



Effects of neurofeedback and working memory-combined training on executive functions in healthy young adults

Shirley Gordon^{1,4} · Doron Todder^{2,3} · Inbal Deutsch⁴ · Dror Garbi^{1,4} · Oren Alkobi⁵ · Oren Shriki⁵ · Anat Shkedy-Rabani⁵ · Nitzan Shahar¹ · Nachshon Meiran¹

Received: 15 July 2018 / Accepted: 19 March 2019 / Published online: 3 May 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Given the interest in improving executive functions, the present study examines a promising combination of two training techniques: neurofeedback training (NFT) and working memory training (WMT). NFT targeted increasing the amplitude of individual's upper Alpha frequency band at the parietal midline scalp location (Pz), and WMT consisted of an established computerized protocol with working memory updating and set-shifting components. Healthy participants ($n = 140$) were randomly allocated to five combinations of training, including visual search training used as an active control training for the WMT; all five groups were compared to a sixth silent control group receiving no training. All groups were evaluated before and after training for resting-state electroencephalogram (EEG) and behavioral executive function measures. The participants in the silent control group were unaware of this procedure, and received one of the training protocols only after study has ended. Results demonstrated significant improvement in the practice tasks in all training groups including non-specific influence of NFT on resting-state EEG spectral topography. There was only a near transfer effect (improvement in working memory task) for WMT, which remained significant in the delayed post-test (after 1 month), in comparison to silent control group but not in comparison to active control training group. The NFT + WMT combined group showed improved mental rotation ability both in the post-training and in the follow-up evaluations. This improvement, however, did not differ significantly from that in the silent control group. We conclude that the current training protocols, including their combination, have very limited influence on the executive functions that were assessed in this study.

Introduction

Executive functions are defined as the abilities responsible for organizing, monitoring and regulating lower level cognitive functions such as perceptual, motor processes to fit information processing demands of the current task (Friedman and Miyake, 2017). People with good executive functions are characterized as being flexible since they are better able than those with poor executive functions in dealing with complex situations requiring creative, “out-of-the-box” solutions. In contrast, poor executive functions characterizes immaturity (e.g., Diamond Thomas & Munro, 2007) as well as many forms of psychological dysfunction including psychopathology (e.g., Kashdan and Rottenberg, 2010), rigid problem-solving style and poor reasoning ability (Friedman and Miyake, 2017). Due to the benefits of having good executive functions abilities, studies trying to enhance executive functions by training have become a popular topic in neuroscience. Such training includes computerized cognitive training (see meta-analysis by Au et al., 2015), meditation

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00426-019-01170-w>) contains supplementary material, which is available to authorized users.

✉ Shirley Gordon
Shirley.gordon@gmail.com

¹ Department of Psychology and Zlotowski Center for Neuroscience, Ben-Gurion University of the Negev, 84105 Beersheba, Israel

² Mental Health Center, Beer Sheva, Ministry of Health, Beersheba, Zlotowski, Israel

³ Center for Neuroscience, Ben-Gurion University of the Negev, Beersheba, Israel

⁴ IDF Medical Corps, Tel Hashomer, Ramat Gan, Israel

⁵ Department of Cognitive and Brain Sciences, Ben-Gurion University of the Negev, Beersheba, Israel

(e.g., Goyal et al., 2014; Fox et al., 2016), neurofeedback (e.g., Gruzelier, 2014a) and transcranial direct current stimulation (e.g., Au, Karsten, Buschkuehl & Jaeggi, 2017). In the current study, we were inspired by Miyake et al.'s (2000) taxonomy of executive functions. According to Miyake et al. (2000), there are three main components of executive function: updating working memory, switching and inhibition. Therefore, we chose specific training protocols that have previously been found to improve these abilities. The present study evaluates the use of neurofeedback training (NFT), in the individually defined upper Alpha band, and computerized working memory training (WMT)—including their combination—for the improvement of executive functions. Specifically, the current investigation raises questions as to whether a combination of NFT + WMT would prove more effective than a training protocol employing a single approach (WMT, NFT only).

Previous literature demonstrated that upper Alpha (10–12 Hz) synchronization has a key part in executive function processes. Klimesch, Doppelmayr, Schwaiger, Auinger & Winkler (1999) for example, demonstrated an increase in upper Alpha power in a memory task that involved switching between different memory sets. Support for the role of upper Alpha on inhibition comes from Hummel, Andres, Altenmüller, Dichgans, and Gerloff (2002) who requested participants to withhold a response while performing a motor task. EEG results showed de-synchronization in upper Alpha during movement and synchronization in upper Alpha during the inhibition phase. These effects were detected in central scalp locations, which is in agreement with other literature showing that strong upper Alpha synchronization is detected mostly in central and parietal locations on the scalp (Klimesch, 1999; Cooper, Croft, Dominey, Burgess, & Gruzelier, 2003; Basar, 2006). From these findings, one can anticipate that an enhancement of upper Alpha in the parietal and central scalp sites by means of NFT can improve executive functions in a healthy young population (for review, see Gruzelier, 2014a; Mirifar, Beckmann & Ehrlenspiel, 2017; Enriquez-Geppert, Huster & Herrmann, 2017). Indeed, increasing upper Alpha amplitude by NFT was found to be effective in improving performance on a mental rotation test (Hanslmayr, Sauseng, Doppelmayr, Schabus & Klimesch, 2005; Zoefel, Huster & Herrmann, 2011), known to be related to visual working memory (Just and Carpenter, 1985; Hyun & Luck 2007). Moreover, Nan, Rodrigues, Ma & Qu (2012) demonstrated improvement in a working memory capacity task (forward and backward digit-span tasks), using 20 short 3-min sessions of upper Alpha NFT in a central scalp location. The authors did not find significant differences between pre-training EEG baseline to post-training EEG in the NFT group and suggested that the overall training (60 min) was relatively short and, therefore, no physiological changes were detected relative to a silent

control group. Escolano, Aguilar & Minguez (2011) also found working memory enhancement as a result of NFT. Furthermore, increase in upper Alpha was evident immediately after training in the NFT group. In sum, short-term changes (behavioral and/or electrophysiological) were demonstrated in a healthy young population. However, there are still no indications of long-term effects of such training on executive functions as well as on resting-state EEG spectral measurements.

As aforementioned, one of the three main components of executive functions according to Miyake et al. (2000) is updating working memory. Another method aimed to improve working memory abilities is computerized training (e.g. N-back training, Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; complex span training; Harrison et al., 2013; etc.). Albeit being a popular topic, there is a heated debate regarding computerized training effectiveness. First, mixed results have been found in multiple meta-analyses regarding the generalization effects of training (Karbach and Verhaeghen, 2014; Au et al. 2015; Melby-Lervåg and Hulme 2016). Second, it is unclear whether training leads to significant changes in brain physiology (for meta-analyses see Buschkuehl, Jaeggi, Jonides, 2012; Patel, Spreng & Turner, 2013). In the current study, we used a computerized WMT task developed by Shahar and Meiran (2015), which involves working memory updating (n-back) and set-shifting (task-switching) elements. Previous results using this protocol showed both short- and long-term improvements (10–11 months after training) in a near transfer¹ choice reaction time task as compared to a waiting list group. Specifically, improvement was localized to the aspects of the reaction-time distribution known to be involved in working memory, the right tail of the distribution (the Tau parameter in the Ex-Gaussian distribution; Schmiedek et al. 2007).

Combinations of training methods

Several studies have investigated the effects of a combination of different types of training protocols on the improvement of executive functions. Results are encouraging, with combined training protocols showing more promising outcomes than single-approach training protocols (e.g., Park, Seo, Kim & Ko, 2014; Ditye, Jacobson, Walsh & Lavidor, 2012; Hosseini, Pritchard-Berman, Sosa, Ceja, & Kesler, 2016). However, most of these studies were performed on clinical or elderly populations (Alfonso, Caracuel, Delgado-Pastor & Verdejo-García, 2011; Johnstone, Roodenrys, Johnson, Bonfield, & Bennett, 2017; Park et al., 2014). As far as we know, only one study tested the effectiveness of a combined

¹ Improvement in the training task leads to improvement on tasks that are similar to the one participants were trained on.

intervention in a healthy young adult sample (Hosseini et al., 2016). In that study, near-infrared spectroscopy in the prefrontal cortex was employed for NFT, and was combined with computerized WMT. Specifically, participants simultaneously received behavioral feedback regarding their performance on the working memory task as well as their brain activity (increased oxygenated hemoglobin in the prefrontal cortex). Twenty healthy adult participants (from which only 10 received sham NFT) underwent 4 sessions of NFT (100 min total) and their performance on executive functions tests were measured before and after completion of training. Results indicated that participants who underwent NFT showed significant improvements in executive functions performance (in N-back task and a switching task) as compared to participants who received sham feedback. Moreover, when compared to the sham group, the NFT group significantly reduced activity in the right middle and inferior frontal regions. The authors suggested that training leads to greater efficiency of storage, access, and updating of working memory representations and that this improvement was mediated by the right middle frontal gyrus. In contrast to Hosseini et al. (2016), the current study examined the effects of a combination of EEG-NFT of the parietal individual upper Alpha frequency and WMT (NFT + WMT), administered one after the other and not simultaneously. We chose a serial training design in which NFT came before WMT, a feature which enabled us to test the immediate effectiveness of the NFT on working memory performance. Another reason for employing this combination was based on Miyake et al.'s (2000) taxonomy. Of interest are claims in the literature that task switching involves both working memory retrieval and inhibition. Specifically, task switching requires the retrieval of the new goal (Altmann & Gray, 2008) and the new task rules (Mayr & Kliegl, 2000, 2003), and, therefore, involves working memory, especially working memory retrieval. In addition, task switching also involves inhibition of the rule that was relevant in the previous trial (Mayr & Keele, 2000; see; Koch, Gade, Schuch and Philipp, 2009, for review), as well as inhibition of previously relevant rules that could generate competing responses (Meiran, Hsieh & Dimov, 2010; Meiran, Hsieh & Chang, 2010, 2011). According to the findings presented above (Hummel, 2002; Klimesch, 1999; Cooper et al., 2003; Basar, 2006), there is a reason to believe that upper Alpha NFT may influence inhibition. As such, we hypothesized that when both inhibition and working memory retrieval would improve as a result of training, task-switching ability would also improve. Specifically, the current investigation had several goals. (A) To investigate whether upper Alpha NFT has a direct influence on the WMT task that comes afterwards. We were able to examine this question by calculating the mean level of difficulty reached at the end of the 10th (i.e. final) session of WMT for each participant in

the NFT + WMT and in the WMT-only groups and by comparing these values between the groups; we hypothesized that the combined group (NFT + WMT) would reach higher scores in the WMT as compared to WMT-only group (B) to investigate whether the combination of training methods would create near and far transfer² effects for the short and long term. This was done by comparing NFT + WMT group to NFT + active control training group and actually, also by comparing NFT + WMT group to WMT-only group and to silent control group; we hypothesized that the combined training group (NFT + WMT) would show greater benefits in working memory, inhibition and switching tasks as compared to the NFT + active control training, WMT-only or silent control groups (C) to improve our examination of the WMT. This goal was achieved by comparing between WMT and active control training groups. Thus, the present study constitutes a methodological improvement relative to previous investigations of this specific WMT task (Shahar & Meiran, 2015) that used a waiting list group for control. We hypothesized that the WMT group would show greater benefits in working memory task relatively to the group that received active control training. (D) In a similar way, a comparison between NFT group and active control training group constitutes a methodological improvement relative to previous NFT investigations (Nan et al., 2012). We hypothesized that the NFT group would have greater benefits in inhibition and working memory tasks relative to the group who received active control training only.

Methods

Participants and selection procedure

One hundred and sixty-five young adults participated in the study (95 males, mean age = 22.12 years, $SD_{age} = 2.54$, range 19–28). The participants were Israeli Defense Force soldiers who were invited, through an email ad, to take part in a large brain training study. All participants reported having normal or corrected-to-normal vision, with no history of psychiatric disorders, head trauma, central nervous system disorders or use of psychotropic medications. All participants were proficient Hebrew speakers. 86.5% were right handed³. The potential to improve brain functions was the main motivator for participation. All participants gave informed consent prior to their participation and were

² Improvement in the training task can be generalized to improvement on tasks that are very different from the one participants were trained on, but still require similar cognitive processes.

³ Handedness was determined based on the preferred hand for writing, assessed on a three-point scale (1 = right, 2 = left, 3 = either hand/ambidextrous).

Table 1 Descriptive statistics of the different training groups

Training protocol	Dropout/total (%)	Duration of training (min)	Final number of participants	Age in years mean (SD)	Gender (male, female)	Intelligence rating score mean (SD)	Individual alpha frequency mean (SD)
NFT + WMT	2/22 (9.10)	40	20	21.60 (1.93)	11,9	70.00 (11.69)	9.65 (1.23)
NFT + ACT	9/29 (31.03)	40	20	22.60 (2.93)	13,7	68.00 (13.22)	9.40 (1.14)
NFT	1/21 (4.76)	20	20	21.65 (2.45)	10,10	67.00 (14.18)	9.65 (0.75)
WMT	0	20	20	23.15 (2.88)	11,9	68.50 (12.26)	9.30 (0.98)
ACT	3/23 (13.04)	20	20	21.47 (2.35)	13,7	66.50 (12.68)	9.85 (1.14)
Silent control	10/50 (20.00)	–	40	21.98 (2.51)	27,13	69.75 (10.74)	9.48 (1.02)
Dropout (Total)	25/165 (15.15)	–	–	22.48 (2.52)	10,15	67.20 (13.69)	9.63 (0.79)

Dropout represents the number of participants who voluntarily dropped out from training relative to the total number of participants that started the training

NFT neurofeedback training, *WMT* working memory training, *ACT* active control (visual search) training

informed that they could withdraw from the study at any time. The institutional review board of the Israeli Defense Force Medical Corps approved the study. The study was conducted according to the guidelines of the declaration of Helsinki (World Medical Association, 2013). Since it was a training study, another inclusion criterion was having at least an average Intelligence Rating Score⁴ of 50. The mean in our sample was accordingly above average and showed restricted range ($M = 68$, $SD = 12.22$, as compared with $SD = 20$ in the non-restricted population). Participants were randomly assigned to one of six groups (Table 1). There was a 15% dropout rate from the study, therefore, we continued to recruit participants until equal group sizes were reached (20 participants per group and 40 in the silent control). Silent control was doubled in size, to ensure statistical power in focused comparisons. A final sample of 140 participants (85 males) remained for data analysis. The critical statistical analyses concerned planned contrasts which compared the training groups (each $N = 20$) to either silent control ($N = 40$) or to the active control ($N = 20$). We thus focused on these comparisons, though in practice, quite a few of the analyses increased power by pooling groups. The conclusions from the power analyses thus refer to the minimal sensitivity, whereas in practice, sensitivity was often higher. Sensitivity

analyses conducted with G-Power, assuming Power = 0.80, $\alpha = 0.05$, and a pre-post $\rho = 0.70$ (see more below) indicated that the study could detect partial Eta-squared of 0.03 for $N = 40$ (comparison to active control) and Eta-squared of 0.02 for $N = 60$ (comparison to silent control). When assuming $\rho = 0.50$ between pre- and post-test, and $N = 40$ (the worst case scenario), the analyses indicated sensitivity to detect partial Eta-squared = 0.05. All these analyses examined the between (group) by within (pre-post) interaction as examined in a 1-df (denominator) F test.

