Title

Working memory updating training modulates a cascade of event-related potentials depending on task load

Authors and Affiliations

Juha Salmi[1,2,3*], Adrià Vilà-Balló[1,4,5,6*], Anna Soveri[7], Carles Rostan[5], Antoni Rodríguez-Fornells[6,8], Minna Lehtonen[10], Matti Laine[1,11]

* Authors contributed equally to the manuscript

[1] Department of Psychology, Åbo Akademi University, Tehtaankatu 2, Turku, Finland
[2] Department of Psychology, University of Turku, Turku, Finland
[3] Institute for Advanced Studies, University of Turku, Turku, Finland
[4] Headache and Neurological Pain Group, Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain
[5] Department of Psychology, Faculty of Education and Psychology, University of Girona, Girona, Spain
[6] Cognition and Brain Plasticity Group [Bellvitge Biomedical Research Institute – IDIBELL], L'Hospitalet de Llobregat, Barcelona, Spain.
[7] Department of Clinical Medicine, University of Turku, Turku, Finland
[8] Department of Cognition, Development and Educational Psychology, Faculty of Psychology, University of Barcelona, Barcelona, Spain
[9] Catalan Institution for Research and Advanced Studies, ICREA, Barcelona, Spain
[10] MultiLing Center for Multilingualism in Society across the Lifespan, Department of Linguistics and Scandinavian studies, University of Oslo, Oslo, Norway
[11] Brain and Mind Center, University of Turku, Turku, Finland

Corresponding author

Juha Salmi
E-mail: juha.salmitaival@utu.fi
Phone: +358405118678
Address: Department of Psychology, Åbo Akademi University, Fabriksgatan 2, 20500 Åbo, Finland

Abstract

The brain mechanisms of working memory (WM) training in humans remain unclear. Here we examined how WM updating training modulates a cascade of event-related potentials (ERPs) elicited at different processing stages. We hypothesized that WM updating training results to decreases in the early responses reflecting stimulus selection and response preparation, and increases the late slow responses reflecting maintenance of to-be-remembered materials. Healthy adults were randomized to a WM updating group that trained an adaptive dual n-back task (n=20), and an active control group that played a computer game (n=20). Both groups performed three 25-min training sessions per week for five weeks. Pretest-posttest comparisons showed that the training group significantly improved their performance as compared to the active controls, but this was limited to the trained task. In line with our hypothesis, P2-N2-P3 complex showed changes from pre- to posttest. In the training group this was observed as decreased load-effect while in the control group there was an opposite pattern at some latencies. Slow waves elicited during the maintenance were decreased in the easy task and increased in the difficult task. Taken together, our findings suggest that the early and late ERPs are differentially affected by training. When task demands are high, training may lead to an improved ability to actively maintain several stimuli in memory, and when they are low, training results in more efficient processing and automatization.

Keywords: ERP, load, n-back task, training, working memory updating

Introduction

Working memory (WM) allows us to maintain, manipulate, and update information contents in mind (Baddeley 1986). Due to its fundamental role in cognition, coupled with its limited capacity, WM has been a key target in intervention studies aiming to induce plasticity in human executive functions (Constantinidis and Klingberg 2016). However, as the behavioral outcomes of WM training have been disputed (Melby-Lervåg and Hulme 2013, Melby-Lervåg et al. 2016, Soveri et al. 2017a), a better understanding of the underlying brain mechanisms of WM training is called for. Hence, the most recent meta-analyses suggest that training-related transfer is observed mostly in the tasks that are structurally similar to the trained tasks (Soveri et al. 2017a). Studies investigating the underlying brain mechanisms could potentially pinpoint in more detail the influence of training at different stages of the perception-action continuum (see Salmi et al. 2018).

Human brain imaging studies have found modulations of large-scale brain networks by WM training (Constantinidis and Klingberg 2016, Salmi et al. 2018). While sensory-motor training not requiring WM has been traditionally associated mostly with activation decreases (Chein and Schneider 2005), in WM training studies activation increases have also been frequently reported (Salmi et al. 2018). Decreased brain activity following training is thought to reflect higher automaticity (Chein and Schneider 2012), possibly explained by increased sensitivity to detect stimuli (Rainer and Miller 2000). Higher activity, in turn, has been associated with an enhanced role of controlled processing (e.g., attention, WM), possibly reflecting higher number of neurons engaged or higher firing rates (Qi et al. 2011, Meyers et al. 2011). In the WM training literature, it has also been debated whether the practice effects result in the

