Doubting the double-blind: Introducing a questionnaire for awareness of experimental purposes in neurofeedback studies

Timo L. Kvamme a, b, *, Mesud Sarmanlu a, Morten Overgaard a

a Cognitive Neuroscience Research Unit, CFIN/MINDLab, Aarhus University, Aarhus, Denmark
b Centre for Alcohol and Drug Research, Aarhus University, Aarhus, Denmark

ARTICLE INFO

Keywords:
Neurofeedback
Awareness
Contingency Awareness
Consciousness
Metacognition

ABSTRACT

Double-blind subjects to the experiment’s purpose is an important standard in neurofeedback studies. However, it is difficult to provide evidence that humans are entirely unaware of certain information. This study used insights from consciousness studies and neurophenomenology to develop a contingency awareness questionnaire for neurofeedback. We assessed whether participants had an awareness of experimental purposes to manipulate their attention and multisensory perception. A subset of subjects (5 out of 20) gained a degree of awareness of experimental purposes as evidenced by their correct guess about the purposes of the experiment to affect their attention and multisensory perceptions specific to their double-blinded group assignment. The results warrant replication before they are applied to clinical neurofeedback studies, given the considerable time taken to perform the questionnaire (~25 min). We discuss the strengths and limitations of our contingency awareness questionnaire and the growing appeal of the double-blinded standard in clinical neurofeedback studies.

1. Introduction

Reviews of neurofeedback studies generally argue that employing a “double-blind” design is an advisable experimental standard (Sitaram et al., 2017; Sorger et al., 2019; Thibault et al., 2016). “Double-blind” is typically defined as the subject and experimenter being unaware of which treatment condition or placebo-control condition the subject is allocated (Sorger et al., 2019; Thibault et al., 2016). The importance of the double-blind design is accentuated by a recent discussion on whether double-blinded designs reveal that clinical neurofeedback provides no larger effect than placebo-control, i.e., “sham-neurofeedback” (Arns et al., 2009; Micoulaud-Franchi, 2014; Schabus, 2017; Schabus et al., 2017; Schönenberg et al., 2017; Sonuga-Barke et al., 2014; Thibault et al., 2017). An important question then becomes; is the double-blind even possible? Or are subjects aware of the purposes of the neurofeedback experiment?

It’s been argued that sham-neurofeedback is not feasible since subjects can detect false feedback (Geladé et al., 2017; Kotchoubey et al., 2001; Lansbergen et al., 2011; Sitaram et al., 2017). However, the double-blind experimental standard is still regarded as “essential” in a recent consensus checklist aimed at improving the quality of neurofeedback studies (Ros et al., 2020). Fortunately, methods have been developed which employ neurofeedback implicitly, where the subject is claimed to have no awareness of experimental purposes when performing the neurofeedback (Birbaumer et al., 2013; Ramot et al., 2016; Shibata et al., 2011, 2019; Taschereau-Dumouchel et al., 2018b, 2018a; Watanabe et al., 2017) (for a theoretical and critical review of the role of awareness in
neurofeedback, see (Muñoz-Moldes and Cleeremans, 2020)). Implicit neurofeedback is intriguing and controversial because it runs counter to the first neurofeedback study, which showed a link between awareness (e.g. discrimination) of being in a certain brain state and control of the neurofeedback-derived brain activity (Kamiya, 1969, 2011) (replicated here: Frederick, 2012; Frederick et al., 2016, 2019)). Implicit neurofeedback might be advantageous in certain situations because it can be questioned whether providing an explicit cognitive strategy is always advisable (deCharms et al., 2005; Oblak et al., 2017; Scharnowski et al., 2012; Scheinost et al., 2013; Sulzer et al., 2013a). Cognitive strategies have, for instance, been found to be inconsistent across subjects (Kübler et al., 2001; Neumann and Birbaumer, 2003), uncorrelated with neurofeedback performance (Kober et al., 2013), and remain constant despite the reversal of reward contingencies (Ramot et al., 2016; Sniatchkin et al., 2000). Moreover, as proponents of the implicit neurofeedback approach claim, learning is possible even when the subject has no explicit knowledge of the relation between the neural activity and the feedback (Amano et al., 2016; Ramot et al., 2016; Shibata et al., 2011). The addition of the implicit neurofeedback approach could thus become necessary for overcoming the criticism that explicit neurofeedback paradigms are contaminated by placebo, demand characteristics and other biases (Thibault et al., 2016). In short, implicit neurofeedback allows neurofeedback researchers to live up to the principle of clinical RCTs by providing feedback in a double-blinded manner (Tischarceau-Dumouchel et al., 2018b).

As the double-blinded procedure is a vital experimental standard, the methodology of assessing unawareness should be rigorously validated. However, it is difficult to establish that humans are without awareness of specific information conclusively. The common problem is that subjects are not always adequately incentivized to provide extensive details about their experiences (Muñoz-Moldes and Cleeremans, 2020; Timmermans and Cleeremans, 2015). It is thus essential to develop questionnaire methods that further incentivize participants’ reports of their awareness of their strategies and awareness of the experimental purposes to examine whether subjects remain blind in neurofeedback studies (Ros et al., 2020). Arguably, the development of assessment methods for detecting awareness of experimental purposes (AoEP) in neurofeedback studies would benefit from being developed in a non-clinical population where subjects are not recruited due to a disorder that they are continually aware of and which the neurofeedback aims to alleviate (Haug et al., 2021). Moreover, since there is a consensus that subjects can detect false feedback, then the variations in awareness of experimental contingencies should be assessed with an inverse neurofeedback group (Sorger et al., 2019; Thibault et al., 2016). We now discuss two such studies; (Ramot et al., 2016; Shibata et al., 2011), which claimed that subjects were unaware of the experiment’s purpose.