Design and procedure

An email invitation for participation in a training study for cognitive enhancement was sent to the work emails of different Israeli Defense Force Departments. The invitation included details of preliminary requirements (age range, medical constrictions, etc.). Participants who agreed to participate in the study and were found fit for study requirements were randomly assigned to one of the six groups and signed a consent form. All six groups underwent pre, post and follow-up evaluations at the same fixed intervals of the study timeline. Figure 1 depicts a schematic illustration of the study's design and procedure. The pre-training evaluation comprised of a passive resting-state EEG recording followed by administration of the brief executive function battery, mental rotation test, and questionnaires (120 min). The questionnaire results are reported elsewhere (Soffer-Dudek, Todder, Shelef, Deutsch & Gordon, 2018). The formal training phase started 1 day after the pre-training evaluation and continued over five consecutive weeks (ten training session ~30–60 min each, depending on group assignment, twice a week, with at least 1 day between sessions). Between 3 and 5 days after completing the tenth training session, participants completed the post-training evaluation, which was similar to the pre-training evaluation, comprised of a passive resting-state EEG recording followed by administration of

⁴ The Intelligence Rating Score is a well-established tool for intelligence assessment comprised from four sub-tests presented in a multiple-choice format: (1) the Otis-R, which measures the ability to understand and carry out verbal instructions; (2) Similarities-R, assesses verbal abstraction and categorization; (3) Arithmetic-R, which measures mathematical reasoning, concentration, and concept manipulation; and (4) Raven's Progressive Matrices-R, that measures non-verbal abstract reasoning and visual-spatial problem-solving abilities. The sum of the scores for the four tests forms a validated measure of general intelligence, scored on a 9-point scale, scaled between 10 and 90, with a 10-point increment at each score (Gal 1986). The correlation between the general intelligence rating score and the WAIS total IQ is > 0.90 (Kaplan et al. 2002).

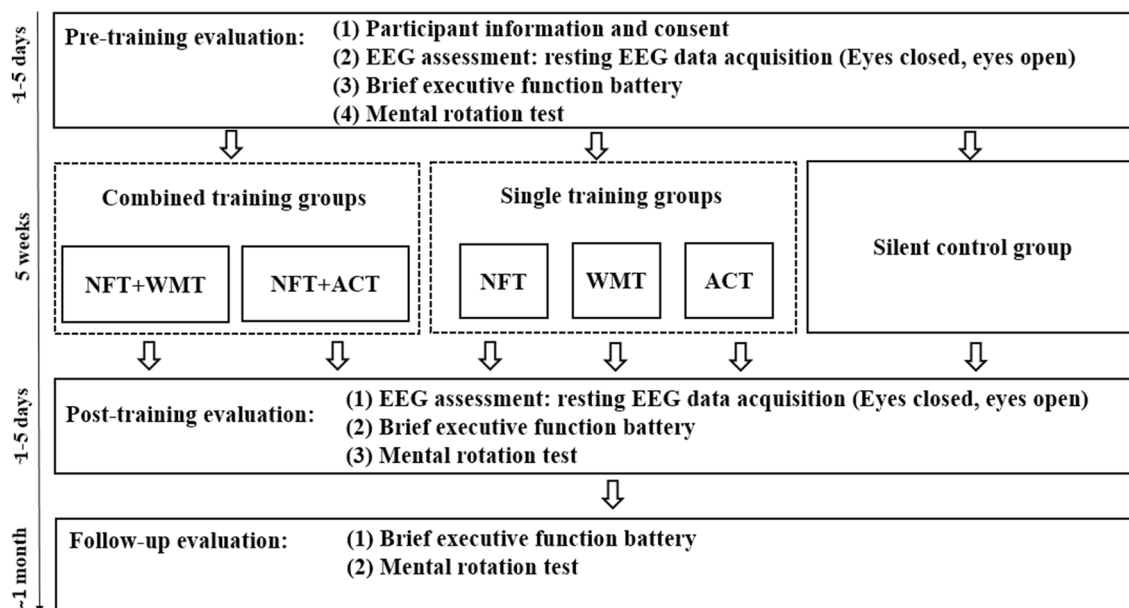


Fig. 1 A schematic illustration of the study's design and procedure. All six groups underwent pre, post and follow-up evaluations at the same fixed intervals. Resting-state EEG was administered twice; pre- and post-training. Five training groups completed ten sessions of

training, ~20–50 min each, depending on group assignment, twice a week, over a period of 5 weeks (ten sessions total). *NFT* neurofeedback training, *WMT* working memory training, *ACT* active control training

the brief executive function battery and mental rotation test. A follow-up evaluation with a brief executive function battery plus mental rotation test similar to that administered in the pre-post-training evaluation (without EEG recording) was conducted 1 month after the first post-training evaluation, to identify long-term effects of the different training protocols. There were two “combined” training groups, both started each session with 21 min of NFT followed by 20 min of WMT (NFT + WMT) or followed by 20 min of active control training (NFT + active control training). To accurately evaluate the effects of a training paradigm, it is recommended to compare the trained group to an active control training group in which participants believe they are participating in a training protocol. By doing so, training effects such as the effects of repetition, intervention (such as regular computer use), and subjective expectations from training can be accounted for (Oken et al., 2008). To control for placebo effects, the training conditions for the trained group and active control training group must be identical, in that participants in both groups undergo the same amount of training time. We also added three single-protocol groups who underwent only one type of training protocol (NFT/ WMT/ active control training). It is important to note that the amount of NFT was strictly equivalent, and the groups thus differed in terms of what other activity they did after the NFT session (WMT/active control training). This structure enabled us to compare between the NFT + WMT and NFT + active control training groups and also between

the three different single training groups. All groups were also compared to the silent control group. Participants in the silent control group were invited to a round of training sessions that were dated after the actual study had ended, without the participants being aware of this difference relative to the other groups. They received a different version of consent form than that of the other groups, with the timeline study starting with pre, post and follow-up measurements in the same fixed intervals of the study timeline and then ten sessions of training. The rationale provided in their consent form was that the study's intention was to first investigate changes in executive functions over time and then start training sessions. We made these adjustments to ensure that the participants in the silent control group would be similar to the training groups in their perception of the goals of the study. We employed a procedure, in which both the participants and the experimenters were ignorant to the training group assignment. After finishing the entire study procedure, participants were offered a full debriefing and received a personal training report.

Apparatus and setting. Participants were seated in front of a computer screen in a quiet, air-conditioned room. Each participant was trained in a separate room. Behavioral training tasks were programmed using E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Stimuli were presented on a black 19-in. (48.26 cm) computer screen. Participants responded using a QWERTY keyboard. NFT was done with the Brainmaster2EBII module in combination with the

Brainmaster3SE software (<http://www.brainmaster.com>). The electrodes were connected to the scalp using gel (<http://www.weaverandcompany.com>) and the scalp location was prepped with Nuprep gel.

Measurements of pre- and post-evaluations

Resting-state EEG recording

Before and after the training intervention phase, participants performed the resting-state EEG recording (Escolano et al., 2011). During EEG recording, participants were placed in a quiet, air-conditioned room, sat on a comfortable chair, in relaxed waking. Resting-state EEG was recorded for a period of 3 min with eyes closed and 3 min with eyes open. In the eyes closed condition, participants were requested to stay relaxed as much as possible, while trying not to fall asleep, and not think of anything specific. In the eyes open condition, participants were requested to look straight at the white wall and not think of anything specific. EEG recordings were conducted using Discovery 24E (<http://www.brainmaster.com>). All electrode computerized electroencephalograph with 19 mono-polar leads were in accordance with the international 10–20 electrode distribution system (FP1, FP2, F3, Fz, F4, F7, F8, T3, C3, Cz, C4, T4, T5, T6, P3, Pz, P4, O1, O2). Impedance for each channel was measured and adjusted until they were kept below 5 k Ω before recording started. EEG activity was digitized at a sampling frequency rate of 256 Hz and band pass filtered online between 0.5 and 100 Hz. An EEG cap (<http://www.electro-cap.com>), one of four different sizes (small, small-medium, medium-large and large), was attached with a gel to the scalp and used to record EEG signals. In addition to these 19 electrodes, two earlobe electrodes (termed A1 and A2) were attached to the two earlobes. Electrodes were referenced to the left earlobe with the ground electrode at the CPZ location.

EEG pre-processing and analyses. EEG processing and analyses were performed offline using the Neuro-guide v-2.5.2 software program (Applied Neuroscience Inc., St. Petersburg, FL, USA). We analyzed our data relative to linked ears references. All EEG recordings were carefully and individually checked for artifacts (eye blinks and muscle artifacts) by visual inspection and then edited to be removed from the data. A minimum of 90 s⁵ of EEG recording was obtained for each participant, with all subsequent calculations being based on the average EEG spectrum computed from these minimum 90 s. Test–retest and split-half

reliability for the entire EEG recording, as well as for all 19 electrode locations separately, were generated automatically by the Neuro-guide software for every EEG recording. Split-half reliability coefficients were calculated as the ratio of variance of all the even 1-s segments of EEG recording divided by all the odd 1-s segments of EEG recording. Variance was calculated as the sum of the square of the deviation of each time point from the mean of all selected time points. Test–retest reliability coefficients were calculated by dividing the EEG recording in half, treating each half as a separate occasion. Reliability was calculated by comparing the variance of the beginning half of the selected EEG recording to the variance of the end half of the selected EEG recording with variance calculated as above. Good split-half and test–retest reliability is considered > 0.90 with an edited sample length of more than 60 s. All test–retest and split-half reliability coefficients used in the analysis were above 0.90. A Fast Fourier Transform was computed on 2-s epochs thus yielding a 0.5 Hz frequency resolution over the frequency range from 0 to 30 Hz for each epoch. The 75% sliding window method of Kaiser and Serman (2001) was used to compute the Fast Fourier Transform in which successive 2-s epochs (i.e. 256 points) were overlapped by 500 ms steps (64 points) to minimize the effects of the Fast Fourier Transform windowing procedure. Relative power was computed from the 19 scalp locations in the following frequency bands: Theta (4–7 Hz); Alpha (7.0–13 Hz); lower Alpha (7–10 Hz); and upper Alpha (10–13 Hz). Relative power was computed as the ratio of power in a given band/sum of power from 1 to 30 Hz (i.e. total power) \times 100. The advantage of using relative power values and not absolute power values is that it eliminates the potential contribution of individual differences in skull thickness and volume conduction. Additionally, to create an individual upper Alpha NFT protocol, we extracted for each participant the Individual Alpha Frequency (IAF). We used peak (i.e., maximum amplitude) frequency in determining the IAF. The Alpha range maximum peak frequency was measured over a 4–14 Hz band, using 0.1 Hz jumps. We computed eyes closed spectrum and eyes open spectrum and then calculated the difference between them and extracted the individual value for the frequency of the maximum Alpha activity peak. The frequency of the maximal power in the difference spectrum was taken as an anchor representing the individual peak of the participant (for more details see Klimesch, 1999). We mainly focused on the analysis of Alpha (at posterior sites) and Theta (frontal) bands given their established role in training studies that examined executive functions (Langer, von Bastian, Wirz, Oberauer, Jäncke, 2013; Klimesch et al., 1999; Escolano et al., 2011).

⁵ In general, the test re-test reliability of quantitative EEG is an exponential function of sample length in which 20 s epochs are approximately 0.8 reliable, 40 s approximately 0.9 reliable and 60 s asymptotes at approximately 0.95 reliability (Burgess and Gruzelier 1993; Van Albada and Robinson 2007).

Brief executive function battery

After completing the EEG recording, participants performed the brief executive function battery and the mental rotation test. The order of administration was fixed for all participants to minimize any error due to participant by order interaction (Miyake et al., 2000). The brief executive function battery⁶ includes tests measuring switching, inhibition and working memory functioning. The rationale for choosing these executive function tasks was inspired by Miyake et al.'s (2000) influential taxonomy of executive functions. We decided to rely on Miyake et al.'s (2000) taxonomy of executive functions partly because it is empirically based, and also because it appears to be highly influential in executive function literature discussion.

The brief executive function battery consisted of three tasks: (A) choice reaction time task. This test began with three six-choice reaction tasks (high working memory load) with tasks involving letters, digits and shapes, respectively (Fig. 2, bottom). Each task comprised of 72 trials preceded by 6 practice⁷ trials. In all three tasks, the mapping between stimuli and response keys (on the keyboard) was arbitrary (Shahar, Teodorescu, Usher, Pereg & Meiran, 2014) and thus required keeping this mapping information in working memory. Participants used the index, middle, and ring fingers of their two hands to respond in this task. In the next phase, 3 additional 2-choice reaction tasks (low working memory load) were executed, with tasks involving letters, digits and shapes, respectively (Fig. 2, top) and comprising 36 trials each, preceded by 2 practice trials. Participants used only the index fingers of both hands, with the two choices mapped to the same stimuli as in the previous phase. Thus, working memory load was reduced in the two-choice condition by both having fewer rules to keep in mind (two vs. six) and by the fact that these rules were trained beforehand. The stimuli for the choice reaction time were the Hebrew letters ת, ש, ר, ק, צ, פ, ס, ג, מ, ל, ב, ט, נ, י, ד, א, כ (Hebrew was the language of the participants), and the digits 0–9, which were presented using 48-point Times New Roman font. The shapes were eight symmetrical shapes printed in white against a black background. Each shape was 64 × 64 pixels in size. Each trial included a fixation (500 ms) and target (until response or until 6 s). Errors were followed by a 400 ms visual feedback.

