artefacts or 'non-neural' strategies, trials with movements, eye or muscle artefacts as well as trials with amplitudes exceeding $\pm 200\,\mu V$ were abandoned and a new trial was started automatically. Importantly, during the NFT condition, participants had to enhance EEG amplitudes in the SMR range between 12 and 15 Hz, whereas during the PFT sessions participants had to enhance random frequency ranges between 7 and 20 Hz (but not the 12–15 Hz SMR range); importantly within a PFT session only one frequency was trained and rewarded. The reason for choosing this kind of placebo or sham protocol was to involve patients to a similar degree as in NFT, yet with no specific frequency being rewarded systematically. Rewarding another frequency systematically could have resulted in undesired effects on EEG and behaviour that would render the PFT control condition suboptimal.

To keep motivation balanced across NFT and PFT conditions and training blocks we adjusted the threshold in such a way that within each of the eight 5 min training blocks it was always similar amounts of reward that were given. Specifically, if less than 13 rewards were received in a 5 min block we lowered the threshold to be exceeded. Likewise, if more than 25 rewards were achieved, we increased the threshold that had to be exceeded. All patients (misperception insomniacs and patients with insomnia) had to complete both NFT and PFT sessions and were blind to the condition they were in at any point of the protocol. Contrary to earlier studies (Hoedlmoser *et al.*, 2008; Schabus *et al.*, 2014), the investigators were also

blind to the condition using a NFT/PFT code list simply stating 'training frequency A–F' for each of the 24 training sessions (i.e. 'third party concealment'). There was no monetary reward linked to training success; overall all participants seemed (given their psychological strain related to their chronic insomnia complaints) highly motivated throughout the training protocols.

EEG recordings

EEG recordings were done using Synamps EEG amplifiers (NeuroScan Inc.) with a sampling rate set to 500 Hz. The EEG setup comprised 22 scalp electrodes (Fp1, Fpz, Fp2, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, Oz, O2 plus the mastoids A1 and A2 for later offline re-referencing), one bipolar horizontal (HEOG) and one bipolar vertical electrooculogram (VEOG), one bipolar electrocardiogram (ECG) channel, one bipolar electromyogram (EMG) channel and one respiratory channel (chest wall movements). During the screening polysomnography we used a reduced setup with eight EEG, one bipolar ECG, two bipolar EOG, four respiratory measures (nasal airflow, chest and abdominal wall movements, oxygen saturation) and four unipolar EMG (submental and left/right tibialis nerve) electrodes. All scalp electrodes were placed according to the international 10-20 system (Jasper, 1958). Sleep was scored automatically by the SOMNOLYZER 24*7 (The Siesta Group) and verified manually by a sleep scoring expert following sleep scoring criteria of the American Academy of Sleep Medicine (AASM; Iber *et al.*, 2007).

Sleep spindle detection

Sleep spindles were detected automatically at frontal, central and parietal electrodes (F3, C3, P3) re-referenced to the contralateral mastoid electrodes. The spindle detection algorithm was based on the following criteria: (i) 11 to 16 Hz band-pass filtering; (ii) amplitude $>25\,\mu V$; (iii) duration >0.5 s; and (iv) no muscle (30–40 Hz) and/or alpha (8–12 Hz) artefacts (for details see Schimicek *et al.*, 1994). The algorithm computes spindle activity, which reflects duration and amplitude of spindles, and therefore quantifies the intensity of the spindle process (during non-REM sleep stage N2). Furthermore, we distinguished between slow (11–13 Hz) and fast (13–15 Hz) spindles and provide a measure of spindle density [spindle number per N2 sleep duration (min)].

EEG spectral analyses

EEG analyses were performed with BrainVision Analyzer 2.0 (Brain Products). In a first step, data were bandpass-filtered between 0.5 and 70 Hz and a 50 Hz notch filter was applied. Ocular artefacts were corrected for (Gratton *et al.*, 1983) and remaining artefacts were excluded manually. Afterwards, data were segmented into epochs of 2 s and a fast Fourier transformation was applied to obtain amplitude values in the frequency domain and finally we averaged values in the desired frequency range. We decided to focus on the trained 12–15 Hz frequency band (i.e. SMR) as well as the neighbouring 16–25 Hz beta band, which has been associated with hyperarousal upon falling asleep in insomnia subjects (Perlis *et al.*, 2001). In addition, we checked for changes in the theta (5–7 Hz) range as an increase in theta amplitude is supposed to indicate drowsiness. For analyses we selected electrodes C3 (used for feedback during NFT/PFT) and the contralateral electrode C4.

Statistical analyses

Data were statistically tested (IBM® SPSS® Statistics, Version 23; SPSS Inc., Chicago, Illinois) using repeated measures analyses of variance (ANOVAs) after having controlled for normal distribution of the data using the Kolmogorov-Smirnov test. Generally, we statistically evaluated (i) EEG effects during NFT/PFT; (ii) short-term EEG effects following NFT/PFT; and (iii) long-term effects of NFT/PFT on objective and subjective sleep parameters.

To investigate effects during training, we computed an ANOVA and tested for changes in EEG power in the SMR frequency range at electrode C3 (% change to baseline) during NFT and PFT training as compared to the 3 s baseline preceding each trial. This ANOVA included the factors Time (10 sessions, as due to technical problems data from Sessions 2 and 12 had to be discarded), Feedback (NFT versus PFT) and Group (patient with insomnia versus misperception insomniac). A supplementary analysis also investigated changes in the theta and beta frequency range, and compared effects at the trained electrode site C3 as well as the contralateral site C4 (Supplementary material).

Short-term EEG effects following NFT/PFT in the SMR (12–15 Hz) frequency range, i.e. training effects on the resting state EEG acquired immediately after NFT/PFT were tested in a repeated measures ANOVA with the within-subject factors Feedback (NFT versus PFT), Electrode (C3 versus C4), Time (12 training blocks) and Pre/post-training (before and after each training session). For this ANOVA we pooled both patient groups, i.e. patients with insomnia and misperception insomniacs. An identical ANOVA was computed for 16–25 Hz beta amplitude as this frequency range has previously been related to hyperarousal in insomnia. Resting state amplitudes are normalized to the individual total amplitude (1–30 Hz) to account for unspecific differences (e.g. skull thickness) between participants.

For evaluating long-term effects on objective sleep parameters and sleep spindles we always computed the mean of the two experimental nights preceding or following NFT/PFT.

