Consensus on the reporting and experimental design of clinical and cognitive-behavioural neurofeedback studies (CRED-nf checklist)


*Authors contributed equally. All middle authors are listed in reverse alphabetical order.

Draft dated: 23 Jan 2019

After a protracted history, neurofeedback has begun to attract the attention and scrutiny of the scientific and medical mainstream (Kamiya, 2011; Linden, 2014; Sitaram et al., 2017). A debate now centres on the extent to which neurofeedback alters brain function and behaviour, and the mechanisms through which neurofeedback operates (e.g., neurofeedback-specific versus nonspecific). A series of correspondences in *Lancet Psychiatry* (Micoulaud-Franchi & Fovet, 2016; Pigott et al., 2017; Schönenberg et al., 2017b, 2017a, Thibault & Raz, 2016a, 2016b) and *Brain* (Fovet et al., 2017; Schabus, 2017, 2018; Schabus et al., 2017; Thibault, Lifshitz, & Raz, 2017a, 2017b; Witte, Kober, & Wood, 2018) discusses the theoretical arguments and empirical data backing the involvement of these two mechanisms.

The apparent controversy that the correspondence letters present stems from a well-known phenomenon in neuropsychology: that multiple components can drive the benefits of a treatment (Campbell & Stanley, 1983; Enriquez-Geppert, Huster, & Herrmann, 2013). We depict this hypothesized multi-component model for the context of neurofeedback in Figure 1. We divide the mechanisms driving experimental outcomes into five bins: neurofeedback-specific (related to training a target neurophysiological variable), neurofeedback-nonspecific (dependent on the
neurofeedback context, but independent from the act of controlling a particular brain signal), general nonspecific (including the common benefits of cognitive training as well as psychosocial influences, such as placebo responding), repetition related (e.g., test re-test improvement), and natural (e.g., spontaneous remission, cognitive development) (Micoulaud-Franchi & Fovet, 2018).

Evidence for putatively causal, neurofeedback-specific mechanisms relies on our knowledge of the physiological basis of neural activity and its relevance to cognition (for a review of neurofeedback mechanisms see Ros et al., 2014 and Sitaram et al., 2017). For example, the association between neural activity and cognition in animals (Babapoor-Farrokhran, Vinck, Womelsdorf, & Everling, 2017; Cao et al., 2016) suggests that self-regulation of brain circuits can alter behaviour and cognition. A number of neurofeedback experiments in animals (Schafer & Moore, 2011; Sterman, Howe, & Macdonald, 1970), and humans (e.g., Watanabe et al. 2017; Young et al., 2017) further support this view. Evidence suggesting that mechanisms other than neurofeedback-specific factors account for the effects of neurofeedback come from a number of recent studies and reviews that find comparable benefits between participants who receive veritable neurofeedback from their own brain and those who observe a sham-neurofeedback signal unrelated to their neural activity of interest (e.g., Schabus et al., 2017; Schönenberg et al., 2017b; Thibault & Raz, 2017).

![Figure 1. Multiple mechanisms drive the effects of neurofeedback training. Neurofeedback participants may benefit from: (1) the specific neurophysiological process of training a particular brain signal, depicted in green. Nonspecific factors, including (2) those](image-url)


unique to the neurofeedback environment (e.g. trainer-participant interaction in a neurotechnology context), depicted in dark blue; and (3) those that are common across interventions (e.g., all other benefits from engaging in a form of cognitive training as well as the psychosocial and placebo mechanisms related to participating in an experiment), depicted in light blue. (4) Repetition-related effects, depicted in purple. (5) Natural effects, which can be positive (e.g., cognitive development in childhood) or negative (e.g., cognitive decline in older age), depicted in orange. These mechanisms may interact synergistically to create a greater overall effect, interact antagonistically to lessen the total benefit, or combine additively (see Finnerup, Sindrup, & Jensen, 2010; Rothman, 1974 for a discussion of this topic). By including control groups, carefully designing experiments, and measuring both brain activity and behaviour, researchers can better estimate the contribution coming from each of these mechanisms.