⁶ It is important to mention here that this brief executive functions battery provided a relatively reasonable coverage of the executive function domain. However, this coverage was nonetheless incomplete. For example, instead of measuring working memory updating, we indirectly measured other working memory functions. One of them is conceptually related to capacity (“Alternative Cost”). The other (“Tau”) is described in the main text.

⁷ In the practice phase, all responses were followed by a feedback (correct/error), see Fig. 2.

The instructions were to respond as fast and as accurately as possible. Two indices of working memory efficiency were extracted from the choice reaction time task: “*Alternative-cost*” which is the difference between the mean reaction time of high working memory load (six choices) and the mean reaction time of low working memory load (two choices). High scores represent compromised working memory efficiency. “*6-choice Tau*” represents the rate of exceptionally slow reaction times of high memory load, as quantified with the Tau parameter from the ex-Gaussian model of reaction-time distributions. High Tau has been shown to have a very high correlation with individual differences in working memory⁸ as estimated at the latent variable level (Schmiedek et al., 2007). Additionally, Shahar et al. (2014), using mathematical modeling of the decision process, showed that Tau is linked to working memory retrieval rate, with high Tau indexing slow retrieval of information from working memory.

(B) The switch task. This task was introduced after the choice reaction time task and used the already learned two-choice stimulus–response mapping with task-cues that were already familiar. The task began with a screen displaying the task-cues and the stimulus–response mapping and continued with a sequence of trials in which the task switched in every trial (Fig. 3). Presentation of letters, shapes, and digits as well as error indication were the same as in the choice reaction time task, with the only difference being that each target stimulus comprised of a combination of a shape, a letter and a digit. There were 201 trials in this task, preceded by 6 practice trials. The task cue stimuli were “W” for letters, “%” for numbers and “O” for shapes. Each task-cue was 64 × 64 pixels in size. Stimuli were presented in white on a black background. The task cues were presented at the center of the screen 500 ms prior to the presentation of the stimulus. Each stimulus appeared until response was given or after 6 s had elapsed. Two distraction-related effects were controlled for. Spatial compatibility was controlled for by having one-third of the target stimuli being presented on the same side of the responding hand (spatially compatible), one-third on the opposite side (incompatible) and one-third in the middle. Task-rule congruency effects were controlled by the fact that in one-third of the trials, all stimuli (target

⁸ The validation of Tau as an index of working memory functioning comes from three sources. One is the correlational studies (e.g., Schmiedek et al. 2007), showing a high correlation between individual differences in Tau and in working memory. The second is the fact that Tau (and not Sigma or Mu) is strongly influenced by the working memory demands of the task (mapping arbitrariness, especially when the number of alternatives is high, Shahar et al. 2014, see also, 2018). Lastly, mathematical modeling reported in these papers has linked Tau to the rate of retrieval from working memory, with this model being superior to alternative models such as those linking Tau mainly to attentional lapses or to the rate of evidence accumulation.

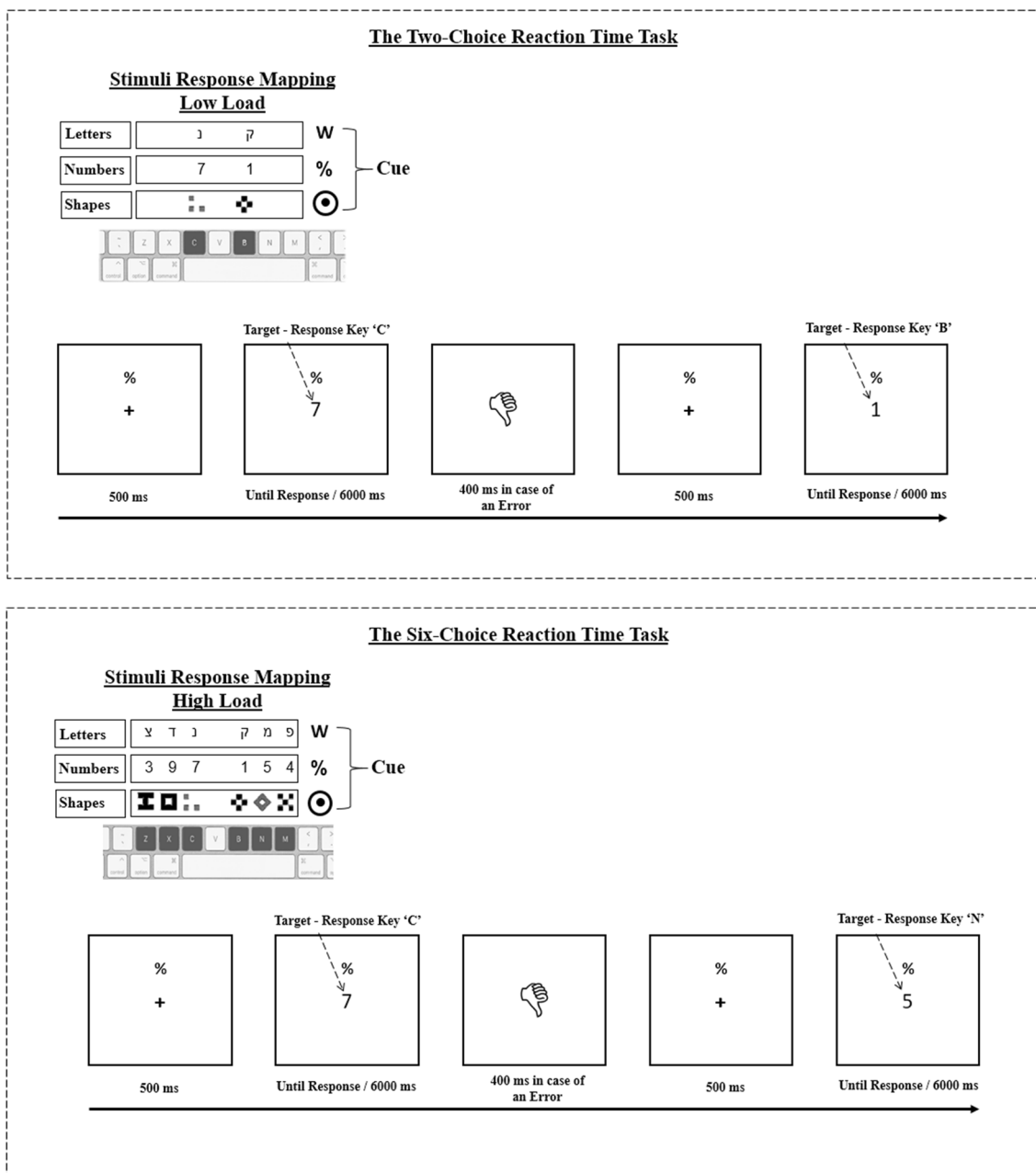


Fig. 2 The choice reaction time task. An example of the number version of the choice reaction time task. The two-choice tasks (top) represent low working memory load. The six-choice tasks (bottom) rep-

resent high working memory load. A task-cue was attached to every task (letters, digits or shapes)

and two distractors) were mapped to the same response, in one-third, one was mapped to the opposite response and in one-third two were mapped to the opposite response. The related effects are not reported in this study. The instructions were to respond as fast and as accurately as possible. Two indices of switching were extracted from the Switch Paradigm: “*Alternation-cost RT*” represents the difference in reaction time between repeated trials taken from experimental blocks without task-switching (“single-task blocks”) and

switched trials (see Meiran et al., 2000). High scores represent difficulty in switching. “*Alternation-cost in errors*” represents the difference in errors between repeated and switched trials. High scores represent poor switching ability.

(C) The anti-saccade task (Miyake et al., 2000). This task measured inhibition success. 96 trials started with a centered fixation mark (+) that appeared for various durations (1000 ms, 1500 ms and 2000 ms), followed by a visual cue presented on one side of the screen (e.g., left) for

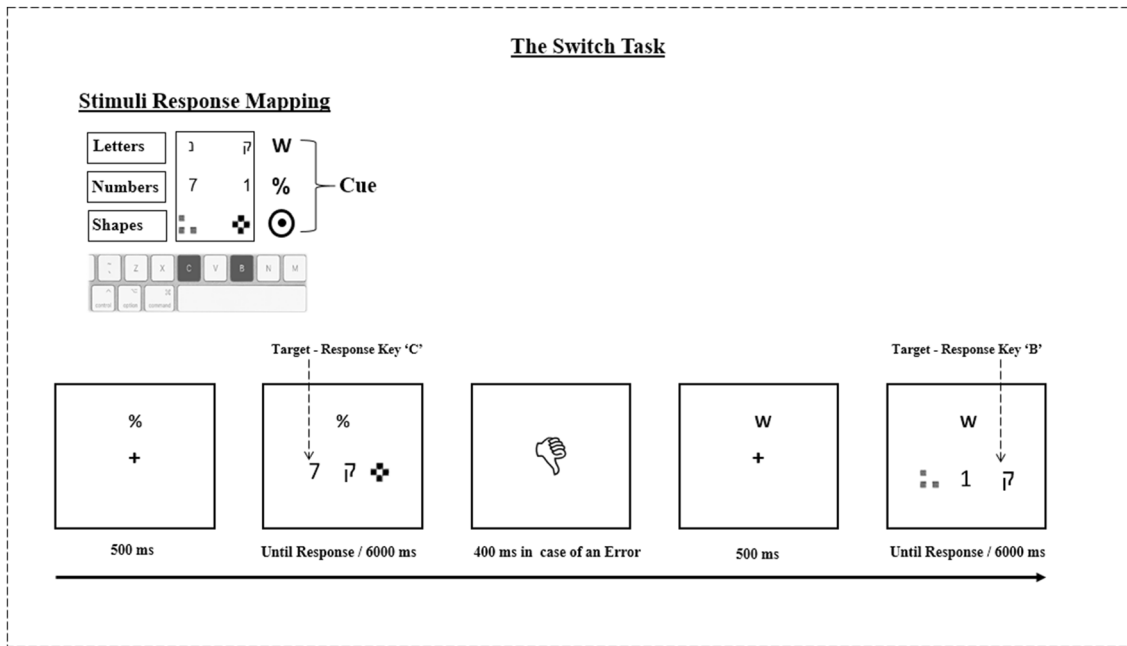


Fig. 3 The switch task. An example of the task with two steps of “word” that switched to “number” task. In the learning phase, cues and stimuli–response mapping presented at the beginning of the

task. There was no time limit for this screen. The cue sign appeared in every trial, along with the target and two other distractors. Errors were followed by a visual feedback (thumbs down)

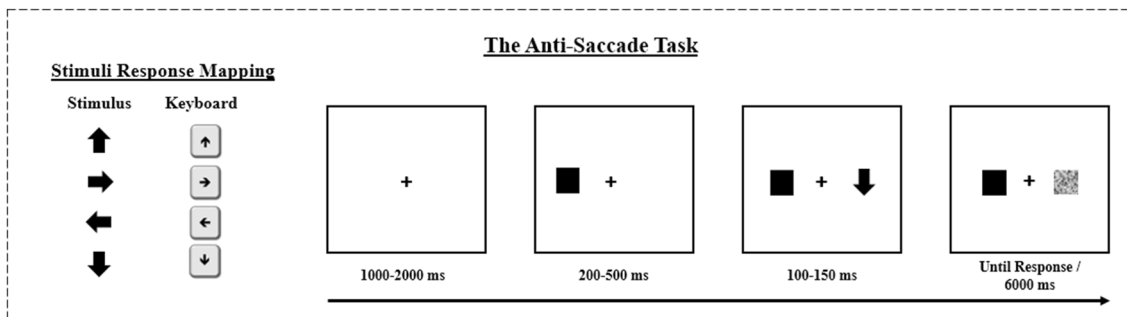


Fig. 4 The anti-saccade task. Each trial began with a centered fixation mark (+) for variable durations (1000 ms, 1500 ms, and 2000 ms), followed by a cue/distractor presented on the right or the left side of the centered fixation mark for variable durations (200 ms,

250 ms, 300 ms, and 350 ms). Then, a target arrow appeared on the opposite side of the centered fixation mark for 100 ms and then masked until response or after 5 s

200–350 ms in 50 ms intervals, followed by the presentation of a target stimulus on the opposite side (e.g., right) for 100 ms before being masked by gray cross-hatching that disappeared after response or after 5 s. The visual cue was a white square (64 × 64 pixels), and the target stimulus was a small white arrow (64 × 64 pixels). The participants’ task was to indicate the direction of the arrow (left, up, down or right) with the keyboard’s arrows (Fig. 4). Participants had to inhibit the reflexive response of looking at the initial cue (a small white square) because doing so would make it difficult to correctly identify the direction of the arrow. The task started with 24 practice trials. We used the proportion

of the incorrect responses as a measure of inhibitory control called “*Anti-saccade in errors*”. High proportion of errors represents bad inhibition.