Specifically, long-term effects of NFT/PFT on objective sleep parameters were evaluated using the nights flanking NFT and PFT, and taking the patient with insomnia and misperception insomniac group separately into account. The resulting ANOVA consisted of the factors Stage (Wake, N1, N2, N3, R in min), Feedback (NFT versus PFT), Pre/post (before and after each training block or Visit 1 to 2, and Visit 3 to 4) and Group (patients with insomnia, misperception insomniac).

Long-term effects of NFT/PFT were also computed on sleep spindles, which were hypothesized to change as a result of SMR-NFT training. An ANOVA was conducted for the NFT as well as PFT training effects with the factors Pre/post, spindle type (slow versus fast spindles), the between-subject factor Group (patients with insomnia, misperception insomniacs) and the dependent measures (i) sleep spindle activity; and (ii) sleep spindle density at (trained) electrode C3.

Last but not least, we computed long-term effects on subjective measures of sleep and life quality. An ANOVA with the factors Pre/post, Feedback and the between-subject factor Group (patients with insomnia versus misperception

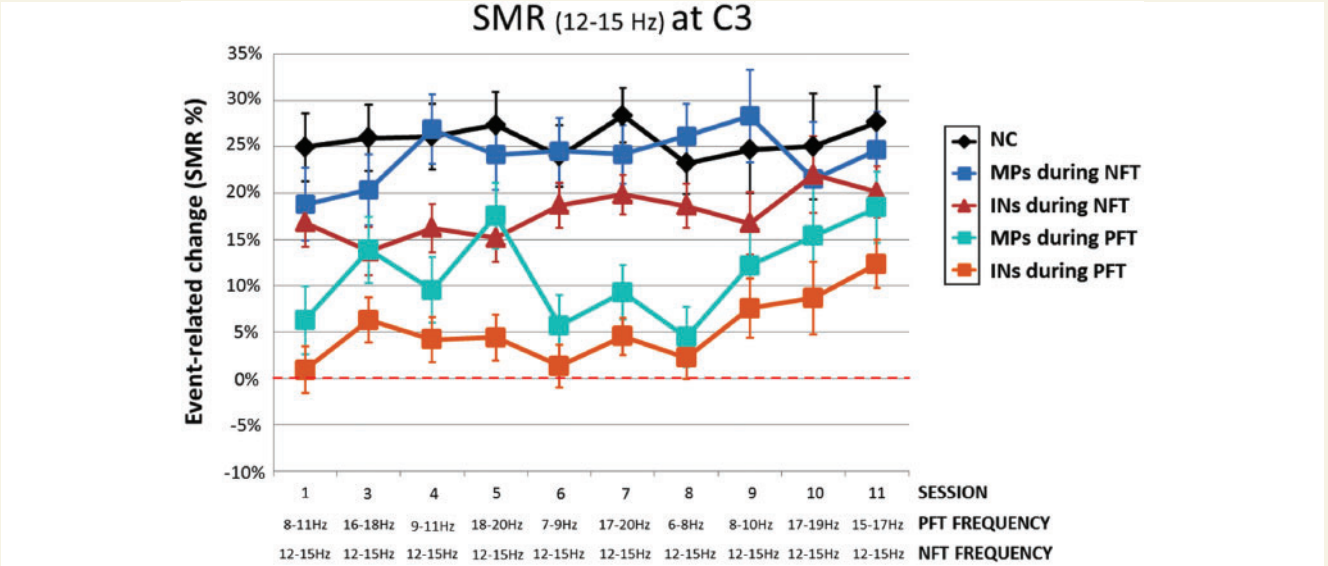

Figure 2 Effects of NFT/PFT training on SMR band power during training. Note that as a group all patients [patients with insomnia (INs), misperception insomniacs (MPs)] were able to significantly increase power in the SMR frequency band during NFT (as compared to PFT). A group of young healthy neurofeedback controls (NC) is plotted for comparison. Note that the *x*-axis informs about the rewarded frequency bands for PFT and NFT. Training sessions 2 and 12 are not displayed due to technical problems.

insomniacs) was run for the dependent measures subjective sleep quality (as assessed by the PSQI) and quality of life (as assessed by the World Health Organization Quality of Life Assessment; Skevington *et al.*, 2004). For 15 patients with insomnia and eight misperception insomniacs we obtained all four PSQI measures (i.e. one from each of the four visits) plus one additional follow-up after ~3 months. To assess the stability of the subjective sleep quality changes we finally performed paired-sample *t*-tests from the last training block (which could be NFT or PFT) to the follow-up. *Post hoc t*-tests following-up significant main effects or interactions of interest were corrected for multiple comparisons using the false discovery rate (FDR) method by Benjamini and Hochberg. The alpha-level used to determine significance was set to 0.05. For the evaluation of effect sizes we report partial eta-squared (η_p^2) as computed by SPSS for repeated measures ANOVAs.

Results

EEG effects during NFT/PFT

Confirming our expectations, we found significant differences between the two training conditions (i.e. NFT and PFT) and our patient groups (i.e. patients with insomnia and misperception insomniacs) when looking at objective EEG changes following NFT and PFT (Fig. 2).

An ANOVA concentrating on the SMR frequency range revealed main effects for the factors Time [$F(9,153) = 4.981$; $P < 0.001$; $\eta_p^2 = 0.227$] and Feedback [$F(1,17) = 20.536$; $P < 0.001$; $\eta_p^2 = 0.547$]. Latter main effect highlights that SMR-power was found to be higher for NFT (mean = 19.638; SD = 2.26) than PFT (mean = 8.753; SD = 2.22) (across all 12 training sessions). This effect was independent from whether patients were classified as patients with insomnia or misperception insomniacs.

Specifically, misperception insomniacs during NFT (mean = 24.29; SD = 7.56) outperformed patients with insomnia ($P < 0.001$) and misperception insomniacs during PFT ($P = 0.013$) as a group, meaning that misperception insomniac patients in the NFT condition were showing higher SMR changes to the baseline than did patients with insomnia and misperception insomniacs during PFT. Patients with insomnia during NFT (mean = 18.04; SD = 9.68) were only showing higher SMR power than when training under PFT conditions (mean = 5.57; SD = 8.68; $P < 0.001$) but did not outperform any other group.

All reported effects are stronger on the trained electrode site C3, yet they do transfer to the contralateral site, that is, C4 as illustrated in Supplementary Figs 2 and 3. In conclusion, we found evidence that both patients with insomnia and misperception insomniacs are able to upregulate the amplitude of the SMR frequency range after 12 training sessions. A supplementary analysis (Supplementary material) further confirms that the amount of SMR enhancement (to the baseline) achieved in our patients is not statistically different from a young healthy control group (neurofeedback control) conducting identical SMR-NFT.