recruitment of new functional systems (re-organization of the WM networks, Kelly et al. 2006, see also Buschkuehl et al. 2012), or if the neuronal plasticity is limited to the WM networks that were activated by the same tasks already prior to training (Constantinidis and Klingberg, 2016). Current empirical evidence mostly stemming from functional magnetic resonance imaging (fMRI) studies supports the latter view (Salmi et al. 2018). It should be noted, however, that the link between increases and decreases of brain activity is still speculative (Constantinidis and Klingberg 2016). The complexity of the issue is further evidenced by the relativity of the neuronal changes to behavioral outcomes, the effect of cognitive load in the testing task on the observed brain activity, and general difficulties in defining baselines. In a typical task-based fMRI study, baseline is defined by another experimental task, meaning that the effects are relative to another condition. Due to the limitations of fMRI in direct comparisons of the pretest and posttest effects, and even more so because of its temporally sluggish signal, more evidence of how WM training affects particular temporal patterns in brain activity is clearly needed.

Flexible updating of WM contents is a key aspect of executive functioning (Miyake et al. 2000). Updating refers to refreshment of WM contents so that the information maintained can be linked to the ongoing task or goal at hand (e.g., Morris and Jones, 1990). Other WM component functions associated with updating include selection of incoming information, inhibition of the irrelevant information, and continuous monitoring of performance. Specific component functions required in updating are thought to vary depending on the task demands: when the executive demands are low, there are more resources available for the active maintenance of relevant information, and vice versa (Ecker et al. 2010, Ecker et al. 2013, Bailey et al. 2016, Botto et al. 2014,

Vilà-Balló et al. 2018). Indeed, a few studies specifically focusing on WM maintenance have reported enhanced event-related potential (ERP) responses in the easy condition in which the participant should be able to keep the previous stimulus actively in mind, as compared with the more difficult condition in which constant maintenance is getting difficult because of the intervening stimuli (see Bailey et al. 2016, Vilà-Balló et al. 2018). WM training studies conducted with fMRI have provided evidence that decreased brain activity is mostly observed in brain areas involved in earlier processing stages, while increased brain activity is observed in brain areas such as the prefrontal cortex that are involved in higher-level functions (Salmi et al. 2018). This raises a question as to whether training could improve maintenance in the difficult condition, and in the easier tasks, in turn, lead to automatization of perceptual processing. However, the limited temporal resolution of fMRI has not been able to address the training effects on specific WM subfunctions or processing stages.

Despite its better temporal resolution, previous ERP studies on WM training have not focused on separating between WM component processes (e.g., Gevins et al. 1997, Langer et al. 2013). Yet one ERP study found that training modulated the contralateral delay activity that presumably reflects early WM processes such as maintenance of active task-relevant information in WM (Kundu et al. 2013). In addition, another ERP study observed that training influences interference control by increasing activity over the posterior regions (Oelhafen et al. 2013). Although active maintenance is amongst the most thoroughly examined WM component functions (Levy and Goldman-Rakic 2000), there is currently no direct evidence whether WM updating training influences maintenance mechanisms in the human brain. Considering that updating has been extensively used as a method to train WM due to its role in refreshing the WM contents

and linking those to the goal or task at hand (e.g., Soveri et al. 2017), it is surprising that it has not been at focus in previous ERP studies.

WM modulates a cascade of ERP responses from early to late latencies. One of the earliest components is the P2 (a positive waveform peaking at 200 ms post stimulus) response that is reflective of sensory cortical functions contributing to selection of information (see Crowley and Colrain 2004). P2 has been reported, not only in updating tasks (e.g., McEvoy et al. 1998 Rämä et al. 2000, Lenartowich et al. 2010, Luu et al. 2014, Dong et al. 2015) but also in other types of WM tasks (e.g., Lefevbre et al. 2005, Marchand et al. 2006, Mecklinger and Pfeifer 1996, Ruchkin et al. 1995). Similarly, the following N2 response is modulated in WM tasks (Dong et al. 2015, Luu et al. 2014, Mecklinger and Pfeifer 1996), and it is thought to reflect, for instance, maintenance of context information (Azizian et al. 2006) and detection of novel stimuli (Folstein and van Petten 2008). Along the same lines, modulations of the subsequent P3 component are frequently observed. Especially its latter subcomponent, P3b, is associated to WM updating, allocation of attentional resources, and/or amount of resources demanded in the current task (Donchin et al. 1986, Dien et al. 2004, Lenartowicz et al. 2010, Daffner et al. 2011). There is some evidence, mostly from studies utilizing differential experimental approaches such as the Sternberg task (Shiran and Brezniz 2011), a go/no-go task (Liu et al. 2017), but also a very recent study utilizing the n-back task (Covey et al. 2018), suggesting that ERPs especially at N2-P3 latencies are modulated by WM training. Although those fMRI studies have mostly encountered activation decreases at the early processing stages (Salmi et al. 2018), there are several ERP studies that have reported increased responses in the direct comparisons between the pretest and posttest responses (e.g., Berry et al. 2010, Shiran and Brezniz 2011, Covey et al. 2018). The

links between these ERP effects and fMRI findings, as well as neurophysiological recordings in non-human primates remains to be discovered.