The mental rotation test

Another measurement included as part of the pre-post-evaluations was the mental rotation test. Previous studies have shown improvements in mental rotation abilities after elevating posterior upper Alpha by NFT (Hanslmayr et al., 2005; Zoefel et al., 2011). Since the current NFT protocol was not an exact replication of these earlier studies, our purpose

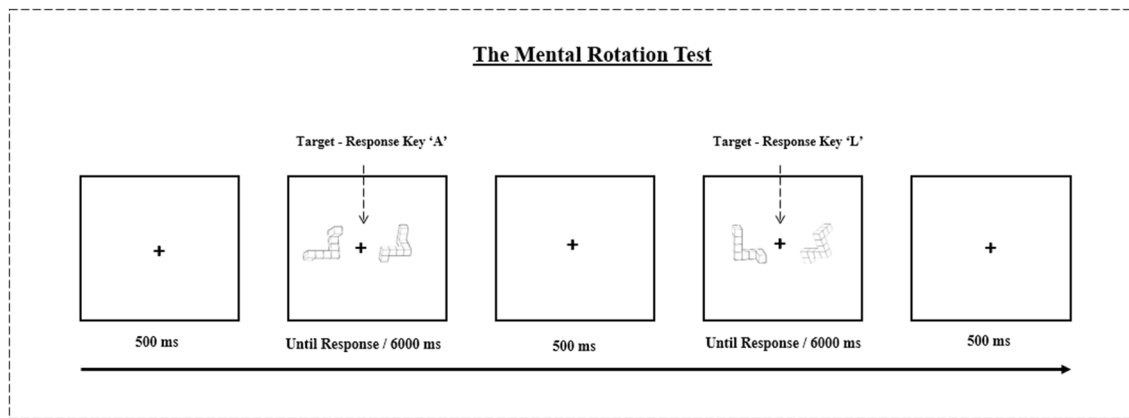


Fig. 5 The mental rotation test. The instructions presented on screen requested the participants to decide whether a pair of three-dimensional shapes were identical (even if rotated) or different. In this figure, ‘A’ represented identical shapes, ‘L’ represented different

was to first make sure that it has an influence on mental rotation abilities as well as on working memory abilities, thus replicating prior findings (Escolano et al., 2011; Nan et al., 2012, respectively). Moreover, since mental rotation ability is known to be related to visual working memory (Hyun & Luck, 2007), we also used it as an indicator of far transfer effects of WMT, in addition to the near transfer working memory task (choice reaction time task). The mental rotation test was a modification of the original task by Shepard and Metzler (1971). The instructions presented on the screen requested participants to decide whether a pair of three-dimensional shapes, each composed of ten cubes, were identical (even if rotated) or different. The rotation angle of the three-dimensional shape was on one axis only each time (horizontal or vertical). Participants responded by pressing left (A) or right (L) on the keyboard. The assignment of keys (A, L) to YES and NO responses was counterbalanced between participants (Fig. 5). The participants received a visual example for correct rotation and incorrect rotation before entering the task. There was no time limit for the instructions screen. The task consisted of a single block of 64 trials, without practice trials⁹. The stimuli were presented side by side, horizontally, with a 20-pixel separation between them on the computer monitor. Each shape was 290 × 290 pixels in size. The stimuli were presented until a response was given or until 10 s had elapsed. Accuracy and response times were measured as a function of the rotation angle (55°, 100°, 145°, and 190°). Each pair was presented such that one shape was on the right and the other was on the

shapes. The rotation of the three-dimensional shape was only on one axis each time (horizontal or vertical). Accuracy and response times were recorded as a function of the rotation angle (55°, 100°, 145°, and 190°)

left with a 7.6° visual angle between the shapes centers. The interval between the response and the next pair of shapes was 500 ms. “*Mental Rotation RT*” represents the reaction time of the correct responses. High scores represent poor visual working memory (longer time to mentally rotate the shape). “*Mental Rotation in errors*” represents the proportion of the incorrect responses. High proportion of errors represents poor visual working memory.

Training protocols

Working memory training (WMT)

We used a WMT task developed by Shahar and Meiran (2015), which involves working memory updating (n-back) and set-shifting (task-switching) elements (Fig. 6). This training task had four components aimed to tap working memory load. (1) Task-switching component: participants were asked to switch between two 2-alternative choice reaction tasks (an object classification task and a spatial classification task). In each trial, a task cue preceded the target notifying participants which one of the two tasks should be performed. (2) *N*-Back component: participants were asked to respond according to the stimuli presented *N*-trials beforehand. Whether the *N*-back aspect referred to the target stimulus or the cue stimulus was randomly selected at the beginning of each training block. The *N* value in the current block was set according to the participants’ performance in the previous block. (3) Stimulus and response novelty: to increase task novelty, each training block included a randomly selected new set of task cues, target stimuli (objects/locations) and response keys. (4) Task difficulty was adapted according to accuracy performance. If the participant made 12 errors or less, difficulty level increased in the following

⁹ We considered the first 4 trials of the mental rotation test as practice trials and, therefore, analysis was conducted without them, on 60 trials per participant.

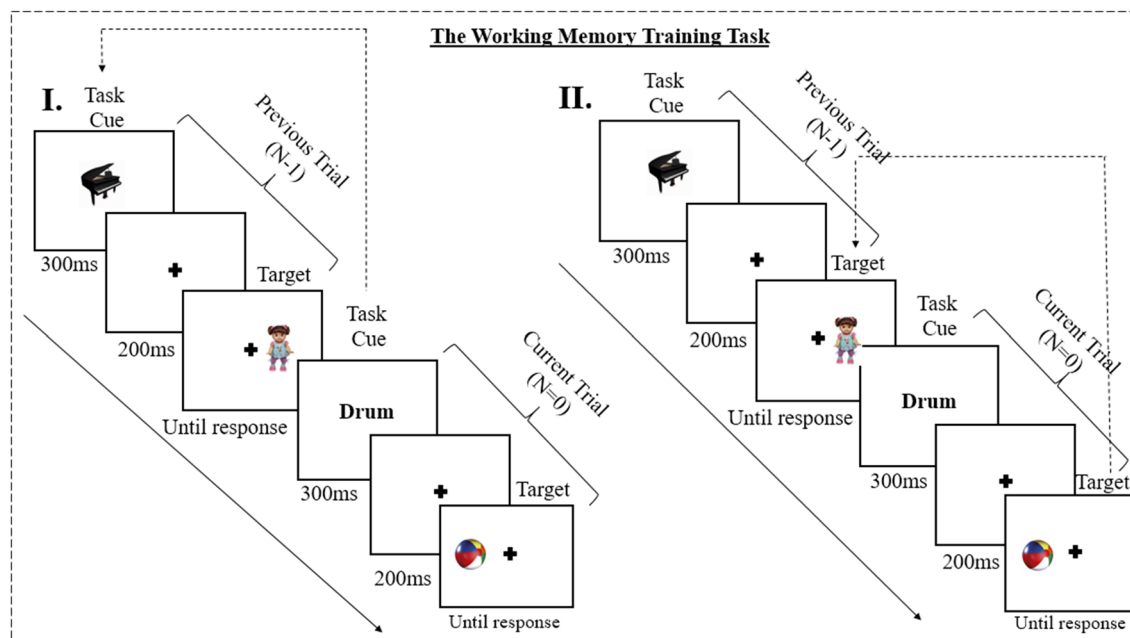


Fig. 6 The working memory training task. I Task sequence and the correct response when $N=1$ according to “task-cue”. II Task sequence and the correct response when $N=1$ according to “target”

block. If 24 errors or more were made, the difficulty level decreased. Otherwise, difficulty level remained unchanged. In determining difficulty, the following parameters were adjusted: N , the proportion of incongruent trials, task-cue compatibility and stimulus–response compatibility. Maximum level of difficulty was 10. In each condition, participants completed a 6 trial practice phase followed by test phase of 100 trials. Each trial consisted of a fixation (250 ms), target (until response or until 6 s had elapsed) and a blank screen inter-stimulus interval (250 ms). A 400 ms beep signaled an error. For details see Shahar et al. (2018).

Neurofeedback training (NFT)

The frequency of upper Alpha bands was extracted individually for each participant prior to the beginning of the NFT, to create a tailored training protocol using the dominant EEG frequency in the Alpha band (IAF) as an anchor point (for more details see Klimesch, 1999; Hanslmayr et al., 2005; Nan et al., 2012). We decided to use a NFT protocol, targeting an increase in the amplitude of the IAF at the parietal midline location on the scalp (Pz). Our decision was based on former studies showing that Alpha power is typically highest during high mental workload at parietal sites (Gevins et al., 1979a, b, c; Gevins and Smith, 2000). The duration of NFT were similar to Guez et al. (2015) who used also upper Alpha protocol at parietal scalp location. All electrode impedances were kept under 5 k Ω . The channel for training was connected to the Pz location and referenced to the left

earlobe. The ground channel was connected to the right earlobe. EEG was recorded at 256 samples per second with a 50 Hz Notch filter. At first arrival, the neurofeedback device was set to the IAF + 2 Hz of the participant. The threshold of upper Alpha amplitude for reinforcement was set in the first 3 (active) min of each training session and remained fixed throughout the entire session. This procedure is similar to Escolano et al.’s (2011) and done to examine the within training effect of a NFT session. Figure 7 depicts an illustration of the NFT timeline sequence of the session. At the beginning, participants were asked to look at a thermometer presented on the computer screen, which reacted to the upper Alpha amplitude of the participant. Participants were instructed to try and raise the thermometer line as high as possible, while staying relaxed and calm. After three minutes, the threshold for reinforcement was set at the 70th percentile of the mean upper Alpha amplitude measured during this time. Therefore, for the remaining duration of the training session, every upper Alpha amplitude that fell in the range of the top 30% of the amplitudes in the first 3 min, received a reward. Hence, this configuration created a challenging training task. After the first 3 min of active threshold assessment, the participant played three games, 5 min per game, resulting in 15 min of training (Fig. 7). These are simple/standard games (packmen and ball filling), that can be found in the Brainmaster3SE software (<http://www.brainmaster.com>). The training session ended in another 3 min of active post-training measurement with the thermometer. Therefore, the total amount of NFT session was 21 min. In

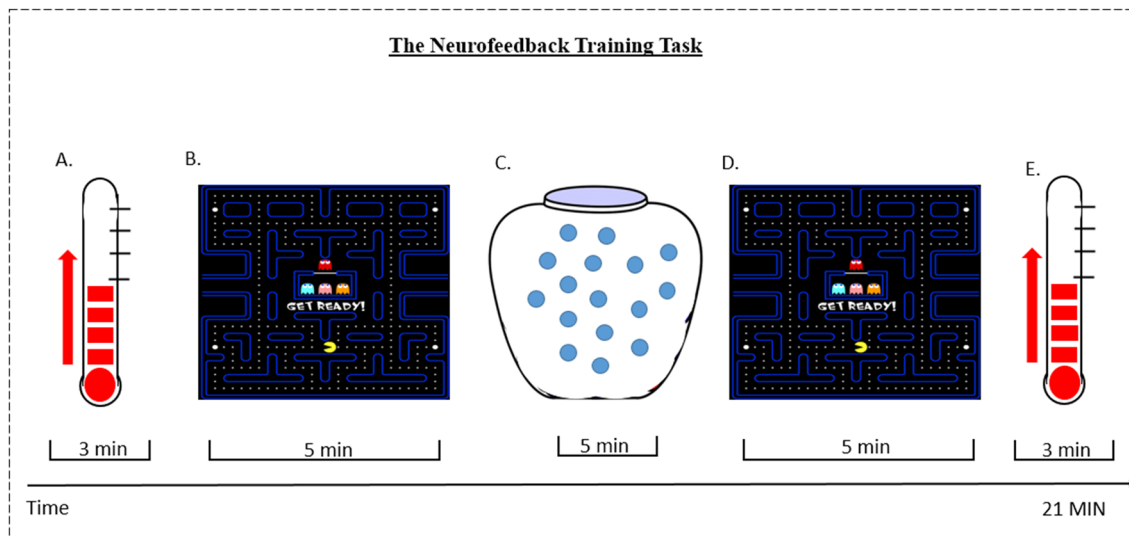


Fig. 7 The neurofeedback training task. Illustration of the timeline sequence of training

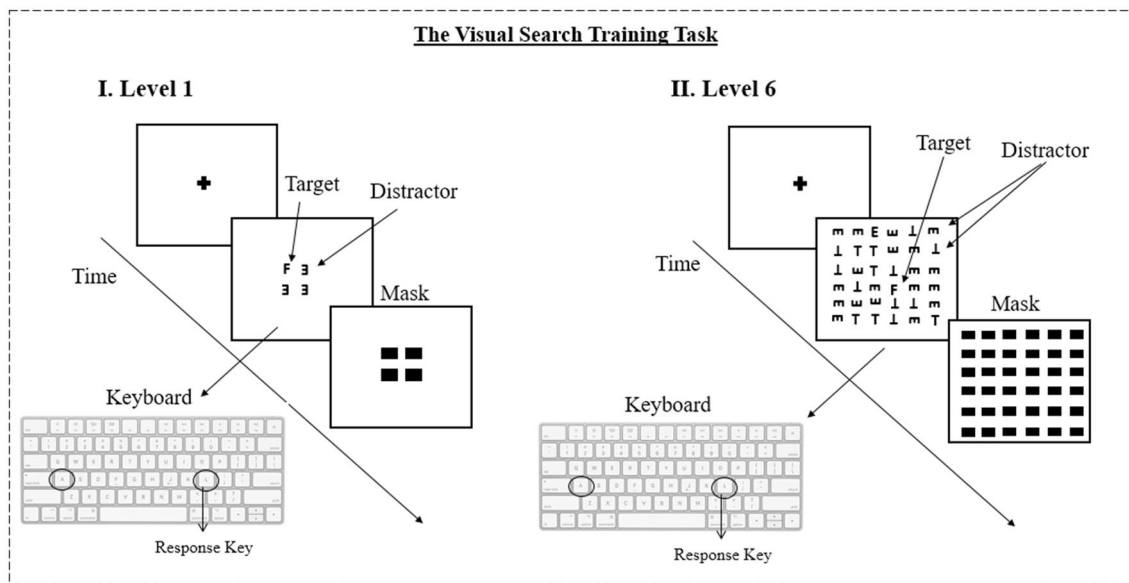


Fig. 8 The visual search training task. An active control training protocol (I). Task difficulty is low (II). Task difficulty is higher. Maximum level difficulty was 10

the first two training sessions, participants were told that gaining points and progressing in the game are good signs and that no specific effort on their part is required since the learning process appears to be mostly subconscious. In later sessions, no instructions were needed.

Active control training

We adapted the visual search training task (Fig. 8) as an active control training task from Redick et al. (2013). On

each trial, participants were asked to report whether a target letter (‘F’) was facing right or left, using a right or left key press, respectively. Difficulty level (array size and distractors composition) was set according to participants’ performance in the previous block. The criteria for level increase were if accuracy was higher than 87.5% and the criteria for level decrease were if accuracy was lower than 75%. Otherwise, difficulty level remained unchanged. Each training session comprised of 16 blocks, with 24 trials each.

Data handling

EEG learning

As recommended in Gruzelier's (2014b) review, we used all three types of measurement for EEG learning: within-session, across-sessions and resting-state EEG measures. We added also learning specificity of NFT as described in Escolano et al. (2011) and Zoefel et al. (2011) studies.