Short-term effects of NFT/PFT on EEG

To evaluate short-term effects of NFT/PFT training, we checked for differences in the EEG measurements between

Table 1 Sleep parameters before and after PFT and NFT

	Patients with Insomnia												Misperception Insomniacs									
	Pre-PFT		Post-PFT				Pre-NFT		Post-NFT				Pre-PFT		Post-PFT				Pre-NFT		Post-NFT			
	Mean	SD	Mean	SD	t (p)		d		Mean	SD	t (p)		d		Mean	SD	t (p)		d		Mean	SD	t (p)		d	
TIB	474.0	18.4	478.4	4.9	−1.00	(0.33)	−0.33		482.2	15.8	479.0	3.9	0.84	(0.41)	0.29		479.4	4.2	479.9	5.8	−0.22	(0.83)	−0.1		478.4	5.0	480.6	3.4	−1.97	(0.08)	−0.88
TST	389.6	49.3	396.7	36.0	−0.66	(0.52)	−0.22		389.6	47.8	398.8	33.3	−0.85	(0.41)	−0.29		451.5	9.1	451.4	11.1	<0.01	(0.98)	<0.1		447.2	15.6	451.1	16.8	−0.47	(0.65)	−0.21
SEff	82.2	9.9	82.9	7.3	−0.31	(0.76)	−0.10		80.8	9.3	83.3	7.1	−1.16	(0.26)	−0.40		94.2	1.7	94.1	1.7	0.20	(0.85)	0.09		93.5	3.2	93.9	3.6	−0.22	(0.83)	<0. 1
SOL	27.7	18.4	24.3	16.0	0.7	(0.50)	0.22		26.9	16.9	25.7	17.4	0.20	(0.84)	0.07		11.5	6.9	10.2	8.8	0.90	(0.40)	0.42		11.2	5.8	16.4	16.7	−1.05	(0.32)	−0.47
WASO	60.9	40.7	62.8	33.5	−0.22	(0.83)	−0.05		72.0	41.7	62.5	35.4	1.08	(0.30)	0.24		19.1	7.6	20.7	8.1	−0.67	(0.52)	−0.20		22.2	14.1	19.3	11.0	0.48	(0.64)	0.24
NOA	21.5	7.8	23.2	9.2	−0.95	(0.36)	−0.32		23.7	10.4	20.8	5.3	1.42	(0.17)	0.49		*15.4*	7.7	16.2	6.6	−0.66	(0.53)	−0.10		16.5	6.6	16.8	7.1	−0.19	(0.86)	−0.09
N1%	13.9	5.9	14.2	5.1	−0.35	(0.73)	−0.12		14.6	5.02	14.2	6.0	0.33	(0.74)	0.11		9.9	2.7	10.3	2.5	−0.36	(0.72)	−0.15		12.6	4.4	11.0	4.1	1.00	(0.34)	0.45
N2%	46.9	7.2	48.6	6.3	−1.49	(0.15)	−0.50		44.6	7.5	45.9	6.0	−0.70	(0.50)	−0.24		45.9	7.2	48.1	5.7	−1.45	(0.19)	−0.31		46.6	5.5	45.5	7.0	0.55	(0.60)	0.25
N3%	21.0	7.4	19.4	7.7	1.66	(0.12)	0.55		21.2	7.7	19.9	6.2	1.10	(0.29)	0.38		22.1	5.1	20.4	5.5	1.82	(0.11)	0.31		19.2	7.1	20.8	6.1	−0.96	(0.36)	−0.43
R%	18.1	5.5	17.7	4.6	0.56	(0.59)	0.19		19.7	5.9	20.1	5.2	−0.37	(0.72)	−0.13		22.1	4.8	21.2	2.8	0.54	(0.61)	0.29		21.6	4.4	22.7	5.5	−0.72	(0.49)	−0.32
Subjective TST	*358.0*	*92.4*	*371.0*	*8.3*	*−0.87*	*(0.40)*	*−0.23*		*376.0*	*89.2*	*357.0*	*79.8*	*1.50*	*(0.16)*	*0.40*		*382.5*	*78.2*	*390.0*	*69.9*	*−0.55*	*(0.60)*	*−0.20*		*382.5*	*69.4*	*393.8*	*58.8*	*−0.60*	*(0.57)*	*−0.22*
Subjective SOL	*27.0*	*19.5*	*19.8*	*16.1*	*1.82*	*(0.09)*	*0.48*		*30.3*	*29.3*	*23.5*	*14.8*	*1.00*	*(0.34)*	*0.29*		*37.8*	*42.9*	*31.2*	*35.1*	*1.54*	*(0.17)*	*0.71*		*46.0*	*57.0*	*35.9*	*36.8*	*1.33*	*(0.22)*	*1.34*

Each column illustrates the mean of the two experimental polysomnography nights preceding or following NFT/PFT. The values of age and sex matched healthy controls (SC) are provided in the Supplementary material for comparison. All participants that completed experimental polysomnography nights before and after at least one condition (PFT and/or NFT) are included. Values in bold provide paired-sample *t*-statistics as well as Cohen's *d* effect size. Subjective total sleep time (TST) and sleep onset latency (SOL) are derived from PSQI questionnaires that were completed pre- and post NFT and PFT training blocks. These subjective tests were also completed at follow-up [patients with insomnia (IN): $mean_{TST}$ = 394.00, SD = 89.90 and $mean_{SOL}$ = 18.90, SD = 15.41; misperception insomniacs (MP), $mean_{TST}$ = 401.25, SD = 53.03 and $mean_{SOL}$ = 28.06, SD = 38.15]. NOA = number of awakenings; TIB = time in bed; Seff = sleep efficiency; WASO = wake after sleep onset.

the resting state recordings that preceded and immediately followed the trainings but found no changes across the 12 NFT/PFT training sessions. Please note that we here will exclusively focus on the trained SMR frequency range (see Supplementary material for additional information in the beta range). An ANOVA of the SMR frequency band showed that the type of feedback used during training (i.e. NFT versus PFT) did not have an effect in this frequency band that outlasted the training itself. Specifically, neither the main effects of Time [$F(11,242) = 0.624$, $P = 0.792$; $\eta_p^2 = 0.028$] nor Pre/post-training were significant [$F(1,22) = 0.677$, $P = 0.419$; $\eta_p^2 = 0.030$]. However a nearly significant main effect for the factor Electrode [$F(1,22) = 4.233$, $P = 0.052$; $\eta_p^2 = 0.161$] indicated that directly following training SMR was higher on the trained electrode C3 than on C4. No further interactions reached significance.