To advance the field of neurofeedback, scientists can benefit from designing future studies with the methodological rigour capable of disentangling the various mechanisms driving the effects of neurofeedback. As authors of the correspondence letters, alongside other researchers active in the field, we propose a standardized checklist outlining best practices in the experimental design and reporting of neurofeedback studies. We believe that widespread adoption of this checklist will help advance our scientific understanding of how neurofeedback affects brain function and behaviour.

Objectives of the checklist

This checklist is intended to encourage robust experimental design and clear reporting for clinical and cognitive-behavioural neurofeedback experiments (for a methodological review see Enriquez-Geppert et al., 2017). Because all neurofeedback aims to train brain activity, these guidelines generalize across EEG (electroencephalography), MEG (magnetoencephalography), fMRI (functional magnetic resonance imaging), fNIRS (functional near infrared spectroscopy) and other neurofeedback modalities. The checklist focuses mainly on aspects unique to the neurofeedback context (as general standards for each imaging modality already exist, e.g., Gross et al., 2013; Nichols et al., 2017; Pernet et al., 2018). It serves as a complement, rather than alternative, to the Consolidated Standards of Reporting Trials (CONSORT) guidelines (Schulz, Altman, & Moher, 2010). When submitting neurofeedback results for publication, we encourage researchers to include the checklist below and fill in the boxes with the page number identifying where in their manuscript each point is addressed. This checklist does not aim to inhibit the exploration of novel directions in neurofeedback research. On the contrary, it advocates robust designs and clear reporting to promote informed research decisions that can effectively build upon previous work. These guidelines are a first iteration. As neurofeedback research progresses, we invite the community to provide comments for improving this checklist. We hope these guidelines will help disentangle the relative contribution of the mechanisms outlined in Figure 1.
Consensus on the Reporting and Experimental Design of clinical and cognitive-behavioural Neurofeedback studies (CRED-nf) best practices checklist 2019*

<table>
<thead>
<tr>
<th>Domain</th>
<th>Item #</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-experiment</td>
<td>1a</td>
<td>Pre-register experimental protocol and planned analyses</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1b</td>
<td>Justify sample size</td>
<td></td>
</tr>
<tr>
<td>Control groups</td>
<td>2a</td>
<td>Employ control group(s) or control condition(s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2b</td>
<td>When leveraging experimental designs where a double-blind is possible, use a double-blind</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2c</td>
<td>Blind those who rate the outcomes, and when possible, the statisticians involved</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2d</td>
<td>Examine to what extent participants and experimenters remain blinded</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2e</td>
<td>In clinical efficacy studies, employ a standard-of-care intervention group as a benchmark for improvement</td>
<td></td>
</tr>
<tr>
<td>Control measures</td>
<td>3a</td>
<td>Collect data on psychosocial factors</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3b</td>
<td>Report whether participants were provided with a strategy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3c</td>
<td>Report the strategies participants used</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3d</td>
<td>Report methods used for online-data processing and artifact correction</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3e</td>
<td>Report condition and group effects for artifacts</td>
<td></td>
</tr>
<tr>
<td>Feedback specifications</td>
<td>4a</td>
<td>Report how the online-feature extraction was defined</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4b</td>
<td>Report and justify the reinforcement schedule</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4c</td>
<td>Report the feedback modality and content</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4d</td>
<td>Collect and report all brain activity variable(s) and/or contrasts used for feedback, as displayed to experimental participants</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4e</td>
<td>Report the hardware and software used</td>
<td></td>
</tr>
<tr>
<td>Outcome measures</td>
<td>5a</td>
<td>Report neurofeedback regulation success based on the feedback signal</td>
<td></td>
</tr>
<tr>
<td>Brain</td>
<td>5b</td>
<td>Plot within-session and between-session regulation blocks of feedback variable(s), as well as pre-to-post resting baselines or contrasts</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5c</td>
<td>Statistically compare the experimental condition/group to the control condition(s)/group(s) (not only each group to baseline measures)</td>
<td></td>
</tr>
<tr>
<td>Behaviour</td>
<td>6a</td>
<td>Include measures of clinical or behavioural significance, defined a priori, and describe whether they were reached</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6b</td>
<td>Run correlational analyses between regulation success and behavioural outcomes</td>
<td></td>
</tr>
<tr>
<td>Data storage</td>
<td>7a</td>
<td>Upload all materials, analysis scripts, code, and raw data used for analyses, as well as final values, to an open access data repository, when feasible</td>
<td></td>
</tr>
</tbody>
</table>