Within-session learning. To determine whether NFT resulted in learning reflected by increasing upper Alpha amplitude, we extracted the mean amplitude of the individual upper Alpha band at the Pz location for each participant from the beginning (first 3 min) and end (last 3 min) of the neurofeedback session and looked at the change in Alpha amplitude within each session over the entire training course. This learning index reflects 'trainability' of the participant within each session (for more details see review by Gruzeiler, 2014b and also Escolano et al., 2011).

Across-session learning. This measurement was defined as the change pattern along the entire training course. It was measured by calculating the slope (on a logarithmic scale) of each participant across the training sessions, with higher slope indicating higher learning rate across the entire training time. We chose to calculate the logarithmic slope and not the linear slope because it is more accurately represents the learning function occurring during NFT (Wan, Nan, Vai, & Rosa, 2014). For calculating the logarithmic slope, we first plotted the data of each participant with the axis scaling determined by x scale log, y scale linear. A linear function was then fitted to the scaled data. Then, the slope of the line in the log-scale units was taken as an index.

Learning specificity of NFT. Learning specificity was assessed in the frequency bands close to the upper Alpha band [IAF, IAF + 2] as described in Escolano et al. (2011) and Zoefel et al. (2011) studies, (1) Lower Alpha [IAF-3, IAF-1], and (2) Lower Beta [IAF + 3, IAF + 5]. The EEG frequency spectrum was extracted from the first 3 min of the NFT session, taken from the first and the last NFT session.

NFT influence on resting-state measurements. As in Hanslmayr et al. (2005), the resting-state measurements were calculated as the post-to-pre-change in the resting-state EEG spectral power measures. Specifically, to investigate whether changes in resting-state Alpha occurred as a result of NFT, we computed the relative power of upper Alpha band (10–12 Hz) and the lower Alpha band (7–10 Hz) extracted from the training electrode (Pz). A separate index was computed for each condition, eyes closed/eyes open. We presented only the results of eyes closed condition in all resting-state EEG analysis for brevity, yet found similar results in the eyes open condition. These results presented in the supplementary materials on line (NFT influence on resting-state measurements, eyes open).

Behavioral learning

Across-session learning. To address the question of whether WMT resulted in improvement in the training task, we calculated the mean level of difficulty reached at the end of each training session as a dependent variable. Same procedure was done for the active control training (visual search) protocol.

Impact of NFT on working memory learning. To investigate whether receiving NFT before WMT has an influence on working memory learning, we calculated the mean level of difficulty reached at the end of the 10th (i.e. final) session of WMT for each participant in the NFT + WMT group and in the WMT-only group and compared between the groups.

WMT influence on resting-state EEG

Similar to previous WMT studies (e.g. Langer et al., 2013), we computed the relative power of frontal Theta extracted from five frontal leads (mean average of FP1, FP2, F3, F4, and Fz) and compared between WMT group and active control training group, in eyes closed condition. Furthermore, we investigated the within changes (post-pre-evaluations) in frontal Theta in the WMT group. For that, we calculated the relative power of frontal Theta amplitude at Fp1 and Fp2 locations, in the eyes closed condition (as previously done in Langer et al.'s, 2013).

Transfer effects

Brief executive function battery and mental rotation test. Only correct responses were taken into account in the calculations of response time measures. Trials with response time shorter than 100 ms or higher than 4 SD from the mean score in every response time index and also trials after error were discarded. As a result¹⁰, 8.9% of the trials were discarded from the choice reaction time, 12.8% of the trials were discarded from the switching task, and 22% of the trials were discarded from the mental rotation test. Most of the discarded trials were post-error trials, but these had to be excluded since it is known that after an error there is a post-error slowing (Laming, 1979; Rabbitt, 1966). We computed seven indices, extracted from the brief executive function battery and mental rotation test as mentioned in the [Methods](#) section. To test for reliability of the indices, we computed both retest correlations and Cronbach's alpha for each index. Table S1 in the supplementary materials online demonstrates that all indices were both sensible (in the sense of reproducing well-established findings) and reliable. To test for group equivalence before training, an analysis of

¹⁰ The proportion of discarded trials is the sum of all three evaluations (pre, post, follow-up).

variance (ANOVA) was conducted for each executive functions index separately with group as a between-subjects independent variable. Results indicated lack of group differences in all the executive functions measures. For details, see Table S2 in the supplementary materials online.

Statistical analysis

Ten participants (from the overall 60 participants who performed NFT) had several (1 up to 3) neurofeedback sessions with extreme values (above 100 mv), therefore, we replaced these values by interpolated values, based on the adjacent sessions. Five participants were entirely removed from the NFT analysis, since the quality of their data was low (all of their NFT sessions were extremely high, above 100 mv). Two were from the NFT + WMT group, two from the NFT + active control training group and one from the NFT group. Therefore, statistical analysis of NFT was conducted on 55 participants. Three participants (from an overall of 40 participants who performed WMT) had 1 session missing and we replaced the missing value by an interpolated value.

Bayes factors analysis

For all statistical analysis computed in this study, aside from standard significance tests, we computed Bayes Factors (Rouder, Morey, Speckman, Province, 2012) using JASP 0.8.1.2 (JASP Team, 2018). In “standard” null hypothesis testing, one can either reject H_0 or remain undecided. The advantage of Bayesian statistics and especially Bayes Factors statistic is that they allow for three outcomes: accepting H_0 , accepting H_1 , or remain undecided. Specifically, when multiplying the prior odds (the a priori belief regarding the relative odds of H_1 and H_0) by Bayes Factors, one obtains the posterior relative odds of H_0 and H_1 , given the data. For simplicity, and since we did not have grounds to assume otherwise, the prior odds ratio was assumed to be one, meaning that the Bayes Factors represents the posterior odds of H_0 and H_1 , given the data. We report BF_{10} (the relative odds in favor of H_1), but it is important to keep in mind that BF_{01} (the relative odds in favor of H_0) equals $1/BF_{10}$. The accepted criteria are for $BF_{10} > 3$ to indicate some support for H_1 (and concomitantly, $BF_{10} < 0.33$ to accept H_0), and $BF_{10} > 10$ to indicate strong support for H_1 (and $BF_{10} < 0.10$ for accepting H_0). Additionally, we report usual null hypothesis testing, namely p values, and effect size partial Eta-squared.

Results

A final sample of 140 participants (85 males) was used for data analysis. Their mean age was 22.06 years with a range of 19–28 years. Table 1 presents the descriptive statistics

of the different training groups. There were no significant differences between groups in age [$F(5, 134) = 1.38$, $p = 0.241$, $BF_{01} = 5.92$], gender (Cramer's $V = 0.14$, $p = .762$, $BF_{01} = 118.9$), Intelligence Rating Score [$F(5, 134) = 0.31$, $p = 0.902$, $BF_{01} = 30.91$] or Individual Alpha Frequency [$F(5, 96) = 0.34$, $p = 0.891$, $BF_{01} = 20.81$]. In fact, the Bayes Factors indicate significant support for lack of group differences. However, significant differences were found in dropout rates between groups [$p = 0.031$, by Fisher–Freeman–Halton test, correction of Chi-squared]. Importantly, the highest dropout rates were in the “control groups”; NFT + active control training (31%) and silent control (20%). Given that participants are more likely to drop out of training if they have poor executive functions (Arbiv and Meiran, 2015), participants in these control groups were probably higher (after the drop out), on average, in executive functions as compared to the training groups. Therefore, claims of improvement as being related to drop out rate are less probable.

In this section, we first present evidence for EEG learning found in all NFT groups including the within and across-session learning effects and evidence for NFT learning specificity (i.e., in the Alpha band and not in the adjacent frequency bands). In addition, we show evidence for non-specific influence of NFT on resting-state EEG spectral topography. Then, we present the significant working memory learning effects found in the WMT groups (across-sessions) as well as the learning effects found in the active control training groups. We also show lack of impact of NFT on working memory learning, as well as the lack of influence of WMT on resting-state EEG spectral measures. Finally, we show the lack of impact of NFT on near and far transfer tasks and evidence for limited and short-lasting transfer effect found only in the WMT group and the lack of specific effect of the combined training group on far transfer task (mental rotation test).

EEG learning

Within-session learning

The main analysis was a three-way mixed model ANOVA with NFT Group (NFT + WMT, NFT + active control training, NFT; 20 participants in each group) as a between-subjects independent variable, and Sessions (1–10) and Time (beginning/end of session) as a within-subjects independent variables. A within-session improvement was found across groups [Time: $F(1, 52) = 25.39$, $p < 0.001$, $\eta_p^2 = 0.34$, $BF_{10} = 9.57 \times 10^{32}$] indicating strong support for H_1 and no significant difference in this improvement was found between the groups [Time by Group: $F(2, 52) = 0.50$, $p = 0.601$, $\eta_p^2 = 0.02$, $BF_{10} = 0.59$]. However, we could not

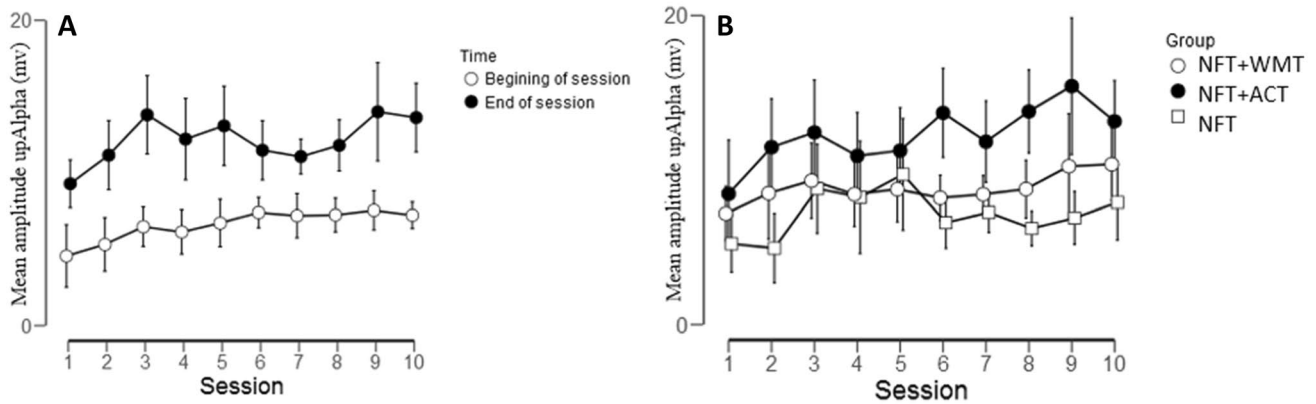


Fig. 9 NFT effects collapsed over Time (A) and over Group (B). Error bars represent 95% credible intervals. *ACT* active control training group

accept H_0 either in this case (since $BF_{10} < 0.33$ is required for accepting H_0).

Across-session learning

We assessed the level of change in the learning index at the group level through a one-sample t test. The logarithmic slope variable was not distributed normally and was, therefore, analyzed with the nonparametric Wilcoxon signed-rank test. According to our predictions, we revealed a significant across-session improvement [$V(52) = 742$, $p < 0.001$, $BF_{10} = 14.49$] indicating strong support for H_1 . To control for possible contribution of group differences to the effect, one-way between-subjects ANOVA analyses were conducted that included the Group variable (NFT + WMT, NFT + active control training, NFT). The analyses revealed support for lack of differences between the NFT groups in the across-session learning index [$F(2, 52) = 0.14$, $p = 0.879$, $\eta^2 = 0.007$, $BF_{10} = 0.19$]. Figure 9 depicts the results of NFT improvement (within-session) collapsed over Time (A) and the results of NFT improvement (across-sessions) collapsed over Group (B).

Learning specificity of NFT

Statistical significance has been assessed using paired, one-tailed t tests. We used one-tailed t tests since we expected to demonstrate an increase in the Alpha band. Significant differences were found both in the upper and lower Alpha bands, respectively, [$t(48) = 2.39$, $p = 0.012$, $BF_{10} = 4.01$ indicating some support for H_1 ; $t(48) = 3.43$, $p = 0.001$, $BF_{10} = 22.39$ indicating strong support for H_1]. No significant difference was found in the lower Beta band [$t(48) = 1.94$, $p = 0.061$, $BF_{10} = 1.07$, although we cannot accept H_0]. Figure 10 depicts the frequency spectrum of the IAF ± 6 Hz range. These results are in line with those reported by Escolano et al. (2011) and Zoefel et al. (2011)

demonstrating learning specificity (i.e., in the Alpha frequency and not in the adjacent frequency bands) as a result of individual upper Alpha NFT at posterior scalp location.

NFT influence on resting-state measurements

As opposed to the analyses above, which focused on EEG measured as a part of NFT, in the following analyses, we focused on the resting-state EEG. The main analysis was a three-way mixed model ANOVA with NFT (Yes/No), WMT (Yes/No) and Time (pre/post-training). Upper Alpha power (Pz) was the dependent variable. The critical effect is the interaction between NFT and Time, and this interaction was found to be non-significant (null hypothesis testing) and showing support for H_0 (Bayesian statistics, $BF_{10} < 0.33$) [$F(1, 72) = 0.41$, $p = 0.532$, $\eta_p^2 = 0.005$, $BF_{10} = 0.31$]. Simi-

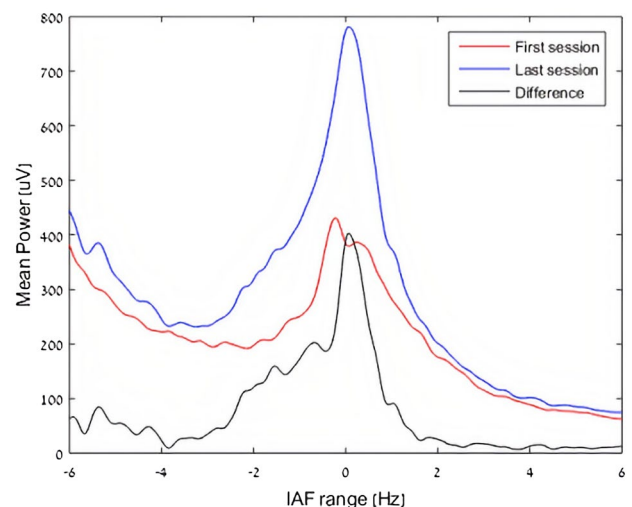


Fig. 10 Frequency spectrum in the range [IAF-6–IAF+6] for the active first 3 min of the first and last NFT sessions

lar results were found when comparison was done with the lower Alpha band as the dependent variable [$F(1, 72) = 0.37, p = 0.552, \eta_p^2 < 0.005, BF_{10} = 0.29$]. Additionally, we compared between NFT-only group and silent control group in a repeated measures ANOVA with Time (pre/post-training) and Group as the independent variables and upper Alpha power as the dependent variable. No significant interaction between Group and Time was found [$F(1, 57) = 0.07, p = 0.801, \eta_p^2 = 0.001, BF_{10} = 0.28$, showing support for H_0]. Comparison conducted with the lower Alpha band as the dependent variable, revealed similar null results, supporting H_0 [$F(1, 57) = 0.01, p = 0.918, \eta_p^2 < 0.001, BF_{10} = 0.28$]. Furthermore, we compared between NFT and active control training groups in the same way. No significant interaction between Group and Time was found in the upper Alpha power [$F(1, 37) = 0.30, p = 0.583, \eta_p^2 = 0.01, BF_{10} = 0.36$, supporting H_0]. Similar non-significant null results were found when analysis conducted for the lower Alpha band [$F(1, 37) = 0.24, p = 0.632, \eta_p^2 = 0.01, BF_{10} = 0.42$], however, we could not fully endorse H_0 . In sum, we found no NFT-related group differences in posterior (Pz) resting-state Alpha power, which was true for both upper and lower Alpha.