In a study by Shibata et al. (2011), the subjects were shown a green feedback circle, which was linked with patterns of occipital cortex activity, and were told to “somehow regulate activity in the posterior part of the brain to make the solid green disc as large as possible” (Shibata et al., 2011). The subjects were thus not cued to use any specific cognitive strategy but could instead find their own way to gain control over the feedback display (Shibata et al., 2019; Watanabe et al., 2017). As reported by Shibata et al. (2011), when told afterwards that the experiment intended to improve the perception of line orientations, subjects could not guess the specific line orientation that they were trained on even though the neurofeedback was successfully improved their perception of that specific line orientation. In other words, their guesses were random.

In a study by Ramot et al. (2016) the subjects were not told that they were receiving neurofeedback but were instead told that the experiment was meant to test their reactions to different rewarding stimuli (Ramot et al., 2016). Unbeknownst to the subjects, the rewarding stimuli were contingent on the ratio of brain activity between two category-selective regions, the fusiform face area (implicated in face-related imagery) and the parahippocampal place area (implicated in places/house-related imagery). In subsequent post-scan interviews, when told that the experiment was a neurofeedback study relying on activity from two areas of their brain, the subjects reported having no notion what was driving the feedback. When given a five-alternative forced choice where subjects were asked to select a cognitive function related to the reward stimuli, the subjects’ guesses were random. The subjects were allowed to select among the two correct options; face-related imagery and places/house-related imagery, as well as three decoy options; abstract visual forms, language, and body/motion-related imagery.

Shibata et al. (2011) and Ramot et al. (2016) represent the state-of-the-art neurofeedback procedures. However, improving the method of assessing awareness in neurofeedback may benefit future studies. In the two studies, it is unclear how subjects were queried about their strategy for performing the neurofeedback. How were the particular questions precisely phrased? And what specific information was provided, and how was it presented to subjects before awareness of purposes was assessed? Moreover, the requirement that subjects designate the specific cognitive function which correlated with the brain activity they regulated, such as face-related visual processing or orientation of lines, seems challenging. Such insight into the specific cognitive functions and how they are distinguished from other functions seems more befitting a highly self-reflective neuroscientist than a layperson participating in a brain-scanning study.

To improve methods for assessing awareness of purposes in neurofeedback studies, we suggest that insights from neurophenomenology are employed, where questions fit more appropriately with the colloquial language typically employed by subjects (Bagdasaryan and Le Van Quyen, 2013; Micoulaud-Franchi et al., 2014; Varela, 1996). We hypothesized that residual degrees of awareness might be detectable if subjects’ experiences are extracted using “the elicitation interview” method, where the experimenter rigorously guides the subject into an open introspective state to gather precise verbatim “conscientizable” dimensions of their experiences that are otherwise hidden in “pre-reflective” experiences (Petitmengin and Lachaux, 2013; Petitmengin et al., 2013; Vermersch, 1995). Our prediction was not that subjects could develop an awareness of the experimental purposes to alter a cognitive
process as specific as the perception of line orientations or specific face/place-related imagery presented by Shibata et al. (2011) and Ramot et al. (2016). Instead, we predict that if awareness measures are sensitive enough, it might be possible to detect the awareness of experimental purposes in neurofeedback studies aiming to train cognitive processes that humans routinely have an awareness of, such as their attention.

To investigate this we developed a detailed contingency awareness questionnaire to assess subjects’ awareness of experimental purposes following a MEG-neurofeedback paradigm commonly used to train attention by lateralizing alpha activity (Bagherzadeh et al., 2020; Okazaki et al., 2015; Schneider et al., 2020a). In a companion paper (Kvamme et al., 2022), which has been published elsewhere, we investigated whether this type of alpha lateralization training, which has been found to increase attention to one visual hemifield, could also extend to influence multisensory perceptual experiences in the sound-induced flash illusion (Cecere et al., 2015; Lange et al., 2013, 2014; Rohe et al., 2019). We, therefore, reasoned that subjects might experience an increase in their attention towards the trained hemifield and that they might develop an awareness of the purposes of the experiment to affect their multisensory perceptions. The companion paper (Kvamme et al., 2022) describes the neurofeedback parameters, neuroimaging analysis, and behavioral effects on multisensory perception. The experiment is the same, but the research objectives of the two papers are different. In sum, many neurofeedback studies underline the necessity of the double-blinded design; however, the practice of assessing whether awareness of purposes developed for subjects is often neglected or underdeveloped. We developed a detailed contingency awareness questionnaire to improve this practice to assess whether subjects became aware of experimental purposes to bias their attention and multisensory experiences to one visual field following alpha lateralization neurofeedback.

2. Materials and methods

2.1. Subject groups

We recruited 20 healthy volunteers (7 males) with normal or corrected-to-normal vision; participants had an average age of 26.1 ± 6.4 years). The local ethics committee approved the study in accordance with Danish Law. Subjects participating in the study were not incentivized monetarily based on their performance during neurofeedback. The subjects were, in a double-blinded manner, randomized into two groups. The L-NFB group (n = 10) was trained to increase alpha power in the left relative to the right parietal sensors, whereas the R-NFB (n = 10) was trained in the opposite direction. The sample size of the present study (N = 20) was determined based on the expected neurofeedback effect of the companion paper (Kvamme et al., 2022) based on prior work (Bagherzadeh et al., 2020).

2.2. Experimental design

The MEG experiment lasted approximately 70 min and consisted of three phases. The first phase (~20 min) was a pre-test of the multisensory sound-induced flash illusion (SIFI) paradigm. The second phase (~25 min) was the neurofeedback training phase, followed by the third phase (~20 min), post-testing the SIFI paradigm identical to the first phase. Immediately after the MEG scan, subjects were questioned regarding their strategy for and experience with performing neurofeedback (~10 min) called the “Immediate Post-Scan Interview”. Subjects were part of a more extensive study on multisensory integration, and they participated in several other experiments lasting ~60 min after the MEG scan. Finally, subjects performed a post-scan interview and questionnaire (~20–25 min), called the “One-Hour Post-Scan Interview and Questionnaire” regarding their strategy for and experience with performing neurofeedback and their assessment of the purpose of the experiment (see Fig. 1).