*Darker shaded boxes represent Essential checklist items; lightly shaded boxes represent Encouraged checklist items. If a checklist item is “not addressed” in the manuscript, enter “N” in place of the page number. We recommend using this checklist in conjunction with the CRED-nf article, which explains the motivation behind this checklist and provides details regarding many of the checklist items.
Description of checklist items

Pre-experiment

Item 1a. Pre-register experimental protocol and planned analyses
Pre-register, for example, on a platform such as www.osf.io, as a randomized controlled trial (RCT) on ClinicalTrials.gov or the European Union Clinical Trials Register (EUCTR), or by submitting a registered report (see www.cos.io/rr for information concerning registered reports). Essential for clinical and replication studies, encouraged for others. Clearly label primary and secondary outcome variables. Indicate the number, frequency, and duration of neurofeedback sessions. In the publication, report which analyses were pre-registered and which were exploratory.

Item 1b. Justify sample size
Justify the sample size with a power analysis based on the smallest effect size of interest (e.g., minimal clinically important differences, see item 6a) or another method (e.g., Bayesian sequential sampling). Otherwise, label the experiment as a pilot study. If the pre-registered sample size is not met, state so. We do not recommend selecting a sample size based on an expected effect size derived from previous literature. Due to the publication bias that remains common across research fields, this practice can leave experiments underpowered (Albers & Lakens, 2018; Algermissen & Mehler, 2018).

Control groups

Item 2a. Employ control group(s) or control condition(s)
Employ a control group (between subjects) or control condition (within subjects). This could include a placebo-control (e.g., sham-neurofeedback, neurofeedback from a largely unrelated brain signal, or inversing the neurofeedback reward contingency) or another active non-neurofeedback control (e.g. a similar type of computerized cognitive training, biofeedback, or medication). See Sorger et al. (2018) for an in-depth review of control groups in neurofeedback research. Consider the potential for, and report any, adverse effects in both the experimental and control groups.

Item 2b. When leveraging experimental designs where a double-blind is possible, use a double-blind
For example, in experiments with a placebo-neurofeedback control group or within participant control conditions.

Item 2c. Blind those who rate the outcomes, and when possible, the statisticians involved
Indicate which individuals were blinded, how blinding was achieved and whether the blind was maintained.
Item 2d. Examine to what extent participants and experimenters remain blinded
For an overview on reporting whether blinding was successful, see Kolahi, Bang, & Park (2009).

Item 2e. In clinical efficacy studies, employ a standard-of-care intervention group as a benchmark for improvement
This design helps establish whether neurofeedback is superior to, or at least non-inferior to, standard treatments.

Control measures

Item 3a. Collect data on psychosocial factors
For example, participant motivation, treatment expectation, effort exerted, and subjective sense of success.

Item 3b. Report whether participants were provided with a strategy
If strategies were provided, report the details of the strategies.

Item 3c. Report the strategies participants used

Item 3d. Report methods used for online-data processing and artifact correction
For example, detection and rejection/correction of ocular and muscular artifacts (EEG, MEG), and of cardio-respiratory and movement artifacts (fMRI).

Item 3e. Report condition and group effects for artifacts
Report condition and group effects for the artifacts detailed for item 3d (to test whether artifacts are more prevalent in certain participants and conditions).