To investigate whether resting-state Alpha at posterior locations had been elevated as a result of NFT, we performed paired sample t tests in the NFT groups. Here, we decided to use two-tailed t tests since we had no prior expectations. The dependent variable was the relative power of upper Alpha (Pz). Null hypothesis testing showed no significant difference between pre-post-evaluations [$t(19) = 1.39, p = 0.901, BF_{10} = 0.53$], and the Bayesian analysis also indicated an undecided result. Interestingly, significant difference was found in the right frontal Theta (Fp2) [$t(19) = 3.39, p < 0.001, BF_{10} = 42.67$, indicating strong support for H_1]. Similar significant results were found in frontal Theta, in the NFT + active control training group [Fp1, $t(19) = 2.70, p = 0.011, BF_{10} = 3.80$; Fp2, $t(19) = 2.70, p = 0.012, BF_{10} = 3.52$, both indicating some support for H_1]. Yet, there was no significant indication (null hypothesis testing) for a specific change in upper Alpha band (Pz) in this group [$t(19) = 2.06, p = 0.061, BF_{10} = 1.32$], although Bayes Factors remained undecided. In the NFT + WMT group, however, significant difference was found between pre-post-evaluations in the right parietal lower Alpha [P4, $t(19) = 2.86, p = 0.013, BF_{10} = 5.12$, with some support for H_1]. No significant changes in parietal Alpha and/or frontal Theta were found in the WMT, active control training and silent control groups. Figure 11 depicts the topo plots of paired t tests spectral analysis results (pre-post-change) of the different training groups. These results are in line with the significant within and across-session EEG learning effects found in all NFT groups. Nevertheless, they demonstrated that the

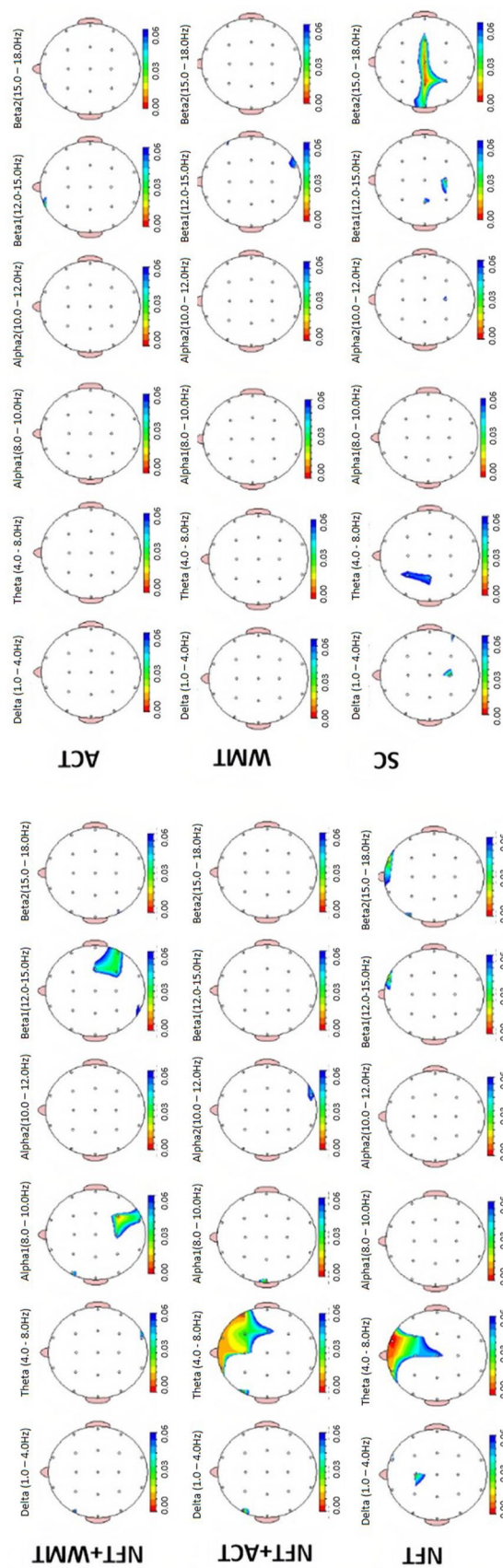


Fig. 11 Topo plots of paired t tests results (pre-post-change) of the different training groups created by Neuro-guide program. ACT active control training group, SC silent control group

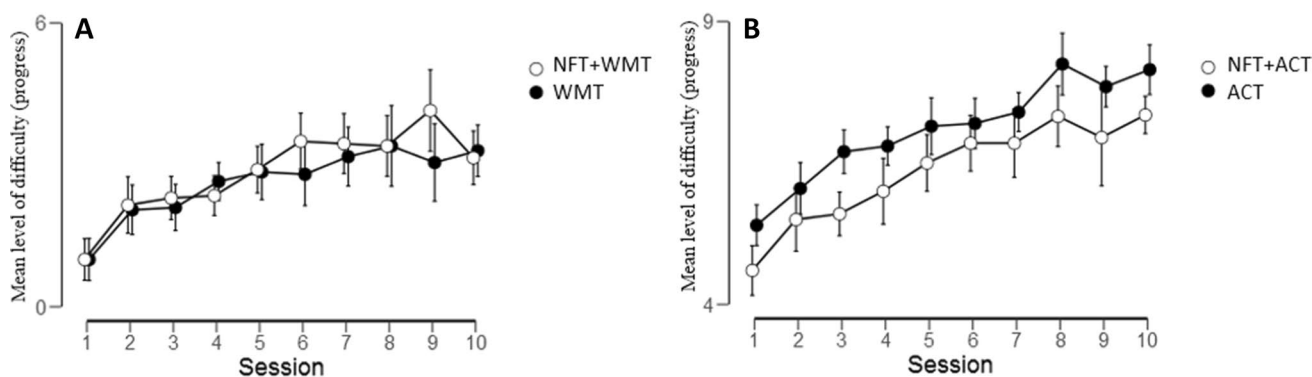


Fig. 12 Learning effects of the computerized training protocols; **a** WMT and **b** active control training. Error bars represent 95% credible intervals. *ACT* active control training group

influence of NFT on resting-state EEG spectral topography is not necessarily specific (e.g., Egner et al. 2004).

Behavioral learning

Across-session learning

To test for the across-session WMT learning, we analyzed the final WMT level achieved in a session with a two-way mixed model ANOVA with WMT Groups (NFT + WMT, WMT) as a between-subjects independent variable and Sessions (1–10) as a within-subjects independent variable. Significant effect was found for Session [$F(9, 342) = 15.61$, $p < 0.001$, $\eta_p^2 = 0.29$, $BF_{10} = 3.47 \times 10^{18}$, indicating strong support for H1], and there was strong evidence for lack of Group differences¹¹ in this learning effect [$F(9, 342) = 1.06$, $p = 0.391$, $\eta_p^2 = 0.03$, $BF_{10} = 0.06$, with strong support for H0]. The results of active control training were analyzed with a two-way mixed model ANOVA with active control training Groups (NFT + active control training, active control training) as a between-subjects independent variable and Sessions (1–10) as a within-subjects independent variable. Significant effect was found for Session, indicating a learning effect [$F(9, 342) = 29.59$, $p < 0.001$, $\eta_p^2 = 0.44$, $BF_{10} = 2.59 \times 10^{35}$, and a strong support for H1], and there was decisive support for lack of Group differences¹² in this effect, as indexed by the interaction [$F(9, 342) = 0.44$, $p = 0.853$, $\eta_p^2 = .01$, $BF_{10} = 0.02$, with a strong support for H0]. Figure 12 depicts the results of learning effects in the

WMT groups (A) and the results of learning effects in the active control training groups (B).

Impact of NFT on working memory learning

These analyses test for immediate effects of NFT on working memory performance. We used a simple t test to compare between NFT + WMT and WMT-only groups. No significant differences were found between the groups in the mean level of difficulty reached at the end of the 10th WMT session, with a result actually showing lack of group differences [$t(38) = 0.32$, $p = 0.749$, $BF_{10} = 0.32$, supporting H0].

WMT influence on resting-state EEG

Following Langer et al. (2013) who showed that WMT affects resting-state EEG, we conducted a repeated measures ANOVA with Time (pre/post-training) as the independent variable and frontal Theta as the dependent variable conducted to compare between WMT and active control training group. This analysis revealed a non-significant interaction between Group and Time [$F(1, 38) = 2.13$, $p = 0.151$, $\eta_p^2 = 0.05$, $BF_{10} = 0.73$] and we could not accept H0 either. We used a similar analysis to compare between WMT and silent control. Here, too, we found a non-significant interaction between Group and Time [$F(1, 58) = 1.97$, $p = 0.162$, $\eta_p^2 = 0.033$, $BF_{10} = 0.73$] and again we could not accept H0. Furthermore, we performed a paired sample t test to investigate whether frontal Theta had been elevated as a result of WMT. No significant differences were found between pre-post-evaluations as a result of WMT, with results supporting

¹¹ Since two participants from the combined training (NFT + WMT) had extreme NFT results, we calculate the interaction without them and the results remained similar [$F(9, 324) = 0.98$, $p = .461$, $\eta_p^2 = .02$, $BF_{10} = 0.06$, indicating strong support for H0].

¹² Similar results were found after excluding the two participants from the combined training (NFT + active control training group) who had extreme NFT results [$F(9, 324) = 0.51$, $p = .864$, $\eta_p^2 = .01$, $BF_{10} = 0.02$, indicating strong support for H0].

Table 2 Comparison between WMT group and silent control group in executive functions indices

Executive functions indices and mental rotation test	Silent control		WMT		<i>t</i> test ^a	<i>p</i> value	BF10
	Mean difference post-pre	SD	Mean difference post-pre	SD			
Alternative-cost RT (ms)	-53	91	-94	137	1.37	0.168	0.06
6-Choice tau (ms)	-15	83	-110	111	-2.85	<0.001	14.28
Alternation-cost RT (ms)	-134	377	-243	293	-1.13	0.258	0.79
Alternation-cost in error proportion	-0.04	0.16	-0.02	0.09	0.43	0.656	0.21
Anti-saccade error proportion	-0.07	0.11	-0.1	0.06	-1.17	0.239	0.83
Mental rotation RT (ms)	-712	798	-699	762	0.06	0.951	0.28
Mental rotation in errors	-0.09	0.16	-0.06	0.09	0.39	0.395	0.35

^aMinus signs indicating an improvement in the WMT group from pre-test to post-test in comparison to the silent control group

H0 [Fp1, *t* (19) = 0.16, *p* = 0.871, BF10 = 0.23; Fp2, *t* (19) = 0.283, *p* = 0.779, BF10 = 0.24]. These results indicate that there were no changes in resting-state frontal Theta as a result of WMT.

Transfer effects

In the following analyses, we only looked for time by group interactions. Table S3 in the supplementary materials online presents the means and SDs of the executive functions indices, separately for each of the three evaluations (pre-post and follow-up). We extracted six indices from the executive function battery. All indices had very good retest reliabilities (ranging between 0.52 and 0.77) and internal reliability (0.78–0.95). For details also see Gordon et al. (2018).

WMT

To examine the influence of WMT on executive functions, we calculated a difference score (post minus pre-training) for each participant for all the executive functions indices and conducted a simple comparison between WMT and silent control groups. As in Shahar and Meiran’s study (2015), the only executive functions index showing a significant WMT-related improvement relative to silent control was the 6-choice Tau index. No other effect reached significance, and notably, in most cases, we could even endorse the null hypothesis (BF10 < 0.33, Table 2). For further investigation, we also calculated a mean difference score (post minus pre-training) for the 6-choice Tau index in the different training groups. As can be seen in Table 3, the only significant difference was found between WMT and silent control [*t* (58) = 2.85, *p* < 0.001, BF10 = 14.28] indicated strong support for H1. The improvement, as in Shahar and Meiran’s study (2015), remained significant for the long-term as well [*t* (58) = 2.76, *p* = 0.012, BF10 = 11.11, for the results concerning the delayed post-test]. We also compared between

WMT and active control training group in the same manner. No significant differences were found between these groups for the short term [*t* (38) = 0.65, *p* = 0.262, BF10 = 0.53], although H0 could not be endorsed either. Still, in the follow-up evaluation, a significant difference was found, and even though it tended slightly in favor of H1, it was still quite weak and did not pass the Bayesian threshold [*t* (38) = 1.79, *p* = 0.038, one sided, BF10 = 2.03]. In sum, although we succeeded in replicating previous results, there was no difference between active control training and WMT groups in the 6-choice Tau index (found only for the short term), emphasizing the potential impact of placebo or other non-specific factors in this transfer effect.