2.3. Neurofeedback data

Real-time MEG signals were acquired in a magnetically shielded room using an Elekta Neuromag TRIUX (Stockholm, Sweden) 306-channel system (204 planar gradiometers, 102 magnetometers), at a sampling rate of 1 kHz. Eye movements were recorded using an eye-tracking device tracking gaze position at 1000 Hz (EyeLink1000, SR Research, Ontario, Canada). The rtMEG software segmented MEG data into 500 ms blocks (Sudre et al., 2011). To reduce the environmental magnetic noise and perform signal space separation (SSS), we used a custom “maxwell_filter” function written in mne-python (Gramfort et al., 2013). Estimates of alpha power (8–12 Hz) for each 500-ms data segment were performed using the welch method (single Hann window) taken from the parietal sensors (we excluded the six anterior parietal sensor triplets) (Welch, 1967). Only gradiometer sensors (20 in each hemisphere) were used because magnetometers are too sensitive to environmental noise (Hämäläinen et al., 1993).

An index of alpha asymmetry index (AAI) was calculated as the log10(αL/αR), where αL refers to the mean of alpha estimates from ipsilateral, and αR refers to the mean of alpha estimates from contralateral parietal sensors. αL and αR were different for the L-NFB and R-NFB group (e.g., αL was left for the L-NFB group). We computed a standardized AAI (“zAAI”) score in real-time by correcting individual differences in baseline alpha asymmetry between the two hemispheres. In the first neurofeedback, block standardization was performed using the mean and standard deviation of AAI scores calculated from data segments in the pre-stimulus period (500 ms before flash onset) from trials in the pre-test of SIFI. In the subsequent three neurofeedback blocks, standardization was performed using a 20 s reference recording taken before blocks. A cumulative distribution function (CDF) was used to convert the zAAI to a value...
between 0 and 1, where values over 0.5 represented positive zAAI scores (Okazaki et al., 2015). The derived CDF value determined the visibility of a Gabor pattern presented to subjects. A negative zAAI score resulted in 0% visibility, whereas a positive zAAI resulted in visibility from 50% to 100%.

In the latter offline analyses, we used power spectral density (PSD) maps to calculate the average of trials during the last 4 s (-4 to 0 s) with respect to the last epoch in a given neurofeedback trial) and a 20 s rest period before training as a baseline, thus deriving baseline corrected (Δ) PSDs for each parietal hemisphere (Kvamme et al., 2022). We could calculate a baseline corrected asymmetry index of parietal cortex alpha power from this. We used these offline source-based MEG data to assess whether the alpha asymmetry neurofeedback training induced asymmetric differences in parietal cortex alpha power during neurofeedback for each trained participant (Kvamme et al., 2022).

---

**Fig. 1. Experimental Procedure Flowchart.** The flow of participants through different experimental procedures. Participants completed a pre-test of the sound-induced flash illusion (SIFI). Participants were then allocated to an L-NFB group where attention was increased to the left hemifield and an R-NFB group where attention was increased to the right hemifield. N = 10 in each group. The neurofeedback training for both groups was intended to decrease contralateral alpha activity. The study’s main hypothesis was that increased attention would result in additional illusory flash percepts in the trained hemifield for both groups in the SIFI post-test. In the immediate post-scan interview, participants were queried as to their strategy for neurofeedback. Participants then completed other multisensory tasks for approximately 60 min. In the one-hour post-scan interview and questionnaire session, participants were queried about their experience of strategy and attention during neurofeedback. Participants then completed the contingency awareness questionnaire where they were queried about audio/visual strategy, their assessment of their own illusion probability, assessment of their confidence, their attentional focus during neurofeedback, the confidence token game, and whether the experiment intended to change their attention to the left or right.
2.4. Neurofeedback task

Subjects were instructed to fixate at a central fixation cross. When the cross turned black, subjects were required to use “mental effort” to increase the visibility of the Gabor pattern. Before entering the scanner, subjects were informed that the feedback was related to their ongoing brain activity, and a slight delay in the feedback would occur. The neurofeedback phase was composed of 68 trials split into four blocks. A trial began with 5 s of rest, followed by ten epochs of neural feedback. Feedback was calculated based on the recent 500 ms of data and shown to the subject (delay = 187 ms). We omitted 200 ms following the visual display update to reduce the influence of the visual-evoked response on the following data segment (Okazaki et al., 2015). A single neurofeedback trial thus took ~9 s. A white fixation cross was presented for an inter-trial interval (ITI) between 3 and 5 s following a trial. We used the eye tracker for gaze-contingent correction of subjects’ gaze behaviors to ensure subjects’ gazes remained stable. Live video recording was monitored to ensure that subjects were not closing either eye for long periods. Finally, subjects were required to press either the left or right response box (i.e., a text appeared, e.g., “press left key”) at the end of each trial. Following each neurofeedback block, subjects rated their sense of control of the neurofeedback display using the 6-point sense of control scale (SCS) (Dong et al., 2015). Subjects were encouraged to try their best following their SCS rating.