Long-term effects of NFT/PFT on objective and subjective sleep parameters

Long-term effects on objective sleep parameters

For all objective measures of sleep architecture we did not find any significant changes across PFT or NFT training blocks. Only a Stage × Group interaction reached significance [$F(4,92) = 9.736$, $P < 0.001$; $\eta_p^2 = 0.30$] and indicated that patients with insomnia presented with more wake time, yet less N2, N3 and REM than misperception insomniacs. Subjective total sleep time and sleep onset latency (derived from the PSQI) were also evaluated before and after the NFT/PFT training blocks but likewise revealed no significant changes. See Table 1 for a listing of all sleep parameters before and after NFT/PFT and separately for patients with insomnia and misperception insomniacs.

Long-term effects on sleep spindles

Analyses of long-term effects of NFT/PFT training on sleep spindle measures did not indicate that these measures increased by any type of training (Fig. 4). Importantly, that suggests that besides the lack of short-term effects, NFT does not seem to be efficacious regarding these EEG-derived measures in the long run either. An ANOVA for the dependent measure spindle activity in the NFT condition revealed no main effects for the factors Pre/post-training [$F(1,23) = 2.153$, $P = 0.16$; $\eta_p^2 = 0.09$], spindle type [$F(1,23) = 0.005$, $P = 0.94$; $\eta p^2 < 0.01$], the between-subject factor Group [$F(1,23 = 0.077$, $P = 0.78$; $\eta p^2 = 0.03$] or any of the interactions. An ANOVA for the dependent measure spindle density in the NFT condition revealed only a main effect for the factor Type [$F(1,23 = 38.322$, $P < 0.001$; $\eta p^2 = 0.63$] indicating a higher spindle density for the fast spindle type on electrode C3. This was generally expected as the number of fast spindles is usually higher at

centro-parietal sites as compared to (more frontal) slow spindles (Anderer *et al.*, 2001). Neither the factor Pre/post-training [$F(1,23) = 0.244$, $P = 0.63$; $\eta p^2 = 0.01$], nor the between-subject factor Group [$F(1,23) = 0.173$, $P = 0.68$; $\eta p^2 = 0.01$] or any of the interactions were significant.

For the PFT condition an ANOVA for the dependent measure spindle activity revealed no main effect for the factor Type [$F(1,22) = 0.005$, $P = 0.94$; $\eta p^2 < 0.001$], or the between-subject factor Group [$F(1,22) = 0.297$, $P = 0.59$; $\eta_p{}^2 = 0.013$] nor were any of the interactions significant. Yet, one main effect was marginally significant, with the factor Pre/post-training [$F(1,22) = 3.298$, $P = 0.083$; $\eta_p{}^2 = 0.130$] indicating a trend towards decreased spindle activity after PFT.

For the dependent measure spindle density only a main effect for the factor Type [$F(1,22) = 46.572$, $P < 0.001$; $\eta_p{}^2 = 0.679$] was significant, again indicating an increased prevalence of fast spindles on the central recording site C3 as compared to slow spindles. In summary, sleep spindle activity, as well as sleep spindle density, were not affected by NFT or PFT.

Long-term effects on subjective sleep quality and quality of life

Contrasting objective EEG-derived measures, NFT training did have a beneficial effect on subjective measures of sleep quality. To a lesser extent, that was also true for subjective (physical) quality of life. Analyses revealed a main effect for the factor Pre/post-training [$F(1,21) = 7.621$, $P = 0.01$; $\eta_p{}^2 = 0.266$]. Specifically, subjective sleep complaints decreased from pre- to post-training. Yet, no other effects for the factors Feedback [$F(1,21) = 0.169$, $P = 0.69$; $\eta_p{}^2 = 0.008$], or the between-subject factor Group [$F(1,21) = 1.598$, $P = 0.22$; $\eta_p{}^2 = 0.071$] were significant (Fig. 5). Importantly, also the interaction Pre/post training × Feedback [$F(1,21) = 0.109$, $p = 0.75$; $\eta_p{}^2 = 0.005$] was not significant, nor was the three-way interaction Pre/post training × Feedback × Group [$F(1,21) = 0.010$, $P = 0.92$; $\eta_p{}^2 = 0.0$] significant. Overall, the ANOVA thus shows a significant decrease of subjective sleep complaints, which is independent from the type of feedback (i.e. NFT versus PFT) and independent from whether they were patients with insomnia or misperception insomniacs.

Finally, we compared the evolution of the subjective sleep complaints following the last training block to the follow-up after 3 months. Paired-sample *t*-tests revealed a marginally significant further improvement (from mean = 5.50, SD = 2.91 to mean = 4.58, SD = 3.15) for the patients who had ended their training with PFT [$t(11) = 2.200$, $P = 0.050$, $d = 0.299$]. This was however not true for patients who had terminated the study protocol with NFT training [from mean = 7.73, SD = 3.74 to mean = 6.82, SD = 4.22; $t(10) = 1.311$, $P = 0.22$, $d = 0.224$]. In general effect sizes indicate a small positive effect for the change in subjective sleep quality from pre- to post-training.

Despite positive effects of training on subjective measures of sleep quality, quality of life remained unaffected. Subjective quality of life as assessed by the The World Health Organization Quality of Life Assessment did likewise not indicate a systematic increase of life quality over the NFT or PFT sessions (Supplementary Fig. 5). Yet, focusing on the World Health Organization Quality of Life Assessment subdimensions we were able to confirm the unspecific increase of physical quality of life (incorporating facets such as fatigue, physical discomfort or work capacity) from the first experimental polysomnography night to the 3-month follow-up [$t(22) = -3.531$, $P = 0.002$] as reported previously (Schabus *et al.*, 2014) (Supplementary Fig. 6).

Last but not least, subjective awakening quality (i.e. morning sleepiness as assessed by the Stanford Sleepiness Scale, and mood after awakening as assessed by the MDBF questionnaire) was unaffected by NFT/PFT training (Supplementary material).

Discussion

To the best of our knowledge, we here present the first rigorously controlled study on the efficacy of NFT for insomnia since the promising findings obtained in pioneering studies by Sterman *et al.* (1970), Hauri (1981) and Hauri *et al.* (1982). The present study (i) provides support for the principle of NFT, i.e. that participants can learn to control neural processes when they receive adequate feedback; and (ii) finds positive evidence that such feedback can diminish the subjective burden of a disease. Critically though, the results call into question whether the positive findings reported for NFT in the literature are indeed NFT-specific or whether they are due to rather unspecific effects, such as receiving attention, care and support from experimenters.