Furthermore, previous research with the popular delayed matching-to-sample paradigm probing short-term memory has repeatedly demonstrated that maintenance of information is reflected as a slow-wave component that appears in-between stimuli (e.g., Ruchkin et al. 1995, Mecklinger and Müller 1996, Mecklinger and Pfeifer 1996, Barriga-Paulino et al. 2014). Both negative and positive slow waves with varying topographies during WM task performance have been reported and tentatively assigned to different functional roles (see Ruchkin et al. 1992, 1995). Consistent with these findings, studies in non-human primates have reported maintenance of neural activity during the retention period of the delayed matching-to-sample task (see Fuster 2000 for a review). In humans, slow negative components (NSW) persisting throughout the retention interval in WM tasks predict the number of objects to be maintained in memory (Fukuda et al. 2010, Luria et al. 2016). A slow wave related to active maintenance of WM contents between n-back trials has been reported in a few prior studies (Bailey et al. 2016, Vilà-Balló et al. 2018). There is evidence that also the amplitude of the contingent negative variation (CNV, Walter et al., 1964) during the retention interval is smaller when the memory load is higher (Ford et al. 1979; Roth et al. 1975; Roth et al. 1978), but this response can be distinguished from the other slow waves based on its scalp distribution (e.g., McEvoy et al. 1998).

We examined how WM training modulates a cascade of ERP components from early P2-N2-P3 complex locked to the stimulus to slow waves observed during the maintenance stage. Based on previous fMRI studies (for a meta-analysis see Salmi et al.

2018), we expected differential training effects at early and late processing stages. We expected that early sensory-motor processing modulated by attention would be more efficient after training, being reflected as smaller load effect. We also tested a specific hypothesis suggesting that WM updating training would make the maintenance of to-be-remembered materials more effective, seen mainly in improved performance and changes in slow waves linked to WM maintenance. This hypothesis raises from non-human primate studies demonstrating that WM training results in an increase in the number of activated dorsal and ventral prefrontal neurons during maintenance of information in WM (Qi et al. 2011). Our training task was a dual n-back task, because n-back tasks are amongst the most widely used WM tasks and because the slow wave responses related to maintenance has been identified with this paradigm (Bailey et al. 2016, Vilà-Balló et al. 2018). In an n-back task, the participant is to decide whether the current stimulus matches the one *n* steps back in the stimulus sequence. We expected that successful n-back training would enhance the use of active maintenance even under higher n-back load, reflected as slow wave amplitude increase in that condition. In contrast, in the low-load condition, slow waves may even diminish after training due to partial automatization of performance in the trained task during the practice period. To be able to examine the cascade of responses starting from the early latencies we utilized a conventional -200 – 0 ms baseline (see Gómez et al. 2017 for the effect of the analysis approach in WM studies). Furthermore, in the Supplementary Online Material (SOM) we also report analyses with a pre-stimulus baseline focusing on the WM maintenance related effect, that is similar to our prior cross-sectional study in the same participants where we found a link between positive slow wave (PSW) and WM maintenance (Vilà-Balló et al. 2018).

214

215 Materials and Methods

216

217 Participants

218 The present sample included 48 right-handed Spanish university students. Two
219 participants were excluded because of health issues (one had moderate depression
220 symptoms and the other had bulimia nervosa). One participant was excluded because he
221 did not understand the instructions during the first session, and consequently, failed to
222 respond to any target. Moreover, based on the previous literature (Marco-Pallares et al.
223 2011), 5 participants were excluded due to the lack of correct trials (minimum 20) after
224 the artifact rejection. After exclusions based on health issues and poor signal quality, the
225 final sample included 40 healthy participants (see Table 1 and SOM). All participants
226 gave their informed consent prior to the pretest and were reimbursed with 100 € after
227 study completion. The study was reviewed and accepted by the Clinical Research Ethics
228 Committee of the Bellvitge University Hospital, University of Barcelona, Spain.