2.5. Pre- and Post-Test multisensory perception task

In the pre and post-test phases of the MEG, experiment subjects performed a multisensory perception task called the sound-induced flash illusion (SIFI) (Shams et al., 2000, 2002). The sound-induced flash illusion is a visual illusion where a single flash presented in the visual periphery together with more than one auditory beep occasionally leads to the perception of more flashes. Since low alpha-band activity in the posterior cortices is correlated with increased sound-induced flash perceptions and change in measures of attention, we hypothesized that the multisensory flash reports could increase in the contralateral hemisphere to which alpha was down-regulated with neurofeedback (Bagherzadeh et al., 2020; Kaiser et al., 2019; Kramme et al., 2022; Lange et al., 2013; Rohe et al., 2019). The experiment’s purpose was thus that subjects located in either group (L-NFB) and (R-NFB) would see increases in the experience of sound-induced flashes in the left and right hemifield, respectively. To measure this, we employed a version of the SIFI where flashes appeared on both sides/hemifields. We used 208 trials per phase with three illusion trial conditions (2b1f, 3b1f, and 4b1f), two control conditions (1b1f and 2b1f), and one catch trial condition (1b0f) where the index number represents the auditory beeps (b) and the visual flashes (f). Gaze-contingent eye-tracking was employed to ensure that each subject’s gaze remained on the centrally presented fixation cross. Auditory beeps were presented through both left and right earplugs and comprised 7 ms 1 kHz sine waves with 3 ms onset and offset ramps. Visual flashes were presented for 16.7 ms on either left or right side at 6.7° eccentricity concerning the central fixation cross and consisted of white circular disks subtending a visual angle of 1.6°. The onset of auditory stimuli occurred 7.6 ms before the flash, and the delay between the first and second flash was 66.7 ms (Stiles et al., 2018; Yun et al., 2020). Subjects indicated how many flashes they experienced using a response box.

2.6. Immediate Post-Scan interview

Subjects performed an initial short 5-minute “immediate post-scan interview” after the post-test SIFI task. As subjects exited the scanner and removed EMG-electrodes, they were recorded. Recordings were transcribed and translated and put in the following online repository (link): In the immediate interview, subjects were questioned about what the neurofeedback was like for them. Every participant was repeatedly asked which strategy they used to perform the neurofeedback (making the black lines clearer) until they no longer reported any new strategy. Initially, we asked participants (ID 1 and ID 2) if they noticed a change in their attention during the immediate post-scan interview. This question was later moved to the One-Hour Post-Scan Contingency Awareness Questionnaire.

2.7. One-Hour Post-Scan contingency awareness questionnaire

Subjects performed a more detailed ~ 20–25-minute questionnaire on contingency awareness which was delayed after performing other multisensory tasks without scanning (~60 min). We reasoned that an immediate interview that probed subjects’ strategies would help subjects recall their strategies and experiences performing neurofeedback due to the delay caused by performing the post-test. Moreover, we reasoned that the distractive element of performing other multisensory tasks before the final detailed contingency awareness questionnaire would serve as an incubation period for subjects, which would help their recall (Gilhooly, 2016; Sio and Ormerod, 2009).

In the one-hour post-scan interview and questionnaire (see Fig. 1 and the supplementary material for the entire questionnaire), we interviewed subjects with questions related to subject’s perceptions of the tasks, strategies, and awareness of the experiment’s purposes (AoEP). The questionnaire contained a combination of open-ended questions where subjects were allowed to provide a free-form answer and closed-ended yes-or-no questions. The questionnaire featured “decoy questions” or “lures” meant to raise the possibilities of alternative solutions to the final question on the experiment’s purpose than the correct one. Initially, subjects were asked if they had any new recollection of their employed strategies during neurofeedback and the strategies they discussed in the immediate ~ 5 min following the scan. Participants (with ID 1 and ID 2) were asked in the immediate post-scan interview whether they noticed a change in their attention following neurofeedback, while the latter (N = 16) participants were asked in the one-hour post-scan contingency awareness questionnaire. Due to an error, this was not asked for two participants (ID 11 and ID 15). Subjects were asked whether they noticed a change in their attention in an open-ended question. Then subjects were asked whether their strategies were mainly visual or...
Subjects were then told about the sound-induced flash illusion (the multisensory perception task they were tested on before and after neurofeedback). They were asked to estimate how often they experienced the illusion in the task and their confidence in their estimates of flashes across the two phases. Subjects were then asked if they attended more to the left or the right. Participants also reported how often they thought the probability of the small probe would appear on either side using VAS (the probe appeared equally on either side during the neurofeedback). Subjects then completed the so-called Confidence Token Game (CTG) (Kvamme et al., 2019) modified for this specific neurofeedback paradigm. In the CTG, the subject is given six red rectangular cardboard tokens that represent their total confidence. The subject was then faced with a 6 row by 7 column grid where the grid columns contained behaviors that the neurofeedback training was supposedly meant to change. The behaviors were 1) experience more flashes, 2) experience fewer flashes, 3) report faster, 4) report slower, 5) experience more left-sided flashes (correct for L-NFB subjects), 6) experience more right-sided flashes (correct for the other R-NFB subjects) and 7) the feedback was random (see Fig. 2).

Subjects were required to place all six tokens within one or several columns of their choice. The tokens represented which behavior(s) the subjects believed the neurofeedback training was meant to change their behavior from pre-test to post-test. The subjects were told that the experimenter was unaware of their group assignment but that the experimenter knew that a minimum of one behavior was correct and that a minimum of one was incorrect. Subjects were also given seven different neurocognitive explanations of how neurofeedback could have attempted to influence each behavior. An attempt was made to link the neurocognitive explanations to prior questions in the contingency awareness questionnaire. For instance, subjects were told that confidence (which they had already been interviewed on) could be measured in the brain and trained, affecting their reaction time. The concept of control groups in neurofeedback studies where the feedback is random was also described. Subjects who did not believe that the feedback was meant to change any behaviors were encouraged to place tokens in the random category. Subjects rated their confidence in their placement and were told to relocate tokens until the placement represented their most confident guess. Finally, subjects were asked whether the experiment intended to change their attention to the left or right side using a VAS scale. The final question was added after the first subject and was thus not available for this subject. Subjects and experimenters wore facemasks during the interview due to COVID-19 guidelines (N = 16, as the guidelines changed during the experiment). The questionnaire, the interview transcripts (translated to English for Danish speakers), and the questionnaire data are provided in the following online repository (link).