Unfortunately, well-controlled studies that may help disentangle NFT-specific and unspecific therapeutic effects underlying efficacy are sparse in the field, thus eventually rendering a reliable evaluation of the efficacy of this kind of 'neurotherapy' difficult. Although some controlled studies do exist for example for the treatment of attention deficit hyperactivity disorder (ADHD) symptoms (Arns *et al.*, 2009; Gevensleben *et al.*, 2009), still the majority of them lacks a real placebo group (e.g. groups of different training duration or intensity) making it impossible to evaluate the specific effects of NFT training. The importance of well-controlled studies is further underlined by recent reports that even the few studies that actually included placebo conditions do not find neurofeedback to be superior to placebo or 'sham' feedback (Vollebregt *et al.*, 2014; Thibault *et al.*, 2015; Thibault and Raz, 2016). Moreover, sample sizes are often small and many studies solely rely on subjective ratings rather than reporting objective changes in EEG parameters or other neural measures that are expected to change through the feedback.

Generally, we found that participants can obtain control of otherwise unconscious neural processes with the help of feedback thereby supporting the basic rationale behind this family of therapeutic approaches. Specifically, our results suggest that regulation of EEG activity, at least in narrow frequency bands (here 12–15 Hz or SMR oscillations) can be learned quickly so that participants successfully perform according to instructions and feedback. In addition, our results suggest that the ability to learn may also depend on general learning abilities of the individual with young healthy subjects (i.e. the neurofeedback control group) appearing to learn more quickly and exhibiting steeper learning curves than patients with insomnia (Fig. 2). Specifically, healthy neurofeedback control subjects already seem to level off after a single NFT session (i.e. 8 × 5 min training) whereas patients need four (in misperception insomniacs) to six (in patients with insomnia) training sessions to achieve a similar effect. Interestingly, this learning effect was not limited to the NFT condition, but it was also evident in the sham or PFT condition (with 15–20 Hz Beta enhancements) (Supplementary Fig. 3) thus further backing the notion of general learning deficits limiting NFT success.

Beyond this, these findings also question other reports advocating that neurofeedback training needs to comprise at least 20–40 and sometimes even up to 50 training sessions in order to be effective (Hammond, 2011). More training sessions may indeed be necessary when working with participants with more pronounced learning difficulties or if a successful transfer to the real-world without a neurofeedback device is desired. Yet given the absence of even short-term EEG effects immediately following the NFT sessions it is unlikely that increasing the number of training sessions will change the general outcome. One might argue that patients simply did not learn to control their SMR activity well enough to earn benefits such as improved sleep or memory consolidation. However, that seems implausible, given the fact that misperception insomniac patients as compared to young healthy neurofeedback control subjects even reach identical NFT learning levels (over the 12 training sessions).

For NFT to be recommended for the treatment of primary insomnia, NFT has to (i) outperform a placebo condition; (ii) bring about positive effects in the long run; and (iii) ideally be as rapidly acting as the current (non-pharmacological) standard-of-care treatment. In our study, patients who underwent the training protocol (i.e. NFT or PFT) did report an improvement in subjective sleep quality and (physical) quality of life when asked 3 months after having completed the training protocol, i.e. during the follow-up. Generally, this improvement on subjective insomnia complaints is in accordance with earlier results obtained by other groups (Hauri, 1981; Cortoos *et al.*, 2010) as well as in our laboratory (Schabus *et al.*, 2014). As the current study is characterized by an increase in external validity due to the design being highly controlled (i.e. double-blinded and including a placebo or 'sham' condition) and validated in a bigger sample, it provides credible support for the notion that NFT/PFT is able to improve insomnia symptoms. Besides this, similar subjective improvements following EEG neurofeedback have been found in studies involving samples of children with ADHD (Lansbergen *et al.*, 2011; Arnold *et al.*, 2013; van Dongen-Boomsma *et al.*, 2013). Crucially however, revealed improvements were independent of the type of feedback training (i.e. NFT or PFT) patients underwent. Essentially, this suggests that while subjective effects are indeed stable in the long run, the observed improvements were rather due to unspecific therapeutic than to NFT-specific factors. Thus, the current study implies, in line with recently published findings (Arnold *et al.*, 2013; Thibault *et al.*, 2015), that NFT is not systematically superior to other placebo or 'sham' feedback conditions; at least not at durations comparable to current standard-of-care treatments (CBT-I).

Besides using subjective measures, we importantly also aimed at evaluating objective measures of symptom reduction. To our surprise, however, the observed changes in subjectively reported sleep quality were not accompanied by any changes in objective EEG-derived measures of sleep quality. Specifically, we expected positive changes in sleep onset latency (Hoedlmoser *et al.*, 2008) and/or the number of awakenings and the amount of slow-wave sleep (Schabus *et al.*, 2014). The more severe nature of the insomnia symptoms and the higher age in the current study (i.e. a kind of 'learning deficit' as discussed above) may be one reason for the patients in this double-blind protocol not objectively improving despite the earlier findings. On the other hand, (earlier) single-blind designs suffer from the inherent risk that that laboratory staff and experimenters may inadvertently bring about the desired effects, for example by subtly paying more attention to the patients' complaints and needs in the NFT condition. Indeed, we found evidence for this when we looked at subjectively perceived social support in a previous single-blind study from our own group (unpublished results; Schabus *et al.*, 2014).

The last and most important aim of this study was to shed light at the mechanisms underlying the efficacy of NFT by looking at objective measures of the EEG processes participants were to gain control of. Unfortunately, the existing NFT literature does not show a particularly strong tradition in reporting measures of this kind, although without doubt this would greatly benefit the field's credibility. In the present study we found that despite NFT and, importantly also PFT, being beneficial on a subjective level, objective measures of EEG activity outside the training period remain unchanged even when evaluated only minutes after training (Fig. 3). The latter is also true for sleep spindles, which are the first logical candidates for parameters which should change as a consequence of SMR-NFT training (as they lie in the same 12–15 Hz frequency range).