229

230 Training regime

231 Before the training began, all participants took a pretest (see task description below,
232 details in SOM). After the pretest, the participants were randomized into a WM
233 updating training group or an active control group and underwent their respective
234 training for five weeks (3 sessions/week, 20-25 minutes/session). The training period
235 was followed by a posttest employing the same computerized tasks as in the pretest. The
236 task order was randomized for each testing session, and test versions (see below) were
237 counterbalanced across participants.

238

239 **Training tasks**

240 Our training regime was similar to our previous behavioral study (see Soveri et al.
241 2017b, more details in SOM). The WM updating training group practiced with a dual n-
242 back training task. It included a phonological n-back task with syllables presented
243 through headphones, and a parallel visuospatial n-back task where white squares
244 appeared in eight possible locations on the screen. The task was adaptive, i.e., the
245 difficulty level was automatically adjusted according to participant's performance
246 (above 90 % accuracy threshold for n increase and below 75 % threshold for n
247 decrease). The n could vary between 1 and 9, and each training session began with a 2-
248 back sequence. When a training session was over, a result screen was displayed. Each
249 session included 20 sequences, with each sequence containing 20 syllables and 20
250 squares. Each block in the training task included six auditory targets and six
251 visuospatial targets (four in one modality only; two in both modalities at the same time).
252 Training tasks also included lures (n-1 or n+1 targets) that appeared randomly. The
253 active control group played a video game (Bejeweled 2) with a rather low WM load (as
254 compared to the experimental task) for 20 minutes in each training session and recorded
255 their scores in personal training logs. Although Bejeweled 2 provides a score that
256 reflects progress in the game, we did not attempt to analyze these scores as it is unclear
257 which specific cognitive functions they reflect. The same computers were used for
258 playing Bejeweled 2 and for WM training. Both the training and the control sessions
259 were performed in a quiet chamber annexed to the EEG cabin. The training and the
260 control session were performed in groups of maximum four participants.

261

262 **Pre- and posttest measures**

The behavioral pre- and posttest measures were largely similar to those in Soveri et al. 2017b (see SOM). These tasks included (1) a dual n-back task similar to the one used in training but with 10 sequences, (2) a single visual n-back task with digits from 1 to 9 (see Figure 1), (3) a set shifting number-letter task, (4) verbal and visuospatial running memory WM updating tasks (including set shifting, see Soveri et al. 2017b), (5) a number substitution task (Carretti et al., 2007), and (6) verbal and visuospatial simple span tasks (digit span, Corsi block). Single n-back tasks were presented only during the EEG recording. All pre- and posttest tasks were computerized. To examine near transfer, four composite scores based on previous research were created from z-transformed scores (Soveri et al. 2017b, see SOM). As the single digit n-back task was, unlike the other WM tasks, structurally similar to the trained task (near-near-transfer), we did not include it in the composite scores.

Behavioral data analyses. The dependent variables for each behavioral task are described in the SOM section. Regarding the statistical analyses, mixed-model ANOVAs were separately performed for dual and single n-back performance and for each near-transfer composite score. These ANOVAs had one between-subjects factor (group) and one within-subjects factor (session), except for the single n-back performance which had load as another within-subjects factor. The whole sample could be used for the behavioral analyses, as no participant met the criteria of being an extreme outlier in accuracy or RTs (performance more than three times the interquartile range below or above the 1st or 3rd quartile, respectively) at pretest.

ERP experiment

Procedure: At pre- and posttest, we administered a single n-back task with digits that was adapted to simultaneous measurement of ERPs (for details, see Vilà-Balló et al. 2018, which is based on the pretest data of a partially overlapping sample). The participants responded to both target and non-target trials, performing eight 1-back (low load) sequences and sixteen 3-back (high load) sequences. Each trial began with a fixation point. After 450 ms, a digit appeared on the screen for 1500 ms. Stimulus onset asynchrony was fixed to 1950 ms. Each sequence included 48 trials, resulting in altogether 1152 trials. The order of the sequences was randomized for each participant.

Electrophysiological recording: Electroencephalogram (EEG) was recorded continuously (digitized with a sampling rate of 250 Hz, bandpass 0.01-70 Hz) using SynAmp Neuroscan amplifiers from 29 tin electrodes mounted on an elastic cap and located at standard positions (FP1/2, F3/4, C3/4, P3/4, O1/2, F7/8, T3/4, T5/6, Fz, Cz, Pz, FC3/4, FT7/8, CP3/4, TP7/8, FCz, CPz), and the left and right mastoids. Vertical eye movements were monitored by an electrode placed below the right eye. To be able to monitor the mastoid activity during the recording, the EEG was referenced on-line to the right ocular canthus (Morís et al., 2013; Vilà-Balló et al., 2017). Electrode impedances were maintained below 5 kΩ. After, the EEG signal was offline re-referenced to the mean activity at the two mastoid electrodes algebraically subtracting out the on-line reference, being the same as using on-line referencing to mastoids (Luck, 2005; Cohen, 2014).