**Fig. 2. Confidence Token Game.** Subjects placed six red tokens, which represented their confidence, within seven different columns that represented potential behaviors that the neurofeedback might have had the purpose of changing their behavior from before to after neurofeedback. The grey-colored columns were decoy options; the blue-colored column was correct for subjects in the L-NFB group, while the red-colored column was correct for subjects in the R-NFB group. In the table presented to the participants, the columns were not colored (see the supplementary material for the exactly presented questionnaire). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
3. Analysis and results

3.1. Open-ended interview questions

None of the participants (N = 20) indicated that their strategies included visually attending to their trained side in the open-ended interview questions. The most predominant strategies were; Focusing on the black lines (N = 12), Focusing eyes (N = 10), Focusing on the fixation cross (N = 6), Focusing away from the black lines (N = 4), Focusing on an after image (N = 4), Using an inner voice (N = 4) and relaxing eyes (N = 3). One participant reported looking across the black lines as a strategy but without a preference for left/right versus up/down (ID = 17). One participant reported looking down as a strategy (ID = 4). One participant noticed experiencing more illusions following neurofeedback (ID = 3), but not more to any side. One participant in the L-NFB group reported hearing heart pulses and attending more to the left ear (ID = 5). Interestingly, this participant reported in the one-hour post-scan interview experienced that the fixation cross in one of the multisensory tasks was shifted to the left. In the initial open-ended questions, none of the participants
noticed a change in their attention to their trained side when asked if they (N = 18, two lacking due to errors) experienced a change in their attention due to neurofeedback. See the supplementary material for full transcripts and a summary of strategies (link).

3.2. Main awareness questions of interest

The following questionnaire from the one-hour post-scan contained three main awareness questions of interest; 1): The self-reported asymmetric attention during neurofeedback, 2): The CTG question on self-reported awareness of experimental purposes (AoEP) to manipulate flash reports asymmetrically, and 3): The self-reported AoEP to manipulate attention asymmetrically. All values were corrected to positive and negative values concerning each subject’s group and to-be-trained side/hemifield. For example, in the t-NFB reporting AoEP to change attention to the left, the amount of confidence towards the left side counted as positive. In contrast, values to the right counted as negative (for the opposite group). For the second question (from the CTG), which was derived from six confidence tokens representing their total confidence (i.e., 100 %), values were multiplied by 16.66 for each token. We only considered confidence reports on the two non-decoy options (i.e., AoEP to manipulate flash reports left or right). One subject (ID = 14) was unsure about the direction and thus rated with one token for both left and right hemifield for the second question. For this subject, responses were set to zero for the second question (the original responses are kept for the supplementary analysis in the supplementary material).

We performed one-sample t-tests on each of the three questionnaire questions and performed a follow-up exploratory analysis of alpha power asymmetry using a two-sample permutation test in the R statistical computing software (Kohl, 2019; R Core Team, 2013; Valero-Mora, 2010). All one-sample t-tests were conducted against a mean of 0. For the first question, 1): The self-reported asymmetric attention during neurofeedback values had (M = 1.62, SD = 8.9) and the test statistic was not significant t(19) = 0.81, p = 0.436. For the second question, 2): The CTG question on self-reported awareness of experimental purposes (AoEP) to manipulate flash reports asymmetrically values had (M = 6.66, SD = 12.56), and the test statistic was significant t(19) = 2.37, p = 0.028. For the third question, 3): The self-reported AoEP purpose to manipulate attention asymmetrically, values had (M = 4.63, SD = 19.60) and the test statistic was not significant t(18) = 1.03, p = 0.317. As the first and third questions are conceptually similar, we performed spearman’s rank correlation and found that they were significantly correlated (R = 0.47, p = 0.04).

3.3. Subset of subjects with awareness of experimental purposes

In light of our findings that a subset of subjects (5/20), reported in the CTG that they had a degree of the AoEP to manipulate multisensory flash reports asymmetrically, we decided to focus on these AoEP-subjects. We performed a dichotomous calculation of values by assessing whether these subjects as compared to Non-AoEP subjects rated to the left, middle, or right on their AoEP on the previous question on asymmetric attention during neurofeedback and the follow-up question on AoEP to manipulate attention asymmetrically. In a qualitative and descriptive assessment of the dichotomous ratings (see Fig. 3), no clear pattern of t-NFB and R-NFB groups choosing correct when considering all subjects. However, specifically among AoEP-subjects, four guessed correctly on the first question (one guessed middle), whereas all five AoEP-subjects guessed correct on the final question.

We tested the likelihood of a random sample of 5/20 participants correctly answering correctly within the CTG (left/right). Under the null hypothesis, we could randomly permute t-NFB/R-NFB group labels for corrected values relative to each subject’s to-be-trained side/hemifield on the CTG. Repeating this procedure 10,000 times yielded empirical distributions that true values could be compared against. This test was not significant (p = 0.10, n = 5, non-parametric two-sided permutation test).

In an exploratory analysis, we tested whether AoEP-subject had a greater baseline-corrected asymmetry index of parietal cortex alpha power during neurofeedback. Although AoEP had a greater average asymmetry of alpha power during neurofeedback (M = 0.11, SD = 0.10) compared with non-AoEP-subjects (M = 0.017, SD = 0.10), a two-sample permutation t-test was not significant t(19) = 1.74, p = 0.093 (Kohl, 2019). We performed two-sample t-test comparisons between Non-AoEP subjects and AoEP subjects on confidence measures and token placement on questions in CTG than the correct options. None of the tests were significant, although a tendency for Non-AoEP subjects to assign tokens to the random option was noted t(18) = 1.83, p = 0.08. See the supplementary materials for more information.

4. Discussion

In this study, we investigated whether subjects could gain awareness of experimental purposes (AoEP) using a detailed contingency awareness questionnaire (CAQ) following a MEG-neurofeedback paradigm commonly used to alter attention (Bagherzadeh et al., 2020; Okazaki et al., 2015; Schneider et al., 2020a). We hypothesized that subjects might acquire an awareness of our attempt to alter their multisensory processing following neurofeedback if cued to different options through a detailed questionnaire that was sensitive enough to assess AoEP. None of the participants indicated a strategy of attending more to the trained side in the open-ended questions.