In summary, our results show that patients benefitted from any treatment on a subjective level. Objectively,

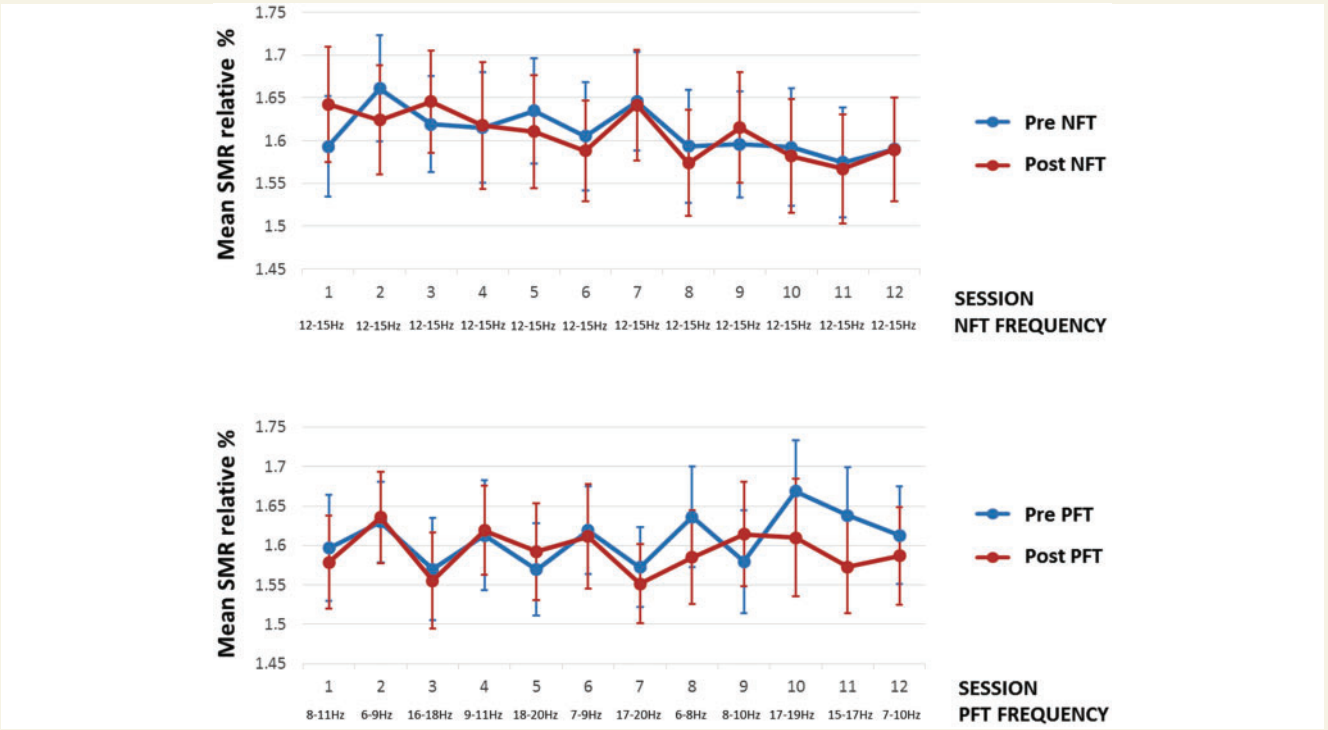

Figure 3 Short-term effects in the SMR band. A resting EEG (with eyes open) was recorded directly before and after each PFT and NFT training block. Analyses revealed that even directly following training, patients with insomnia and misperception insomniacs (here pooled) had SMR amplitude values (on the trained site C3) that did not differ from the values preceding the training blocks. Note that amplitude is normalized to the individual total-amplitude (1–30 Hz) to account for unspecific differences between participants.

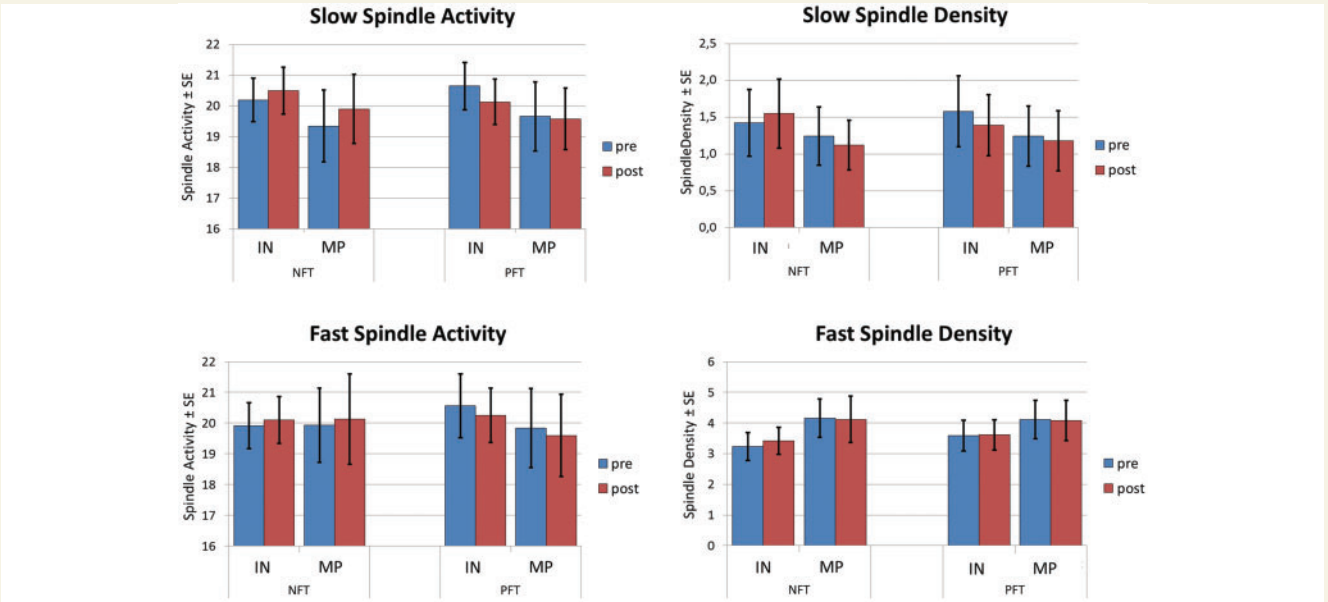

Figure 4 Long-term effects of NFT/PFT training on sleep spindle activity and sleep spindle density. The graphs on the *left* illustrate the slow (11–13 Hz) and fast (13–15 Hz) sleep spindle activity (mean spindle amplitude × duration) for the NFT and PFT conditions. The graphs on the *right* illustrate sleep spindle density (number of spindles / min). All spindle detections have been performed on (trained) electrode C3 and in N2 sleep where spindles are most prevalent. Pre and post refer to the mean of two full polysomnography nights before or after 12 sessions of NFT/PFT. Error bars indicate ± 1 standard error of the mean (SEM).