NFT

In this section, we compared between NFT group to silent control group (in one analysis) and to WMT alone group (in another analysis). To this end, we calculated a difference score (post minus pre-training) for all the executive

Table 3 Descriptive statistics and the *t* test simple comparison results of the mean difference score (post minus pre-training) for the 6-choice Tau index in the different training groups

Training protocol	Mean difference post-pre	SD	<i>t</i> test	<i>p</i> value	BF10
NFT + WMT	-88	129.4	1.85	0.068	3.41
NFT + active control training	-85	76	1.28	0.088	5.32
NFT	-47	132	0.52	0.601	0.45
WMT	-110	107.7	2.85	<0.001	14.28
Active control training	-87	75	1.70	0.092	5.42

Training groups compared to silent control group. The mean difference score differed significantly from zero, indicating an improvement

Table 4 Comparison between NFT group and silent control group in executive functions indices

Executive functions indices and mental rotation test	Silent control		NFT		<i>t</i> test ^a	<i>p</i> value	BF10
	Mean difference post-pre	SD	Mean difference post-pre	SD			
Alternative-cost RT (ms)	-53	91	-72	97	-0.79	0.427	0.55
6-Choice tau (ms)	-15	83	-47	139	-0.67	0.501	0.49
Alternation-cost RT (ms)	-134	377	-181	191	-0.5	0.622	0.42
Alternation-cost in error proportion	-0.04	0.16	-0.05	0.1	-0.28	0.778	0.35
Anti-saccade error proportion	-0.07	0.11	-0.09	0.12	-0.73	0.459	0.53
Mental rotation RT (ms)	-712	798	-597	831	0.49	0.627	0.31
Mental rotation in errors	-0.09	0.16	-0.05	0.1	1.07	0.288	0.41

The dependent variable is the mean difference score (post- minus pre-training)

^aMinus signs indicate an improvement in the NFT group from pre-test to post-test in comparison to the silent control group

functions indices. No significant differences were found between NFT and silent control groups (Table 4), and also between NFT group and WMT group (Table S4 in the supplementary materials online). Given the lack of differences between NFT and WMT, we found no reason for a comparison between NFT and active control training groups.

NFT+WMT

To investigate whether the combination of NFT and WMT added any benefit to the improvement of far-transfer indices beyond the single protocols, we calculated the triple interaction between NFT (Yes/No), WMT (Yes/No) and Time (pre and post-training evaluation). The only index found significant was “mental rotation in errors”, supporting H1 [$F(1, 72) = 11.73, p < 0.001, \eta_p^2 = 0.09, BF10 = 9.38$]. Contrast analysis that probed this interaction revealed significant improvement in mental rotation errors in the difference score (post- minus pre-training) for the NFT + WMT group as compared to NFT + active control training group [$t(38) = 3.73, p < 0.001, BF10 = 44.29$], indicating strong support for H1. Furthermore, these differences remained significant (but only according to null hypothesis testing and not according to Bayesian inference) for the long term [$t(38) = 2.20, p = 0.03, BF10 = 1.85$]. Comparison between NFT + WMT and silent control groups revealed no advantage in the short-term post-test [$t(54) = 0.95, p = 0.35, BF10 = 0.39$], although H0 could not be endorsed either. Similar results were found in the delayed post-test [$t(50) = 0.65, p = 0.52, BF10 = 0.35$]. Although unequal in the total duration of the training session (but strictly equivalent in WMT duration), we also compared between NFT + WMT and WMT groups. Significant improvement (by null hypothesis testing but not by Bayesian inference) in mental rotation errors was found in NFT + WMT group relative to WMT

group [$t(38) = 2.00, p = 0.052, BF10 = 1.46$], with possible small support for H1. These differences did not remain significant for the long term [$t(38) = 0.97, p = 0.34, BF10 = 0.54$], although H0 could not be endorsed either. When comparison was between NFT + active control training and active control training groups, the contrast reached significance and indicated support for H1 [$t(38) = 2.82, p = 0.008, BF10 = 6.03$]. These differences did not remain significant in the long term [$t(38) = 1.17, p = 0.25, BF10 = 0.56$], although H0 could not be endorsed either. Figure 13 presents the results of the proportion of mental rotation errors index in the different training groups.

Discussion

The current randomized controlled training study involved a relatively large sample of healthy young adults. We evaluated the impact of the parietal individual Alpha

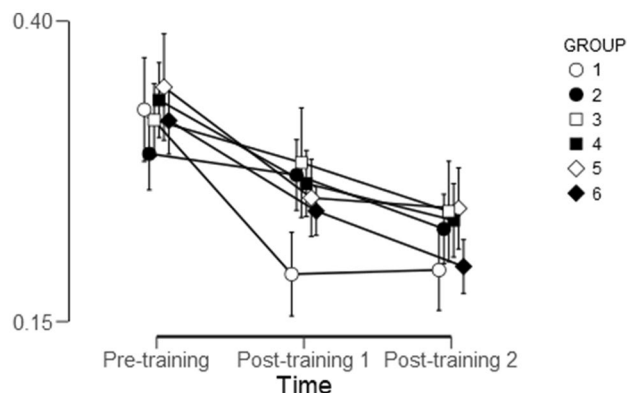


Fig. 13 Proportion of mental rotation errors as a function of time and group. Error bars represent 95% confidence intervals. ACT active control training group

frequency (IAF + 2 Hz) NFT, WMT and their combination (NFT + WMT) on executive functions as compared to silent control, active control training and NFT + active control training groups. Electrophysiological and behavioral effects were evaluated before and after training by resting-state EEG recordings and the brief executive function battery, respectively. NFT (in all three NFT groups) resulted in a significant increase of mean upper Alpha amplitude. Moreover, NFT demonstrated learning specificity in the Alpha frequency band. Similarly, the groups who received behavioral training (WMT/ active control training) demonstrated significant practice effect, i.e., improvement in the training tasks. Results from the groups who received combination of training protocols showed that NFT had no influence on performance in the WMT (or active control training) tasks performed immediately afterwards. Moreover, transfer effects were rare, and when found, small and short lasting. Specifically, while all training groups improved in the executive function transfer tasks (post-training and follow-up evaluations), WMT group was the only group that demonstrated significant improvement in the near transfer working memory task (choice reaction time task) as compared to the silent control group, however, not when compared to the active control training group. The NFT + WMT group was the only group that demonstrated significant improvement in the far transfer visual working memory task (mental rotation test) as compared to the NFT + active control training and WMT-only groups but not when compared to the silent control group. These results indicate that although the combination of NFT and WMT improved mental rotation ability for the short and long term after training, this improvement was similar to that seen in the silent control group. Finally, significant non-specific impact of NFT on resting-state spectral topography was found, suggesting that NFT had an influence on resting-state EEG. Given these findings, it cannot be ruled out that perhaps other (untested) functions have been influenced by training. We conclude that the combination of the training protocols tested in this study has very narrow influence on the executive functions that we assessed.

Concerning the effect of WMT on executive function, we were able to replicate previous results of this WMT protocol (Shahar and Meiran 2015) of significant long-lasting change in the working memory Tau index as a result of WMT. However, there was no difference between active control training and WMT-only groups in this index. Therefore, we conclude that this WMT is task specific (Soveri, Antfolk, Karlsson, Salo & Laine, 2017), and that placebo or other non-specific factors may contribute to the impact of training.

As to upper Alpha NFT, the current results do not support previous findings showing a positive impact of upper Alpha NFT on working memory (Nan et al., 2012) and mental rotation abilities (e.g. Zoefel et al., 2011). The negative findings are important because, at minimum, they cast doubt on the

potential efficacy of this intervention in enhancing executive function in healthy young adults. Still it is important to mention that there are some discrepancies between the NFT parameters that we employed and those used in previous studies. First, a “baseline” of upper Alpha power (i.e., the power that is used for the participants’ feedback) was determined in the first 3 min of the training (i.e., while participants were already trying to increase their upper Alpha power) and not during complete rest. This procedure may have caused an under-estimation of the learning since the baseline was possibly already higher than in rest. However, we demonstrated upper Alpha learning nonetheless, both within and across-sessions, although these effects are possibly underestimated. Second, the visual feedback procedure during NFT was variable, with participants playing several “games”. We, therefore, cannot rule out the fact that the changes in the visual feedback themselves (as a result of changes in participants’ upper Alpha power) might have caused a less effective upper Alpha power learning, as compared with previous studies. Nonetheless, we are unaware of any study showing that such procedure should compromise learning. Actually, we reasoned that this aspect of the procedure would make NFT less boring and would thus enhance participants’ motivation. It is also important to mention that in the NFT literature, there does not appear to be any standard procedure and what seems to be agreed upon is the need to show within-session ‘trainability’ (see Gruzeiler, 2014b and also Wan et al., 2014), which we have done successfully. We acknowledge the fact that the learning effects were not specific to upper Alpha but to the entire Alpha range, see Fig. 10, and that this may have been due to the particularities of our protocol. Furthermore, and more specifically, there are three major differences between Zoefel’s (2011) and the current protocol; (1) the frequency of the training [two sessions per week, each 21 min, for 5 weeks (total training 210 min) vs. five sessions per week, each 5 min (total training 25 min) in Zoefel et al., 2011]; (2) the timing of the post-training measurement of the resting-state EEG and mental rotation test (1–5 days after training vs. immediately after NFT in Zoefel et al. 2011); (3) the mental rotation test (comparison between two three-dimensional rotated shapes vs. a reference cube compared to six other cubes, with at most one of them being convertible into the reference cube). Since the total amount of training was higher in the current study, and previous studies (Klimesch, 1999; Hanslmayr et al., 2005; Zoefel et al., 2011; Escolano et al., 2011), that evaluated mental rotation test directly after NFT, one possible reason for the non-significant changes in mental rotation scores may be the delayed evaluation of mental rotation test in the current study (as mentioned in the **Methods** section, the post-training evaluation conducted between 3–5 days after completing the tenth training session). As to Nan’s (2012) study, a somewhat different set of points is relevant.

They used a protocol involving frequent training [20 sessions, each 8 min, and 3–4 sessions per day, over a period of 15 days (total training 160 min)]. In addition, Nan et al. (2012) used a forward and reversed digit-span as working memory task, emphasizing storage capacity in declarative working memory tasks, while we used a choice reaction time task and emphasized processing (retrieval efficiency) in a procedural working memory task and they also used only a waiting list as control for NFT. Another difference between the studies is the immediate (Nan et al., 2012) vs. delayed (our study) post-test evaluation. It is, therefore, clear that future studies should use both immediate and delayed evaluations and use different working memory tasks to further investigate the discrepancy between the studies. However, the fact that we did not find a direct influence of NFT on the WMT task, at least in a working memory task that emphasize processing and set-shifting (task-switching) elements suggest once again that upper Alpha NFT has a task-specific, immediate-only influence on performance in healthy young adults.

After considering all these factors, it is still important to emphasize that replication was not the main goal of the current investigation. The goal was to examine how two relatively established intervention methods combine. Although the combination of NFT and WMT improved mental rotation ability both for the short and long term, this improvement was not significantly different from that seen in the silent control group. Therefore, the results refute the claim regarding specificity of the combined training. We can conclude that the combinations of the training protocols assessed in this study, have only small and restricted impact on the executive functions that were assessed. Nevertheless, additional explanations to the current results should also be considered; (1) the difference between the current protocols (consecutive training) compared to the simultaneous training of Hosseini et al. (2016); (2) the participants were healthy young adults with an above mean Intelligence Rating Score, a fact that may have left little room for improvement.

In regard to the resting-state EEG measurements (pre-post-training), while previous studies (Finnigan & Robertson, 2011; Langer et al., 2013) found increased resting-state frontal Theta power after WMT, our results found no such relation. Not only that we were unable to detect any WMT-related neural changes in the resting-state EEG frontal Theta following WMT, we demonstrated support for lack of differences (Bayes Factors) between pre- and post-training resting-state frontal Theta spectral measures. In fact, pre-to-post changes in resting-state EEG were only evident in the groups who received NFT. These results are in line with our previous findings concerning a lack of monotonic correlation between executive functions and resting-state EEG spectral power measures (relative power, ratio between different bands, coherence and asymmetry of Beta/Alpha and Theta/

Alpha, Gordon et al., 2018). In that study, we suggested that the widely used spectral analysis approach may not capture the full complexity of behavior-related regularities.

Two implications of the current results can be noted. First, the use of the presently Alpha NFT as a means of improving executive functions seems of limited potential, at least among middle-to high intelligence healthy young adults. Second, in regard to the difference between resting-state EEG and task-evoked EEG, our results suggest that they may represent two different mental states. Specifically, it appears that perhaps upper Alpha indicates executive functioning when it is task-evoked, but resting-state upper Alpha is unrelated to executive functions (see also Gordon et al., 2018). Thus, it remains possible that an efficient training protocol would be one based on task-evoked EEG in a manner analogous to that used by Hosseini et al. (2016), who employed near-infrared spectroscopy for NFT.

Limitations

There are limitations to the current study. First, we can only refer to electrode locations and not to brain locations, due to the characteristics of the EEG used. We used 19-electrode EEG caps, which in turn limited our ability to employ localization procedures. Second, our conclusions are valid only for the parameters we examined and not to resting-state EEG in general. It is possible that different parameters would have yielded significant correlations with executive function, something that future studies should examine. Third, although our executive function battery provided a relatively reasonable coverage of the executive function domain, this coverage was nonetheless incomplete. For example, instead of measuring working memory updating, we indirectly measured other working memory functions such as capacity and retrieval rate. Thus, we cannot rule out the possibility of very small correlations given the sample size that we used. Fourth, the single training groups were not equal in training time as compared to the combined groups. Future studies should determine if this is a true limitation. Fifth, although the purpose of the current investigation was not to detect direct changes in performance as a result of training, it would have been helpful if we had added immediate as well as late measurements to the investigation. Lastly, some NFT studies performed a double-blind procedure (sham-control group as placebo). The inclusion of a sham-control group (fake feedback) is crucial to control for unspecific factors and to be able to estimate the effectiveness of NFT above placebo. However, there are studies demonstrating no significant difference between the two types of control on behavioral measures (e.g. Logemann et al. 2010). In any event, the fact that the training resulted in very narrow

transfer effect on executive function makes this issue less critical in the present case.

Conclusion

In a relatively large group of healthy young adults, we demonstrated that the WMT and NFT protocols including their combinations tested in this study had very limited influences on executive function abilities. WMT improved only to a near transfer task, and NFT by itself had no influence on executive functions. The NFT + WMT group showed short- and long-term improvements to a far transfer task of mental rotation ability. However, this improvement did not differ significantly from the silent control group.