EEG data analyses: ERPs were time-locked to the stimulus presentation first from -200 to 1950 ms time-range (baseline -200 to 0 ms). Waveforms were separately obtained from the 1-back and 3-back conditions. Epochs exceeding ±75 μV in electrooculogram

(EOG) or EEG were removed offline for further analysis using the extreme value function of the EEGlab toolbox. Also in the ERP analyses, only correct trials with RT responses slower than 120 ms or faster than 3 standard deviations from the participant's mean were considered for the analyses. The P2 (220 - 270 ms), N2 (270 - 330 ms), P3 (330 - 430 ms), and NSW (500 - 1000 ms) responses were defined based on the previous literature. The time-windows were centered on the peak activity of each component.

Different repeated measures ANOVAs for the mean amplitudes were carried out for each component. Each ANOVA included the following three within-subject factors: load (1-back vs. 3-back), session (pre vs. post), and electrode (frontal [electrode FZ], central [electrode CZ], posterior [electrode PZ]). In addition, there was one between-subject factor (training vs. control group). The selection of electrodes was based on the topography and previous articles (Vilà-Balló et al. 2018, see also Bailey et al., 2016).

To correct for possible violations of the sphericity assumption (Jennings and Wood, 1976), the Greenhouse–Geisser epsilon correction was used, and the adjusted p-values after the correction are reported. The Cohen's *f* and *d* were used as effect size measures for the ANOVAs and the *t*-tests, respectively (Cohen, 1992).

Results

Behavioral results

Dual n-back task: At pretest, the mean n-back level achieved in 10 blocks across all participants was 2.48 (*SD* = 0.554, see Table 2, Figure 2A/B). Training effects were examined with a repeated measures ANOVA on the maximum n-back level achieved, using session (pre; post) and group (training; control group) as independent variables. The results showed a statistically significant interaction between session and group ($F(1,38) = 146.789$, $p < 0.001$, $f = 1.963$), stemming from higher n-back level for the training group at the posttest, as compared to the control group (Table 2). There was also a statistically significant main effect of session ($F(1,38) = 236.469$, $p < 0.001$, $f = 2.500$), indicating that both groups improved their performance from pretest to posttest. The results also showed a statistically significant main effect of group ($F(1,38) = 48.371$, $p < 0.001$, $f = 1.128$), suggesting that overall the training group performed better than the control group. A follow-up analysis on pretest performance, however, revealed no statistically significant difference in performance between the two groups ($t(38) = 0.282$, $p = 0.780$, $d = 0.009$).

Single n-back task: This task showed the canonical load effects at pretest (Table 2). The training group and active controls performed similarly in the 1-back ($t(38) = 0.967$, $p = 0.339$, $d = 0.304$) and 3-back tasks ($t(29.51) = 0.780$, $p = 0.442$, $d = 0.247$) at pretest. A repeated measures ANOVA did not show significant interaction between group and session ($F(1,38) = 1.513$, $p = 0.226$, $f = 0.199$), neither between group, session, and load ($F(1,38) = 0.612$, $p = 0.439$, $f = 0.084$). However, there were significant main effects of session ($F(1,38) = 18.314$, $p = 0.0001$, $f = 0.694$) and load ($F(1,38) = 45.789$, $p = 0.0001$, $f = 1.283$), and an interaction between session and load ($F(1,38) = 62.567$, $p = 0.0001$, $f = 1.097$).

WM updating composite: The groups did not differ on this measure at pretest ($t(34) = 0.665$, $p = 0.510$, $d = 0.225$, Table 2). Repeated measures ANOVA showed no interaction between group and session ($F(1,34) = 2.748$, $p = 0.107$, $f = 0.285$), neither a main effect of session ($F(1,34) = 0.005$, $p = 0.944$, $f = 0.003$).

WM interference control composite. The groups did not differ on this measure at pretest ($t(33) = 0.530$, $p = 0.599$, $d = 0.181$, Table 2). Based on the results from the repeated measures ANOVA, there was neither interaction between group and session in the WM interference control composite ($F(1,33) = 1.899$, $p = 0.177$, $f = 0.239$), nor main effect of session ($F(1,33) = 0.003$, $p = 0.960$, $f = 0.003$).