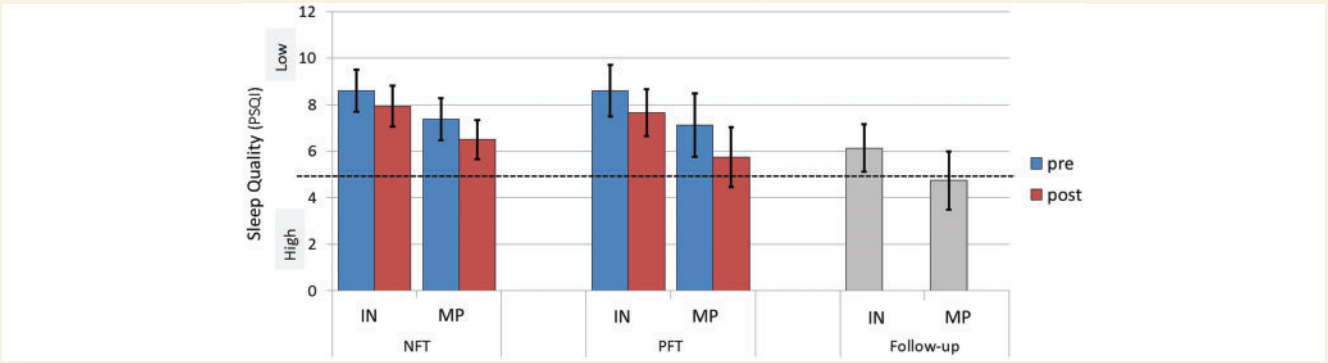

Figure 5 Long-term effects of NFT/PFT training on subjective sleep quality (PSQI). The figure depicts the subjective changes in sleep quality as evaluated with the PSQI. Participants completed the questionnaire before and after the 12 NFT and PFT sessions, as well as 3 months thereafter (i.e. follow-up). Note a general decrease of sleep complaints from pre- to post-training independent from NFT/PFT and a tendency to further decrease to the follow-up. A PSQI total score > 5 is indicative of poor sleep (marked with the dashed line). Error bars indicate $\pm$ 1 SEM. IN = patients with insomnia; MP = misperception insomniacs.

however, this improvement was not reflected in EEG-derived measures. Most importantly, we found that improvement of symptoms was not specific to NFT, but rather seems to have been brought about by unspecific factors such as affection and care. Altogether, it therefore has to be questioned whether (SMR) neurofeedback can be promoted as an alternative to established therapeutic approaches. Our findings may thereby also stimulate a discussion regarding the usefulness of neurofeedback on a more general level. One aspect that may have been widely neglected until now is the above discussed influence of learning abilities that may be compromised in clinical populations in general.

One may argue that a higher number of training sessions or individually tailored NFT protocols may be more successful in evoking the desired objective changes regarding brain activity and symptoms and that this may even have led to a specific effect for NFT. However, compared to CBT-I interventions, our protocol was already rather extensive. Moreover, the feedback protocol used here was designed following an extensive review of the literature and in close coordination with experts in the field. We also followed up on the idea that our groups may have included NFT-responders and non-responders. However, approaches to stratifying our patient group according to various criteria were not successful. This may also be due to there being no commonly accepted criteria for NFT-(non)-responders in the literature (Dempster and Vernon, 2009).

Without doubt, there will be patients who are more compliant and possibly more responsive to a modern 'neurotherapy' technique like NFT than to CBT-I. Yet, this justification is not enough, especially in times where financial resources in the health care systems are limited. We believe that it is essential that the neurofeedback research community backs the often far-reaching promises from little controlled studies by placebo-controlled, double-blind studies along with adequate sample sizes and statistical analyses for each of the disorders that efficacy is claimed for. Similarly, this applies to the ever growing field of EEG-neurofeedback for optimizing performance in healthy individuals (Gruzelier, 2014).

Based on this comprehensive and well-controlled study, we conclude that neurofeedback for treating primary insomnia complaints does not have a specific efficacy beyond unspecific factors. Importantly, we do not find an advantage for NFT over PFT and therefore cannot recommend it over the current non-pharmacological standard-of-care treatment, that is, CBT-I. Refined study designs and neurofeedback protocols should be welcomed by the research community, yet they will have to withstand rigorous testing against real placebo conditions using double-blind designs.

Acknowledgements

We would like to especially thank M.B. Sterman and M. Doppelmayr for their constant support and assistance with our NFT research and NFT protocols in particular. We are also grateful to Daniela Baronner, Christine Blume, Julia Lechinger, Nikolina Luketina and Daniel Körner who supported us with data acquisition, preprocessing, analysis or feedback on the manuscript.

Funding

Research was supported by the FWF research grant P-21154-B18 from the Austrian Science Fund. In addition this study was supported by the Doctoral College 'Imaging the Mind' (FWF; W1233) (D.P.J.H, and M.W.).

Supplementary material

Supplementary material is available at *Brain* online.

References

American Psychiatric Association. Diagnostic and statistical manual of mental disorders, DSM-IV. 4th edn. Washington, DC: American Psychiatric Association; 1995.

Anderer P, Gruber G, Parapatics S, Woertz M, Miazhynskaia T, Klosch G, et al. An E-health solution for automatic sleep classification according to Rechtschaffen and Kales: validation study of the Somnolyzer 24 x 7 utilizing the Siesta database. Neuropsychobiology 2005; 51: 115–33.

Anderer P, Klosch G, Gruber G, Trenker E, Pascual-Marqui RD, Zeitlhofer J, et al. Low-resolution brain electromagnetic tomography revealed simultaneously active frontal and parietal sleep spindle sources in the human cortex. Neuroscience 2001; 103: 581–92.

Arnold LE, Lofthouse N, Hersch S, Pan X, Hurt E, Bates B, et al. EEG neurofeedback for ADHD: double-blind sham-controlled randomized pilot feasibility trial. J atten disord 2013; 17: 410–9.

Arns M, de Ridder S, Strehl U, Breteler M, Coenen A. Efficacy of neurofeedback treatment in ADHD: the effects on inattention, impulsivity and hyperactivity: a meta-analysis. Clin EEG Neurosci 2009; 40: 180–9.

Arns M, Feddema I, Kenemans JL. Differential effects of theta/beta and SMR neurofeedback in ADHD on sleep onset latency. Front Hum Neurosci 2014; 8: 1019.

Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989; 28: 193–213.

Cortoos A, De Valck E, Arns M, Breteler MH, Cluydts R. An exploratory study on the effects of tele-neurofeedback and tele-biofeedback on objective and subjective sleep in patients with primary insomnia. Appl Psychophysiol Biofeedback 2010; 35: 125–34.