Feedback specifications

Item 4a. Report how the online-feature extraction was defined
For example, a frequency band, frequency band ratio, single region of interest, or functional connectivity measure. Was it individualized or fixed across all participants? How was it extracted (e.g., number and location of electrodes)?

Item 4b. Report and justify the reinforcement schedule
For example, justify the reinforcement schedule, or the feedback threshold criteria, in relation to existing neurofeedback literature and practice. Report how the feedback was given (e.g. continuous or periodic, proportional or binary). Report the amount of reward (e.g., percentage) per subject and across subjects.

Item 4c. Report the feedback modality and content
Identify the feedback modality (e.g., visual, auditory, tactile, proprioceptive), and the feedback format (e.g., video clip, simple graphic, melody, tone).

**Item 4d. Collect and report all brain activity variable(s) and/or contrasts used for feedback, as displayed to experimental participants**

Time points may include (i) a pre-training baseline, (ii) rest blocks, (iii) training blocks, (iv) a post-training baseline, (v) transfer run(s) without neurofeedback, and (vi) long-term follow-up. Essential (ii, iii), encouraged (i, iv, v, vi). Report the relevant units.

**Item 4e. Report the hardware and software used**

Include the versions.

**Outcome measures (brain)**

**Item 5a. Report neurofeedback regulation success based on the feedback signal**

Identify the baseline or contrast used (e.g., subject specific data from a previous session, reference data based on averaged data from a normative group). Identify the comparator run (e.g., training run or transfer run).

**Item 5b. Plot within-session and between-session regulation blocks of feedback variable(s), as well as pre-to-post resting baselines or contrasts**

Plotting the session course by comparing the session beginning, middle, and end (for instance, by arbitrarily dividing sessions to segments or using session blocks) allows the assessment of within-session dynamics. Between-session comparisons allow the assessment of the whole training course on a temporally more abstract level.

**Item 5c. Statistically compare the experimental condition/group to the control condition(s)/group(s) (not only each group to baseline measures)**

Comparing experimental and control groups/conditions to their respective baselines, but not to each other fails to test whether the experimental intervention outperforms the control intervention(s) (Nieuwenhuis, Forstmann, & Wagenmakers, 2011).

**Outcome measures (behaviour)**

**Item 6a. Include measures of clinical or behavioural significance, defined a priori, and describe whether they were reached**

For example, by using minimal clinically important differences (MCIDs) to establish the magnitude of an effect to interpret as clinically meaningful (see Engel, Beaton, & Touma, 2018 for overview on establishing MCID values). Moreover, collect data on acceptability, safety, and adverse effects. In this paper, we are using
the term behaviour in the broad sense to encompass all non-physiological measures, including self-reports.

*Item 6b. Run correlational analyses between regulation success and behavioural outcomes*

**Data storage**

*Item 7a. Upload all materials, analysis scripts, code, and raw data used for analyses, as well as final values, to an open access data repository, when feasible*
Contributorship statement

Tomas Ros (TR), Stefanie Enriquez-Geppert (SEG), and Robert T. Thibault (RTT) developed the idea for a checklist of this type. RTT prepared an initial rough draft. TR, SEG, and RTT worked together to produce a complete first draft. Kymberly Young, James Sulzer, Surjo Soekadar, Ranganatha Sitaram, Michael Schönenberg, Frank Scharnowski, Jean-Arthur Micoulaud-Franchi, David M. A. Mehler, Joel Lubar, David E. J. Linden, Rene J. Huster, John Gruzelier, Thomas Fovet, Niels Birbaumer, and Martijn Arns provided comments on the first complete draft. TR, SEG, and RTT worked together to implement the comments and produce a second draft. All authors provided comments on the second draft. TR, SEG, and RTT worked together to implement the comments and produce the final version.

References:


Algermissen, J., & Mehler, D. M. A. (2018). May the power be with you: are there highly powered studies in neuroscience, and how can we get more of them? Journal of Neurophysiology. https://doi.org/10.1152/jn.00765.2017