Passive and active WM composites. There was no group difference in either the passive ($t(36) = 0.949$, $p = 0.349$, $d = 0.310$) or the active ($t(33) = 0.596$, $p = 0.555$, $d = 0.204$) WM composite at pretest (Table 2). Repeated measures ANOVAs did not show interactions between group and session (Passive: $F(1,36) = 0.596$, $p = 0.445$, $f = 0.128$; Active: $F(1,33) = 1.555$, $p = 0.221$, $f = 0.217$) or main effects of session (Passive: $F(1,36) = 0.206$, $p = 0.653$, $f = 0.078$; Active: $F(1,33) = 0.162$, $p = 0.690$, $f = 0.071$).

ERP results

As can be observed in Figure 3, a P2, followed by an N2 and then P3 were elicited during the stimulus selection and response preparation period (200-500 ms). After the P3 and during the maintenance period, there was a frontal NSW. Each of these components were observed in both groups and in both sessions. The following paragraphs will present the load effects (see Figures 4 and 7) and training effects (see Figures 5 and 8) for each ERP component separately.

386

387 *P2 (220 – 270 ms)*

388 For P2, there was a significant main effect of load ($F(1,38) = 31.90$, $p < 0.001$, $f =$
389 0.585), resulting from a higher response amplitudes in the 3-back than in the 1-back
390 condition (Figure 3). We also found a significant main effect of electrode ($F(2,76) =$
391 10.36, $p = 0.001$, $f = 0.522$), indicating that the P2 was larger at fronto-central sites. The
392 lack of significant main effect of group ($F(1,38) = 0.25$, $p = 0.620$, $f = 0.084$), suggested
393 that there were no overall differences in the P2 amplitude when responses were pulled
394 together across the two sessions. However, a significant main effect of session ($F(1,38)$
395 $= 11.35$, $p = 0.002$, $f = 0.547$) indicated that the amplitude of the P2 decreased in the
396 post session. A significant session × group interaction ($F(1,38) = 13.00$, $p < 0.001$, $f =$
397 0.585) was also found, but there were no significant session × load × group ($F(1,38) =$
398 3.54, $p = 0.068$, $f = 0.305$) or session × load × electrode × group ($F(2,76) = 2.10$, $p =$
399 0.147, $f = 0.234$) interactions.

400

401 Post-hoc analyses performed on the mean amplitude from central electrodes and both
402 loads, suggested that the session × group interaction resulted at least partially from
403 reduction of the P2 amplitude in the control group in the post session (post *minus* pre:
404 $t(19) = -5.02$, $p < 0.001$, $d = 0.512$), which was not observed in the training group (post
405 *minus* pre: $t(19) = 1.64$, $p = 0.871$, $d = 0.010$).

406

407 *N2 (270 – 330 ms)*

408 A significant main effect of load ($F(1,38) = 32.23$, $p < 0.001$, $f = 0.921$), suggested that
409 the N2 amplitude was larger in the the 1-back than in the 3-back condition (Figure 3). A
410 significant main effect of electrode was observed ($F(2,76) = 26.25$, $p < 0.001$, $f = 0.832$)

411 which resulted from the centro-frontal distribution of this response. Also in the N2, the
412 main effect of group was not significant ($F(1,38) = 0.33$, $p = 0.567$, $f = 0.095$). We
413 observed a significant main effect of session ($F(1,38) = 12.09$, $p = 0.001$, $f = 0.564$),
414 indicating that the N2 amplitude was larger in the post session. The session × group
415 interaction was not significant ($F(1,38) = 2.95$, $p = 0.094$, $f = 0.279$), but we observed a
416 significant interaction for session × load × group ($F(1,38) = 18.53$, $p < 0.001$, $f =$
417 0.699), suggesting that a training effect was observed when the load was accounted for.
418 No other significant interactions related to the group were detected (session × electrode
419 × group: $F(2,76) = 0.882$, $p = 0.379$, $f = 0.153$; session × load × electrode × group:
420 $F(2,76) = 2.71$, $p = 0.093$, $f = 0.268$).
421
422 Post-hoc analyses for the N2 performed on the mean amplitude of central electrodes,
423 revealed that there were group differences in the load effect (1-back minus 3-back) in
424 the post session ($t(38) = 2.92$, $p = 0.006$, $d = 0.924$) that were not observed prior to
425 training ($t(38) = -1.50$, $p = 0.142$, $d = 0.474$). A second post-hoc analysis was performed
426 to test the pre-post effects for 1-back and 3-back separately for each group. Importantly,
427 in the training group, the amplitude of the N2 was increased in the 3-back task from
428 pretest to posttest ($t(19) = -2.71$, $p = 0.014$, $d = 0.450$). In the 1-back task, no training
429 effect was observed (despite no differences were encountered for the 1-back ($t(19) =$
430 1.37, $p = 0.184$, $d = 0.140$). In the control group there was an opposite load effect: The
431 N2 amplitude increase in the post session was observed for the 1-back ($t(19) = -5.35$, p
432 < 0.001, $d = 0.613$), but not for the 3-back ($t(19) = -1.26$, $p = 0.223$, $d = 0.195$) task.
433
434 *P3 (330 – 430 ms)*