Dempster T, Vernon D. Identifying indices of learning for alpha neurofeedback training. Appl Psychophysiol Biofeedback 2009; 34: 309–18.

Dinges DF, Powell JW. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. Behav Res Methods Instrum Comput 1985; 17: 652–5.

Edinger JD, Bonnet MH, Bootzin RR, Doghramji K, Dorsey CM, Espie CA, et al. Derivation of research diagnostic criteria for insomnia: report of an American Academy of Sleep Medicine Work Group. Sleep 2004; 27: 1567–96.

Gevensleben H, Holl B, Albrecht B, Vogel C, Schlamp D, Kratz O, et al. Is neurofeedback an efficacious treatment for ADHD? A randomised controlled clinical trial. J Child Psychol Psychiatry 2009; 50: 780–9.

Gratton G, Coles MG, Donchin E. A new method for off-line removal of ocular artifact. Electroencephalogr Clin Neurophysiol 1983; 55: 468–84.

Griessenberger H, Heib DPJ, Lechinger J, Luketina N, Petzka M, Moeckel T, et al. Susceptibility to declarative memory interference is pronounced in primary insomnia. PloS One 2013; 8: e57394.

Gruzelier JH. EEG-neurofeedback for optimising performance. I: a review of cognitive and affective outcome in healthy participants. Neurosci Biobehav Rev 2014; 44: 124–41.

Hammond DC. What is neurofeedback: an update. J Neurother 2011; 15: 305–36.

Hauri P. Treating psychophysiologic insomnia with biofeedback. Arch Gen Psychiatry 1981; 38: 752–8.

Hauri PJ, Percy L, Hellekson C, Hartmann E, Russ D. The treatment of psychophysiologic insomnia with biofeedback: a replication study. Biofeedback Self Regul 1982; 7: 223–35.

Hoddes E, Zarcone V, Dement W. Development and use of Stanford Sleepiness Scale (SSS). Psychophysiology 1972; 9: 150.

Hoedlmoser K, Dang Vu TT, Desseilles M, Schabus M. Non pharmacological alternatives for the treatment of insomnia—Instrumental EEG conditioning, a new alternative. In: Melatonin, Sleep and Insomnia. New York: Nova Science Publishers, 2011. p. 69–101.

Hoedlmoser K, Pecherstorfer T, Gruber G, Anderer P, Doppelmayr M, Klimesch W, et al. Instrumental conditioning of human sensorimotor rhythm (12–15 Hz) and its impact on sleep as well as declarative learning. Sleep 2008; 31: 1401–8.

Iber C, Ancoli-Israel S, Chesson A, Quan SF. The AASM Manual for the Scoring of Sleep and Associated Events: Rules Terminology and Technical Specifications. Westchester, Illinois: American Academy of Sleep Medicine, 2007.

Jasper HH. The ten twenty electrode system of the international federation. Electroencephalogr Clin Neurophysiol 1958; 10: 371–5.

Lansbergen MM, van Dongen-Boomsma M, Buitelaar JK, Slaats-Willemse D. ADHD and EEG-neurofeedback: a double-blind randomized placebo-controlled feasibility study. J Neural Transm 2011; 118: 275–84.

Morin CM, Bootzin RR, Buysse DJ, Edinger JD, Espie CA, Lichstein KL. Psychological and behavioral treatment of insomnia: update of the recent evidence (1998–2004). Sleep 2006a; 29: 1398–414.

Morin CM, LeBlanc M, Daley M, Gregoire JP, Merette C. Epidemiology of insomnia: prevalence, self-help treatments, consultations, and determinants of help-seeking behaviors. Sleep Med 2006b; 7: 123–30.

Perlis ML, Smith MT, Andrews PJ, Orff H, Giles DE. Beta/Gamma EEG activity in patients with primary and secondary insomnia and good sleeper controls. Sleep 2001; 24: 110–17.

Riemann D, Nissen C, Palagini L, Otte A, Perlis ML, Spiegelhalder K. The neurobiology, investigation, and treatment of chronic insomnia. Lancet Neurol 2015; 14: 547–58.

Schabus M, Heib DP, Lechinger J, Griessenberger H, Klimesch W, Pawlizki A, et al. Enhancing sleep quality and memory in insomnia using instrumental sensorimotor rhythm conditioning. Biol Psychol 2014; 95: 126–34.

Schimicek P, Zeitlhofer J, Anderer P, Saletu B. Automatic sleep-spindle detection procedure: aspects of reliability and validity. Clin Electroencephalogr 1994; 25: 26–9.

Skevington SM, Lotfy M, O'Connell KA. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004; 13: 299–310.

Sterman MB. Basic concepts and clinical findings in the treatment of seizure disorders with EEG operant conditioning. Clin Electroencephal 2000; 31: 45–55.

Sterman MB, Howe RC, Macdonald LR. Facilitation of spindle-burst sleep by conditioning of electroencephalographic activity while awake. Science 1970; 167: 1146–8.

Steyer R, Schwenkmezger P, Notz P, Eid M. The multidimensional mental state questionnaire (Mehrdimensionale Befindlichkeitsfragebogen (MDBF)): manual. Göttingen, Germany: Hogrefe; 1997.

Tan G, Thornby J, Hammond DC, Strehl U, Canady B, Arnemann K, et al. Meta-analysis of EEG biofeedback in treating epilepsy. Clin EEG Neurosci 2009; 40: 173–9.

Thibault RT, Lifshitz M, Birbaumer N, Raz A. Neurofeedback, self-regulation, and brain imaging: clinical science and fad in the service of mental disorders. Psychother Psychosom 2015; 84: 193–207.

Thibault RT, Raz A. When can neurofeedback join the clinical armamentarium? Lancet Psychiat 2016; 3: 497–8.

Thorndike EL. The elements of psychology. New York, NY: A.G Seiler; 1905.

van Dongen-Boomsma M, Vollebregt MA, Slaats-Willemse D, Buitelaar JK. A randomized placebo-controlled trial of electroencephalographic (EEG) neurofeedback in children with attention-deficit/hyperactivity disorder. J Clin Psychiatry 2013; 74: 821–7.

Vollebregt MA, van Dongen-Boomsma M, Buitelaar JK, Slaats-Willemse D. Does EEG-neurofeedback improve neurocognitive functioning in children with attention-deficit/hyperactivity disorder? A systematic review and a double-blind placebo-controlled study. J Child Psychol Psychiatry 2014; 55: 460–72.