Our results from the detailed questionnaire following the open-ended questions demonstrate that a limited subset of subjects (5 out of 20) could correctly guess which side their multisensory confidence tokens to the correct side in the confidence token game (CTG). Instead, they only assigned one or two tokens, equaling 16.6 % or 33.3 % of their confidence to the correct side. However, none of the subjects guessed incorrectly, i.e., on their untrained hemifield. These AoEP-subjects were also correct on the previous question on their awareness of their attention during neurofeedback and on their assessment of AoEP on the following question on the purpose to alter their attention (albeit one subject did not). Our findings thus provide evidence that our detailed CAQ for neurofeedback is sensitive enough to detect that a limited subset of
subjects acquire AoEP. There are several limitations with the present study worth noting before proceeding with a comparison and discussion of previous studies. In our study, the same participants were first asked about their strategy for performing neurofeedback during an initial interview using open-ended questions. Later on, the same participants went through a detailed questionnaire to cue them to different options, thereby aiding them in potentially acquiring AoEP. To further substantiate these findings, it is necessary to also compare our approach to a simple multiple-choice questionnaire. Moreover, further studies should assess different ways and degrees of cueing or amount of decoy options, or the order of the questions between participants by employing different groups of participants for each type of questionnaire, similar to other approaches (Sandberg et al., 2010). One further option could also be to expand the number of CTG options and then gradually constrict them while testing token placement at every constriction. It might also be that the order of the questions is critical for forming an AoEP in the present study. Our data suggest that those subjects who either had or gained AoEP throughout the questionnaire rated that the AoEP of the experiment to shift their attention to one side was the same side as they had rated their AoEP to shift their multisensory perception asymmetrically.

One possibility is that subjects implicitly become aware of their group assignment of the t-NFB and R-NFB group by reading the subtle social cues from the experimenter performing the questions who knew the experiment’s purpose. However, the experimenter was blinded to the subject’s assignment to the t-NFB or R-NFB group, not to the intention that the two groups were meant to induce differences in multisensory perception. Because the questionnaire featured several questions about side/hemifield, it is not impossible that subjects somehow became aware of their assignment due to a reaction from the interviewer to the initial questions. Here, it should be noted that the double-blind was performed like other neurofeedback studies. It should also be noted that during the majority of the data collection (N = 16), the experimenter and subjects wore facemasks due to Covid-19 restrictions, which would have limited subtle facial cues.

Further studies should make sure that the experimenter who conducts the questionnaire with subjects is blinded to the purpose of the experiment. Ideally, it should also be assessed whether the experimenter can guess the assignment of subjects at different stages of the questionnaire (Ros et al., 2020). One major limitation of the present study is the low sample size (N = 20) which was determined based on the expected neurofeedback effect of the companion paper. The low sample size minimizes the degree to which the present results should influence other studies until the present results are thoroughly replicated. This is especially true given the extensive time the questionnaire takes to administer. A limitation is also that the framing of the interview questions for assessing awareness did vary somewhat from participant to participant. Another related limitation is that as an experimenter performs the questionnaire, they become more adept at querying subjects, posing the risk of discrepant querying in the beginning compared to the end of the experiment. Larger sample sizes could also remedy this. Our use of the sense of control scale (SCS) (Dong et al., 2015), which was used to query subjects about their sense of control of the neurofeedback, could be responsible for part of the development of AoEP in the subset of subjects. A consideration for future studies would be to assess the use of the SCS and control beliefs in general and their relation to the development of AoEP (Witte et al., 2013). Moreover, in our neurocognitive explanations (decoy and correct options), we did not provide references to brain areas (Ramot et al., 2016). Although this could be implemented in future studies, it should be kept in mind that the questionnaire we employed was already considerably lengthy.

Our analysis of confidence and token placement on CTG questions showed that confidence measures and other CTG measures were not significantly different between Non-AoEP subjects and AoEP subjects. However, there was a tendency toward placing tokens on the option for stating that “the feedback was random” for Non-AoEP subjects. One may argue that confidence measures make the questionnaire unnecessarily long; however, the confidence questions also facilitate more participant reflection, which is advantageous. The significant correlation between question 1: The self-reported asymmetric attention during neurofeedback, reports asymmetrically, and question 3: The self-reported AoEP to manipulate attention asymmetrically shows that the two are likely dependent. It could be interpreted that the third question is not necessary. However, we also found that one AoEP subject in the R-NFB group who answered in the middle for question 1 switched to the correct right side for question 3, suggesting that the CTG may help participants gain further AoEP.

In sum, our results provide a new perspective on assessing AoEP in neurofeedback studies. In comparison to the studies by Shibata et al. (2011) and Ramot et al. (2016), we provide a detailed and extensive contingency awareness questionnaire that may be used as a template for further studies into AoEP. As described in the introduction, an assessment of purposes in an implicit neurofeedback study was provided by Shibata et al. (2011). The most substantial evidence Shibata and colleagues (2011) provide for the implicit and double-blinded neurofeedback procedure was that when subjects were told of the purpose of improving their perception of line orientations, subjects could not guess the line orientation they were trained on. In the paper by Shibata and colleagues (2011), when discussing whether subjects became aware of the purpose of the experiment, a reference is made to a question in supplementary material regarding what subjects thought the size of the feedback disc represented. However, no such question is asked in the supplementary material of Shibata et al. (2011); instead, subjects were asked what strategy they tried to employ to increase the size of the feedback disc, which is not the same as assessing AoEP. The subjects reported various strategies but stated that they had no idea of the correct strategy. This is similar to our findings; however, because we devised a detailed contingency awareness questionnaire where we asked subjects what they thought the purpose was, we can provide evidence that a subset of subjects gained a degree of AoEP.