Acknowledgements The study was supported by the Israel Defense Forces (IDF) Medical Corps and Directorate of Defense Research & Development, Israeli Ministry of Defense (IMOD DDR&D). We would like to thank Col. Dr. Erez Carmon, Col. Dr. Eyal Fructer and Lut. Idit Oz for their belief in the value of this research and, therefore, supporting this study. Dr. Arik Eisenkraft and Dr. Linn Wagnert-Avraham from the Institute for Research in Military Medicine, the Hebrew University Faculty of Medicine, Jerusalem for their help in purchasing the equipment needed for the study and handling the budget. As well as Mr. Erez Gordon and Prof. Robert Thatcher for advising in regard to analyzing the data.

Author contributions SG had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: SG, DT, and NM. Acquisition of data: SG, ID, and NS. Analysis and interpretation of data: all the authors. Drafting of the manuscript: SG, NM, ID and DT. Critical revision of the manuscript for important intellectual content: SG and NM. Statistical analysis: SG, OA, DG, NM and AS-R.

Funding The study was supported by the Israel Defense Forces (IDF) Medical Corps and Directorate of Defense Research & Development, Israeli Ministry of Defense (IMOD DDR&D).

Compliance with ethical standards

Conflict of interest All the authors report no financial, personal or other relationships with commercial interests.

References

Alfonso, J. P., Caracul, A., Delgado-Pastor, L. C., & Verdejo-García, A. (2011). Combined goal management training and mindfulness meditation improve executive functions and decision-making performance in abstinent polysubstance abusers. *Drug and alcohol dependence*, 117, 78–81.

Altmann, E. M., & Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review*, 115, 602–639.

Arbiv, D. C., & Meiran, N. (2015). Performance on the antisaccade task predicts dropout from cognitive training. *Intelligence*, 49, 25–31.

Association, W. M. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*, 310, 2191–2194.

Au, J., Karsten, C., Buschkuhl, M., & Jaeggi, S. M. (2017). Optimizing transcranial direct current stimulation protocols to promote long-term learning. *Journal of Cognitive Enhancement*, 1, 65–72.

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuhl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: a meta-analysis. *Psychonomic bulletin & review*, 22, 366–377.

Basar, E. (2006). The theory of the whole-brain-work. *International Journal of Psychophysiology*, 60, 133–138.

Burgess, A., & Gruzelier, J. (1993). Individual reliability of amplitude distribution in topographical mapping of EEG. *Electroencephalography and clinical Neurophysiology*, 86, 219–223.

Buschkuhl, M., Jaeggi, S. M., & Jonides, J. (2012). Neuronal effects following working memory training. *Developmental Cognitive Neuroscience*, 2, 167–179.

Cooper, N. R., Croft, R. J., Dominey, S. J., Burgess, A., P., & Gruzelier, J., H (2003). Paradox lost? Exploring the role of alpha oscillations during externally vs. internally directed attention and the implications for idling and inhibition hypotheses. *International Journal of Psychophysiology*, 47, 65–74.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science*, 318, 1387–1388.

Ditye, T., Jacobson, L., Walsh, V., & Lavidor, M. (2012). Modulating behavioral inhibition by tDCS combined with cognitive training. *Experimental Brain Research*, 219, 363–369.

Egner, T., Zech, T. F., & Gruzelier, J. H. (2004). The effects of neurofeedback training on the spectral topography of the electroencephalogram. *Clinical Neurophysiology*, 115(11), 2452–2460.

Enriquez-Geppert, S., Huster, R. J., & Herrmann, C. S. (2017). EEG-neurofeedback as a tool to modulate cognition and behavior: a review tutorial. *Frontiers in Human Neuroscience*, 11, 51–58.

Escolano, C., Aguilar, M., & Minguez, J. (2011). EEG-based upper alpha neurofeedback training improves working memory performance. In *Engineering in medicine and biology society, EMBC, 2011 Annual International Conference of the IEEE*, pp. 2327–2330.

Finnigan, S., & Robertson, I. H. (2011). Resting EEG theta power correlates with cognitive performance in healthy older adults. *Psychophysiology*, 48, 1083–1087.

Fox, K. C., Dixon, M. L., Nijeboer, S., Girm, M., Floman, J. L., Lifshitz, M., ... Christoff, K. (2016). Functional neuroanatomy of meditation: A review and meta-analysis of 78 functional neuroimaging investigations. *Neuroscience and Biobehavioral Reviews*, 65, 208–228.‏.

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, 86, 186–204.

Gal, R. (1986). *A portrait of the Israeli soldier*. Westport: Greenwood Publishing Group.

Gevins, A., & Smith, M. E. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral Cortex*, 10, 829–839.

Gevins, A., Zeitlin, G. M., Yingling, C. D., Doyle, J. C., Dedon, M. F., Schaffer, R. E., Yeager, C. L. (1979). EEG patterns during 'cognitive' tasks. I. Methodology and analysis of complex behaviors. *Electroencephalography and Clinical Neurophysiology*, 47, 693–703.

Gevins, A. S., Zeitlin, G. M., Doyle, J. C., Schaffer, R. E., & Callaway, E. (1979). EEG patterns during 'cognitive' tasks. II. Analysis of controlled tasks. *Electroencephalography and Clinical Neurophysiology*, 47, 704–710.

- Gevins, A. S., Zeitlin, G. M., Doyle, J. C., Yingling, C. D., Schaffer, R. E., Callaway, E., & Yeager, C. L. (1979). Electroencephalogram correlates of higher cortical functions. *Science*, 203, 665–668. ‏.
- Gordon, S., Todder, D., Deutsch, I., Garbi, D., Getter, N., & Meiran, N. (2018). Are resting state spectral power measures related to executive functions in healthy young adults? *Neuropsychologia*, 108, 61–72.
- Goyal, M., Singh, S., Sibinga, E. M., Gould, N. F., Rowland-Seymour, A., Sharma, R., Ranasinghe, P. D. (2014). Meditation programs for psychological stress and well-being: a systematic review and meta-analysis. *JAMA Internal Medicine*, 174, 357–368.
- Gruzelier, J. H. (2014a). EEG-neurofeedback for optimizing performance. I: a review of cognitive and affective outcome in healthy participants. *Neuroscience & Biobehavioral Reviews*, 44, 124–141.
- Gruzelier, J. (2014b). EEG-neurofeedback for optimizing performance. III: a review of methodological and theoretical considerations. *Neuroscience Biobehavioral Reviews*, 44, 159–182.
- Guez, J., Rogel, A., Getter, N., Keha, E., Cohen, T., Amor, T., & Todder, D. (2015). Influence of electroencephalography neurofeedback training on episodic memory: A randomized, sham-controlled, double-blind study. *Memory*, 23, 683–694. ‏.
- Hanslmayr, S., Sauseng, P., Doppelmayr, M., Schabus, M., & Klimesch, W. (2005). Increasing individual upper alpha power by neurofeedback improves cognitive performance in human subjects. *Journal of Applied Psychophysiology and Biofeedback*, 30, 1–10.
- Harrison, T. L., Shipstead, Z., Hicks, K. L., Hambrick, D. Z., Redick, T. S., & Engle, R. W. (2013). Working memory training may increase working memory capacity but not fluid intelligence. *Psychological Science*, 24, 2409–2419. ‏.
- Hossein, S. H., Pritchard-Berman, M., Sosa, N., Ceja, A., & Kesler, S. R. (2016). Task-based neurofeedback training: A novel approach toward training executive functions. *NeuroImage*, 134, 153–159.
- Hummel, F., Andres, F., Altenmüller, E., Dichgans, J., & Gerloff, C. (2002). Inhibitory control of acquired motor programs in the human brain. *Brain*, 125, 404–420.
- Hyun, J. S., & Luck, S. J. (2007). Visual working memory as the substrate for mental rotation. *Psychonomic Bulletin and Review*, 14, 154–158.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. In *Proceedings of the National Academy of Sciences*, 105, pp. 6829–6833.
- JASP Team. (2018). JASP (Version 0.8.6) [Computer software].
- Johnstone, S. J., Roodenrys, S. J., Johnson, K., Bonfield, R., & Bennett, S. J. (2017). Game-based combined cognitive and neurofeedback training using Focus Pocus reduces symptom severity in children with diagnosed AD/HD and subclinical AD/HD. *International Journal of Psychophysiology*, 116, 32–44. ‏.
- Just, M. A., & Carpenter, P. A. (1985). Cognitive coordinate systems: accounts of mental rotation and individual differences in spatial ability. *Psychological Review*, 92, 137–172.
- Kaiser, D. A., & Serman, M. B. (2001). Automatic artifact detection, overlapping windows, and state transitions. *Journal of Neurotherapy*, 4, 85–92. ‏.
- Kaplan, Z., Weiser, M., Reichenberg, A., Rabinowitz, J., Caspi, A., Bodner, E., & Zohar, J. (2002). Motivation to serve in the military influences vulnerability to future posttraumatic stress disorder. *Psychiatry Research*, 109, 45–49.
- Karbach, J., & Verhaeghen, P. (2014). Making working memory work: A meta-analysis of executive-control and working memory training in older adults. *Psychological Science*, 25, 2027–2037.
- Kashdan, T. B., & Rottenberg, J. (2010). Psychological flexibility as a fundamental aspect of health. *Clinical Psychology Review*, 30, 865–878.
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews*, 29, 169–195.
- Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P., & Winkler, T. (1999). Paradoxical alpha synchronization in a memory task. *Cognitive Brain Research*, 7, 493–501.
- Koch, I., Gade, M., Schuch, S., & Philipp, A. M. (2009). The role of inhibition in task switching: A review. *Psychonomic Bulletin and Review*, 17, 1–14.
- Laming, D. (1979). Choice reaction performance following an error. *Acta Psychologica*, 43, 199–224.
- Langer, N., von Bastian, C. C., Wirz, H., Oberauer, K., & Jäncke, L. (2013). The effects of working memory training on functional brain network efficiency. *Cortex*, 49, 2424–2438.
- Logemann, H. A., Lansbergen, M. M., Van Os, T. W., Böcker, K. B., & Kenemans, J. L. (2010). The effectiveness of EEG-feedback on attention, impulsivity and EEG: a sham feedback controlled study. *Neuroscience Letters*, 479, 49–53.
- Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General*, 129, 4–26.
- Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1124–1140.
- Mayr, U., & Kliegl, R. (2003). Differential effects of cue changes and task changes on task-set selection costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 362–372.
- Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology*, 41, 211–253.
- Meiran, N., Hsieh, S., & Chang, C. (2011). “Smart inhibition”: Electrophysiological evidence for the suppression of conflict-generating task rules during task-switching. *Cognitive, Affective, and Behavioral Neuroscience*, 11, 299–308.
- Meiran, N., Hsieh, S., & Dimov, E. (2010). Resolving task rule incongruence during task switching by competitor rule suppression. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 992–1002.
- Melby-Lervåg, M., & Hulme, C. (2016). There is no convincing evidence that working memory training is effective: A reply to Auet al. (2014) and Karbach and Verhaeghen (2014). *Psychonomic Bulletin and Review*, 23, 324–330.]
- Mirifar, A., Beckmann, J., & Ehrlenspiel, F. (2017). Neurofeedback as supplementary training for optimizing athletes’ performance: A systematic review with implications for future research. *Neuroscience and Biobehavioral Reviews*, 75, 419–432.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49–100.
- Nan, W., Rodrigues, J. P., Ma, J., & Qu, X. (2012). Individual alpha neurofeedback training effect on short term memory. *Current Psychiatry Reports*, 14, 536–542.
- Oken, B. S., Flegal, K., Zajdel, D., Kishiyama, S., Haas, M., & Peters, D. (2008). Expectancy effect: impact of pill administration on cognitive performance in healthy seniors. *Journal of Clinical and Experimental Neuropsychology*, 30, 7–17.
- Park, S. H., Seo, J. H., Kim, Y. H., & Ko, M. H. (2014). Long-term effects of transcranial direct current stimulation combined with computer-assisted cognitive training in healthy older adults. *Neuroreport*, 25, 122–126.
- Patel, R., Spreng, R. N., & Turner, G. R. (2013). Functional brain changes following cognitive and motor skills training: a quantitative meta-analysis. *Neurorehabilitation and neural repair*, 27, 187–199.
- Rabbitt, P. M. (1966). Errors and error correction in choice-response tasks. *Journal of Experimental Psychology*, 71, 264–272.

- Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., Engle, R. W. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *Journal of Experimental Psychology: General*, 142, 359.
- Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, 56, 356–374.
- Schmiedek, F., Oberauer, K., Wilhelm, O., Süss, H. M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, 136, 414–429.
- Shahar, N., & Meiran, N. (2015). Learning to control actions: transfer effects following a procedural cognitive control computerized training. *PLoS One*, 10, e0119992.
- Shahar, N., Pereg, M., Teodorescu, A. R., Moran, R., Karmon-Presser, A., & Meiran, N. (2018). Formation of abstract task representations: Exploring dosage and mechanisms of working memory training effects. *Cognition*, 181, 151–159.
- Shahar, N., Teodorescu, A. R., Usher, M., Pereg, M., & Meiran, N. (2014). Selective influence of working memory load on exceptionally slow reaction times. *Journal of Experimental Psychology: General*, 143, 1837.
- Shepard, J., & Metzler (1971). Mental rotation of three-dimensional objects. *Science*, 701–703.
- Soffer-Dudek, N., Todder, D., Shelef, L., Deutsch, I., & Gordon, S. (2018). A neural correlate for common trait dissociation: Decreased EEG connectivity is related to dissociative absorption. *Journal of Personality*.
- Soveri, A., Antfolk, J., Karlsson, L., Salo, B., & Laine, M. (2017). Working memory training revisited: A multi-level meta-analysis of n-back training studies. *Psychonomic Bulletin and Review*, 1–20.
- Van Albada, S. J., & Robinson, P.A. (2007). Transformation of arbitrary distributions to the normal distribution with application to EEG test–retest reliability. *Journal of neuroscience methods*, 161, 205–211.
- Wan, F., Nan, W., Vai, M. I., & Rosa, A. (2014). Resting alpha activity predicts learning ability in alpha neurofeedback. *Frontiers in Human Neuroscience*, 8, 500.
- Zoefel, B., Huster, R. J., & Herrmann, C. S. (2011). Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance. *Neuroimage*, 54, 1427–1431.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.