

LETTER TO THE EDITOR

Reply: Noisy but not placebo: defining metrics for effects of neurofeedback

Manuel Schabus^{1,2}

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

Sir,

We explicitly intend to keep the current response short as we believe that only reliable, double-blind studies with good controls can shed light on the serious doubts raised by us (Schabus *et al.*, 2017) and others (Vollebregt *et al.*, 2014; Cortese *et al.*, 2016; Thibault and Raz, 2016; Schonenberg *et al.*, 2017; Thibault *et al.*, 2017, 2018).

In the following we will address the main points raised by Witte *et al.* (2018):

- (i) Witte and colleagues correctly mention that in our earlier response (Schabus, 2017) we showed ‘an increase in physical quality-of-life (QoL) ratings across trainings sessions for real as well as sham NFT’. Witte and colleagues criticize that we took this as argument that this is indicative of a major placebo effect and point to the ‘noisy’ subjective data. We are well aware of the ‘noisy nature’ of purely questionnaire-derived QoL data as well as the effect of sample sizes. However, we wanted to emphasize that purely subjective measures—which are reported exclusively, i.e. without ‘neural data’ in most neurofeedback (NFT) studies—will almost always show improvements independent of whether real NFT or a sham-control is presented. Given the results of our earlier study, we additionally speculated that subjects may actually feel more supported in certain subjective dimensions (here, social quality of life including social support questions) if truly double-blind designs are not adopted (Fig. 1, Schabus, 2017; Schabus *et al.*, 2017). We agree that

the variance in such data is huge and that sample sizes well above 20–30 participants per group are highly desired. We are therefore eager to see whether larger (well controlled) studies will come to different conclusions.

- (ii) Witte *et al.* bring up the question what a ‘placebo’ is in the first place. According to Price and colleagues (2008) ‘Placebos have typically been identified as inert agents or procedures aimed at pleasing the patient rather than exerting a specific effect’. According to our understanding, this is exactly what we have reported earlier (Schabus, 2017; Schabus *et al.*, 2017). We see patients with (a) increased subjective sleep quality (PSQI) (Fig. 5, Schabus *et al.*, 2017); and (b) increased QoL (Fig. 1, Schabus, 2017) but importantly, no specific NFT effect. That is, NFT does not (a) bring about larger subjective improvements than sham feedback; nor (b) does power in the trained EEG frequency bands change even minutes after training (Fig. 3, Schabus *et al.*, 2017); or (c) would NFT lead to changes in sleep architecture (Table 1, Schabus *et al.*, 2017); or (d) sleep spindles (Fig. 4, Schabus *et al.*, 2017) during subsequent sleep. We even agree that patients’ outcome expectations or the treatment context may contribute more to the outcome than treatment-specific effects (Schedlowski *et al.*, 2015). Yet, it is then ethically questionable whether such treatment needs high-tech NFT equipment and justifies expensive ‘neurotherapy’ sessions for the patients.

(iii) The last argument addresses what one can actually consider a ‘systematic change’ in EEG-derived parameters after NFT training. This is without doubt a question open to discussion, perhaps even a question the NFT field has ignored for too long. Perhaps this is the reason for the NFT field ‘answering’ the question by simply not reporting EEG parameters at all, but still claiming that their ‘neurotherapy’ brings about neuronal changes. It remains difficult to picture how wellbeing or behaviour should change (due to neural changes) over time if even the trained frequency bands (here, 12–15 Hz) don’t show any statistically relevant change minutes after training has ended. The baseline changes that the authors mention are not an issue according to our experience as they stayed stable across all 12 training sessions in our protocol (*cf.* Fig. 2, Schabus, 2017). One can take the alternative standpoint of Witte *et al.* (2018) that NFT is just about ‘achieving an immediate regulation ability’. Yet, would one then not still expect some NFT-specific and objective changes in symptomatology of the patients? For example, in our case changes in some of the polysomnography-derived sleep quality or memory-related measures.

Altogether, we agree with Witte and colleagues that well-designed studies and standards are highly overdue. However, in our opinion this is not limited to a lack of standards for NFT data analysis but likewise for data acquisition and specific NFT protocols applied to different groups with different ‘outcome aims’. We appreciate that Witte and colleagues openly discuss such important issues and indeed present some ‘neuro’ data in their publications. However, we widely disagree with their definition of ‘high scientific standards’; especially in an area that is looked at with so much doubt from scientists outside of their own NFT in-group. In our view, many more rigorously controlled and pre-registered studies (e.g. Schabus *et al.*, 2017; Schonenberg *et al.*, 2017) as well as robust meta-analyses (e.g. Sonuga-Barke *et al.*, 2013; Cortese *et al.*, 2016) are needed if the field finally strives to establish scientific credibility for their method of choice.

Funding

Research was supported by the FWF research grant P-21154-B18 and Y777 from the Austrian Science Fund.

References

- Cortese S, Ferrin M, Brandeis D, Holtmann M, Aggensteiner P, Daley D, et al. Neurofeedback for attention-deficit/hyperactivity disorder: meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. *J Am Acad Child Adolesc Psychiatry* 2016; 55: 444–55.
- Price DD, Finniss DG, Benedetti F. A comprehensive review of the placebo effect: recent advances and current thought. *Annu Rev Psychol* 2008; 59: 565–90.
- Schabus M. Reply: On assessing neurofeedback effects: should double-blind replace neurophysiological mechanisms? *Brain* 2017; 140: e64.
- Schabus M, Griessenberger H, Gnjezda MT, Heib DPJ, Wislowska M, Hoedlmoser K. Better than sham? A double-blind placebo-controlled neurofeedback study in primary insomnia. *Brain* 2017; 140: 1041–52.
- Schedlowski M, Enck P, Rief W, Bingel U. Neuro-bio-behavioral mechanisms of placebo and nocebo responses: implications for clinical trials and clinical practice. *Pharmacol Rev* 2015; 67: 697–730.
- Schonenberg M, Wiedemann E, Schneidt A, Scheeff J, Logemann A, Keune PM, et al. Neurofeedback, sham neurofeedback, and cognitive-behavioural group therapy in adults with attention-deficit hyperactivity disorder: a triple-blind, randomised, controlled trial. *Lancet Psychiatry* 2017; 4: 673–84.
- Sonuga-Barke EJ, Brandeis D, Cortese S, Daley D, Ferrin M, Holtmann M, et al. Nonpharmacological interventions for ADHD: systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. *Am J Psychiatry* 2013; 170: 275–89.
- Thibault RT, Lifshitz M, Raz A. Neurofeedback or neuroplacebo? *Brain* 2017; 140: 862–4.
- Thibault RT, Lifshitz M, Raz A. The climate of neurofeedback: scientific rigour and the perils of ideology. *Brain* 2018; 141: e11.
- Thibault RT, Raz A. When can neurofeedback join the clinical armamentarium? *Lancet Psychiatry* 2016; 3: 497–8.
- 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.
- Witte M, Kober SE, Wood G. Noisy but not placebo: defining metrics for effects of neurofeedback. *Brain* 2018; 141: e40.