In the study by Ramot et al. (2016), subjects were allowed to select among the five options, where two were the correct options, and three were decoy options. We extend this methodological procedure in our CTG, although we use seven different options. In addition to the procedure presented by Ramot and colleagues (2016), we reasoned that subjects should be told that each option is equally likely to occur. To achieve this, we devised the initial questions before the CTG such that subjects were cued to think that each option could be likely. It was our purpose when providing subjects with a description of the SIFI task and the questions related to their experience that subjects were cued to think about whether they increased or decreased flash perceptions in general. We also questioned subjects about
the confidence of their SIFI reports. We used this information to raise the possibility that the purpose of the experiment might have been to affect their confidence (Cortese et al., 2016) and explained that it would consequently have affected their reaction times. Our short neurocognitive explanations of each decoy and correct options, as well as the random option, thus helped raise the probability that subjects picked the decoy questions rather than the two correct options. We argue that this provides a considerable methodological improvement compared with Ramot et al. (2016), where it is unclear what information subjects were provided before being asked about their AoEP. The complete questionnaire of the present study, all transcripts (translated to English for Danish speakers), and the questionnaire data are provided in the following online repository (link).

Our findings do not contradict the claim by Shibata et al. (2011) and Ramot et al. (2016) that subjects were unaware of the purpose of the neurofeedback procedure in those studies. We argue that it is unlikely that subjects can gain awareness of very complicated experimental purposes like awareness of one’s self-regulation of brain activity related to visual processing like face-related imagery or the orientation of lines. Instead, we argue that certain functions, such as attention, are characterized by affording a kind of meta-cognitive access that is less challenging for subjects to understand and verbalize (van Gaal et al., 2012; Overgaard and Sandberg, 2014). Like a recent review, we argue that theoretical insights from consciousness research would be fruitful for distinguishing between explicit and implicit neurofeedback (Munoz-Moldes and Cleeremans, 2020). Claiming that humans are unaware of a specific type of mental content is a notoriously difficult endeavor. For instance, what was long held as wholly unconscious phenomena, such as dreams or subliminal perception, have been overturned by more sensitive measures which show that degrees of awareness can be detected (Fazekas et al., 2019; Green et al., 1994; Koch et al., 2016; Laberge, 1980; Overgaard et al., 2006; Ramsey and Overgaard, 2004; Sandberg et al., 2010; Siclari et al., 2017; Voss et al., 2013).

Similarly, we argue that a method improvement of assessing AoEP would be advantageous for neurofeedback research. In particular, this should be seen in the light of the appeal of the double blinded neurofeedback procedure as the ultimate way to assess the value of clinical neurofeedback studies (Sitaram et al., 2017). We argue that it is essential to resolve how to measure AoEP in neurofeedback studies. Once it is properly validated, this method should be incorporated into forthcoming double-blinded clinical neurofeedback RCTs such as those in affective disorders. For instance, double-blinded RCTs of pharmacological interventions in major depression have shown that perceived treatment assignment is a better determinant of depressive symptom reduction than the actual treatment (Kirsch, 2019; Laferton et al., 2018). Similarly, to substantiate a potential clinical effect of neurofeedback, it is crucial to address the validity of the RCT principle of the double-blinded design standard.

The issue may also hearken back to the definition of double-blind. It could be seen as too strong a definition for neurofeedback to live up to when the double-blinded experiment is defined as; “An experimental goal where the subject and experimenter are unaware of the nature of the treatment the subject is receiving” (Padhi and Fineberg, 2010), which is the typical definition in neurofeedback studies (Ros et al., 2020; Sorger et al., 2019; Taschereau-Dumouchel et al., 2018b; Thibault et al., 2016). Instead, a soft definition of the double-blind may be more appropriate; “in a double-blind or blinded experiment, information which may influence the subject or the experimenter is withheld until after the experiment is complete” (David and Khandhar, 2020). Or in other words, the recommendation to future neurofeedback studies could simply be; to the extent to which a double-blind is possible, perform a double-blind, and test whether the subject and experimenter became aware of the group assignment and purposes of the experiment (Ros et al., 2020). Some authors have also stressed that the attempt to provide neurofeedback in a double-blinded manner should not replace the practice of adhering to fundamental principles of neurofeedback (Fovet et al., 2017; Pigott et al., 2018). Arguably, suppose it turns out that achieving double-blind in a neurofeedback study in the strong sense is too difficult. In that case, it could be advisable to not compare neurofeedback against blinded placebo-neurofeedback (i.e., sham-neurofeedback), but instead it would be more meaningful to compare neurofeedback against other treatment options (Dohrmann et al., 2007; Geladé et al., 2017; Hartmann et al., 2014; Maurizio et al., 2014; Vanneste et al., 2016). In sum, we provide a novel method for assessing awareness in neurofeedback studies, revealing AoEP in a subset of subjects. We believe that it is essential to develop the method of assessing awareness in neurofeedback to quantify the degree to which subjects become aware of the purposes of neurofeedback experiments. Importantly, this line of research may potentially address and substantiate the role of the double-blind experimental standard in clinical and experimental neurofeedback studies.

CRediT authorship contribution statement

Timo L. Kvamme: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition. Mesud Sarmanlu: Validation, Investigation, Resources, Data curation, Writing – review & editing. Morten Overgaard: Conceptualization, Methodology, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

A grant from Lægeforeningsens Forskningsfond supported this work. We want to thank the participants for their valuable participation in the study. We would also thank the anonymous reviewers for their careful reading and cogent suggestions, which have
improved the present manuscript. The development of the CAQ was inspired by (Grafton et al., 2013).

Appendix A: Supplementary Data

The entire questionnaire, translated transcripts of interviews, a summary of strategies, data on questionnaire responses, R code for statistics and plots, and a supplementary analysis of confidence measures and other C1G measures associated with this article can be found in the following data repository: https://osf.io/h23s6/.

References

Amano, K., Shibata, K., Kawato, M., Sasaki, Y., & Watanabe, T. (2016). Learning to Associate Orientation with Color in Early Visual Areas by Associative Decoded MRI Neurofeedback. *Curr. Biol.*, 26, 1861–1866.