For the P3, there was a significant main effect of load ($F(1,38) = 26.68$, $p < 0.001$, $f = 0.755$), resulting from larger amplitude in the 3-back condition than in the 1-back condition (Figure 3). Similar to P2 and N2, there was also a significant main effect of electrode ($F(2,76) = 114.98$, $p < 0.001$, $f = 1.741$), resulting from centro-posterior distribution. The main effect of group was not significant ($F(1,38) = 0.33$, $p = 0.570$, $f = 0.010$), neither the main effect of session ($F(1,38) = 0.67$, $p = 0.417$, $f = 0.132$) nor the session × group interaction ($F(1,38) = 1.73$, $p = 0.196$, $f = 0.215$). However, the session × load × group interaction was significant ($F(1,38) = 11.21$, $p = 0.002$, $f = 0.636$). This raises from the different effect of session on 1-back and 3-back tasks in the training group, which was not observed in the control group. The two other interactions were not significant (session × electrode × group ($F(2,76) = 1.49$, $p = 0.234$, $f = 0.199$; session × load × electrode × group ($F(2,76) = 2.77$, $p = 0.090$, $f = 0.270$), supporting the result that training had a different effect depending on the load.

A post-hoc analyses carried out for the mean amplitude of the P3 at central electrodes, revealed a reduction in the load effect from pretest to posttest in the training group compared to the control group ($t(38) = 2.61$, $p = 0.013$, $d = 0.827$). In the pretest, no group difference in the load effect was observed ($t(38) = -1.34$, $p = 0.188$, $d = 0.424$). An additional post-hoc analyses performed on the post-pre effects for the two load levels separately detected that the training effect comes primarily from an increase of the P3 amplitude in the 1-back from pretest to posttest ($t(19) = 2.32$, $p = 0.032$, $d = 0.287$). In the 3-back task itself, the amplitude decrease was not quite significant ($t(19) = -1.84$, $p = 0.081$, $d = 0.250$). In the control group, no significant differences were observed for the 1-back task, either ($t(19) = -1.92$, $p = 0.069$, $d = 0.259$) or 3-back ($t(19) = -0.27$, $p = 0.790$, $d = 0.036$).

NSW (500 – 1000 ms)

For NSW, the main effect of electrode was significant ($F(2,76) = 17.25$, $p < 0.001$, $f = 0.673$), which comes from the widespread scalp distribution over the fronto-central-parietal scalp areas (Figure 3). There was also a significant main effect of load ($F(1,38) = 106.03$, $p < 0.001$, $f = 1.670$), caused by the response being larger in the 1-back condition as compared to the 3-back condition. The main effect of group was not significant ($F(1,38) = 2.20$, $p = 0.146$, $f = 0.241$). The main effect of session ($F(1,38) = 1.76$, $p = 0.192$, $f = 0.215$) as well as the session × group interaction ($F(1,38) = 0.37$, $p = 0.548$, $f = 0.101$) were not significant. However, again the session × load × group interaction was significant ($F(1,38) = 9.46$, $p = 0.004$, $f = 0.498$) suggesting a training effect when the task load was accounted for. The other interactions (session × electrode × group ($F(2,76) = 1.65$, $p = 0.203$, $f = 0.209$; session × load × electrode × group ($F(2,76) = 0.25$, $p = 0.695$, $f = 0.078$) were not significant, which supports the session × load × group interaction is due to differential training effects in the two groups.

Post-hoc analyses carried out for the mean amplitudes of the NSW at central electrodes, encountered different load effects (1-back *minus* 3-back) between groups at the posttest ($t(38) = 2.94$, $p = 0.006$, $d = 0.142$) that were not observed in the pretest ($t(38) = -0.15$, $p = 0.886$, $d = 0.928$). The second post-hoc analysis conducted for the training group revealed that the training-related load effect change mainly results from reduction of the NSW amplitude in the 1-back task from pretest to posttest ($t(19) = 3.22$, $p = 0.005$, $d = 0.456$). In the 3-back task, training group showed no difference in the ERP amplitude from pretest to posttest ($t(19) = -1.16$, $p = 0.261$, $d = 0.151$). Moreover, no significant differences between the NSW amplitudes from pretest to posttest were detected in the