Arns, M., de Ridder, S., Strehl, U., Breteler, M., & Coenen, A. (2009). Efficacy of Neurofeedback Treatment in ADHD: The Effects on Inattention, Impulsivity and Hyperactivity: A Meta-Analysis. *Clin. EEG Neurosci.*, 40, 180–189.

Bagdadyan, J., & Le Van Quyen, M. (2013). Experiencing your brain: Neurofeedback as a new bridge between neuroscience and phenomenology (p. 7). *Neurosci: Front. Behav. Brain Res.*, 27, 107–115.

Bagherzadeh, Y., Baldauf, D., Pantazis, D., & Desimone, R. (2020). Alpha Synchrony and the Neurofeedback Control of Spatial Attention. *Neuron*, 105, 577–587.e5.

Birbaumer, N., Ruiss, N., & Sittaram, R. (2013). Learned regulation of brain metabolism. *Trends Cogn. Sci.*, 17, 295–302.

Cecere, R., Rees, G., & Romei, V. (2015). Individual differences in alpha frequency drive crossmodal illusory perception. *Curr. Biol.*, 25, 231–235.

Dong, M. Y., Sandberg, K., Bibby, B. M., Pedersen, M. N., & Overgaard, M. (2015). The development of a sense of control scale. *Front. Hum. Neurosci.*, 9, 554.

Dohrmann, K., Weisz, N., Schlee, W., Hartmann, T., & Elbert, T. (2007). Neurofeedback for treating tinnitus. *Prog. Brain Res.*, 166, 473–554.

Frederick, J. A., Heim, A. S., Dunn, K. N., Powers, C. D., & Klein, C. J. (2016). Generalization of skills between operant control and discrimination of EEG alpha. *Conscious. Cogn.*, 45, 226–234.

Frederick, J. A., Heim, A. S., Dunn, K. N. (2019). Self-prompted discrimination and operant control of EEG alpha. *NeuroRegulation*, 6, 81–92.

van Gaal, S., de Lange, F. P., & Cohen, M. X. (2012). The role of consciousness in cognitive control and decision making. *Front. Hum. Neurosci.*, 6, 121.

Hartmann, T., Lorenz, I., Müller, N., Langguth, B., & Weisz, N. (2014). The effects of neurofeedback on oscillatory processes related to tinnitus. *NeuroImage*, 84, 267–275.

Kirsch, I. (2019). Placebo effect in the treatment of depression and anxiety. *Front. Psychiatry*, 10, 407.

Kamiya (1969). Operant control of the EEG alpha rhythm and some of its reported effects on consciousness. *Alerted States of Consciousness*, 4, 1–8.

Kamiya, J. (2011). The first communications about operant conditioning of the EEG. *J. Neurother.*, 15, 65–73.

Kirsch, I. (2019). Placebo effect in the treatment of depression and anxiety. *Front. Psychiatry*, 10, 407.

Kober, S. E., Witte, M., Niusan, M., Neuper, C., & Wood, G. (2013). Learning to modulate one’s own brain activity: The effect of spontaneous mental strategies. *Front. Hum. Neurosci.*, 7, 695.

Kohi, M. (2019). MKInfer: Inferential statistics.

Kotchoubey, B., Strehl, U., Uhlmann, C., Holzapfel, S., König, M., Fröschler, W., et al. (2001). Modification of slow cortical potentials in patients with refractory epilepsy: A controlled outcome study. *Epilepsia*, 42, 406–416.

Kühler, A., Neumann, N., Kaiser, J., Kotchoubey, B., Hinterberger, T., & Birbaumer, N. (2001). Brain-computer communication: Self-regulation of slow cortical potentials for verbal communication. *Arch. Phys. Med. Rehabil.*, 82, 1533–1539.

Kvamme, T. L., Sarmanlu, M., Bailey, C., & Overgaard, M. (2022). Neurofeedback Modulation of the Sound-induced Flash Illusion Using Parietal Cortex Alpha Oscillations Reveals Dependency on Prior Multisensory Congruency. *Neuroscience*, 482, 1–17.

Laberge, S.P. (1980). Lucid dreaming: An exploratory study of consciousness during sleep. *search.proquest.com*.

Lafontaine, A.C.; Cipriani, S.; Baer, L.; Alcan, A. J.; Cooper, A.; Papakostas, G.; et al. (2018). Mechanisms of Perceived Treatment Assignment and Subsequent Expectancy Effects in a Double Blind Placebo Controlled RCT of Major Depression. *Front. Psychiatry*, 9, 424.

Lange, J., Oostenveld, R., & Fries, P. (2013). Reduced occipital alpha power indexes enhanced excitability rather than improved visual perception. *Journal of Neuroscience*, 33, 3212–3220.

Lange, J., Kell, J., Schnitzler, A., van Dijk, H., & Weisz, N. (2014). The role of alpha oscillations for illusory perception. *Behav. Brain Res.*, 271, 297–301.

Nunes, A. T., van Dongen-Boomsma, M., Buitelaar, J. K., & Slats-Willems, D. (2011). ADHD and EEG-neurofeedback: A double-blind randomized placebo-controlled feasibility study. *J. Neural Transm.*, 118, 275–284.

Maurizio, S., Liechti, M. D., Heinrich, H., Jäncke, L., Steinhausen, H. C., Walitza, S., et al. (2014). Comparing tomographic EEG neurofeedback and EMG biofeedback in children with attention-deficit/hyperactivity disorder. *Biol. Psychol.*, 95, 31–44.
Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.*, 15, 70–73.

Witte, M., Kober, S. E., & Ninaus, M. (2013). Control beliefs can predict the ability to up-regulate sensorimotor rhythm during neurofeedback training. *Neurosci: Front. Hum.*

Yun, K., Bhattacharya, J., Sandkuhler, S., Lin, Y.-J., Iwaki, S., & Shimojo, S. (2020). Causally Linking Neural Dominance to Perceptual Dominance in a Multisensory Conflict. *Neuroreport*, 2020.04.06.027227.