485 control group either for the 1-back ($t(19) = 0.41$, $p = 0.687$, $d = 0.058$) or for the 3-back
486 ($t(19) = 0.46$, $p = 0.650$, $d = 0.057$) task.
487
488
489 **Discussion**
490
491 We examined how WM updating training modulates a cascade of event-related
492 potentials (ERPs) elicited at different processing stages. Based on fMRI studies (Salmi
493 et al. 2018), we expected that WM updating training would result in relative decreases
494 in the early responses reflecting stimulus selection and response preparation, and
495 relative increases in the late slow wave responses reflecting maintenance of to-be-
496 remembered materials, when responses to difficult and easy tasks are compared. We
497 also wanted to clarify how this pattern suggested by an fMRI meta-analysis relates
498 to ERP effects. We observed behavioral improvements only in the trained task. In
499 general, these limited behavioral findings are in accordance with the most recent meta-
500 analyses in the domain of WM training (Melby-Lervåg et al. 2016, Soveri et al. 2017).
501 That is, accumulating evidence suggests that training-related transfer is mostly observed
502 in the untrained variants of the trained tasks. As we expected based on brain imaging
503 studies (Salmi et al. 2018), despite the modest behavioral effects we observed consistent
504 ERP effects at multiple latencies. The load effect in the early responses taken to reflect
505 attentional modulation of sensory-motor processing was decreased in the training group,
506 probably due to a difficult task becoming partly automatized during the training period.
507 NSW elicited during the maintenance period, in turn, showed a decrease in the easy
508 task. Our findings, suggesting that early and late ERPs are differentially affected by

training, provide important evidence of the neural mechanisms associated with WM training.

Training-related modulation of the P2-N2-P3 complex: In our study, training-induced changes were observed already 200-300 ms after the stimulus onset. Such modulations of early responses, especially N2 and P3 responses, have been observed also in prior studies using differential experimental approaches either in training or in measuring its effects (Shiran and Brezniz 2011, Oelhafen et al. 2013, Liu et al. 2017, Covey et al. 2018, Pergher et al. 2018, Covey et al. 2019). We did observe a significant session × group interaction already at P2 latency, although when the groups were examined separately a pre-post comparison showed an effect only in the control group. As Bejeweled game is also demanding, requiring a lot of attention, visual discrimination, enhanced processing speed, and to some extent even WM, it is possible that this modulation of the early P2 response does reflect changes in some cognitive process. However, due to the non-experimental nature of this task, the related effects are difficult to interpret. At the P2 latency, the differential effect of the intervention on the two groups was not affected by the task load. Training effects at N2 and P3 latencies, in turn, were different for the easy 1-back and difficult 3-back tasks, also showing a relative decrease in the difference between the two load levels. While the relative decrease in the load effect might relate to automatization, as fMRI studies have suggested (Salmi et al. 2018), direct comparisons of the pre-post effects revealed a more complex pattern. In N2, the training group showed an increase in the 3-back task from pretest to posttest and no effect in the 1-back task, while the control group showed no effect of intervention in the 3-back task but an increase in the 1-back task. An increase in N2 in the 3-back task was also reported by Covey et al. 2018, both in healthy

participants and in patients with multiple sclerosis, but their analysis focused on the difficult condition (see also Covey et al. 2019). While Covey et al. did not report a significant modulation of P3 by training, possibly because of a relatively modest sample size, our findings suggested a different pattern at these latencies than in N2 latencies, if the load effect is not considered. More specifically, we observed a training-related increase in the 1-back and a trend towards reduced response at posttest in the 3-back task. Although our design was not optimal for distinguishing different psychological phenomena potentially affected by training at these latencies, we provide important evidence that different stages in the cascade of ERP responses are uniquely affected by training. Nevertheless, due to the varying task designs and ERP indicators, as well as a lack of reliable links between brain activity and task performance, more evidence regarding the factors underlying activation increases and decreases is clearly needed.

As in previous studies (Oelhafen et al. 2013, Pergher et al. 2018), we observed P3 amplitude increase in the WM training group. However, while Oelhafen et al. (2013) and Pergher et al. (2018) reported this effect also in a difficult WM task, we observed P3 increase only in the easier 1-back task. Nevertheless, we would like to note that direct comparison between our results and those of previous studies should not be made because the analyses were not similar (e.g., we conducted direct comparisons between pre and post targets, while Pergher et al. 2018 analyzed training effects for target vs. non-target comparisons), the focus of training was different (e.g., Oelhafen et al. 2013 targeted interference effects), the training paradigms (adaptive or not) varied, and the ensuing learning curves were different in the training groups.

The role of maintenance in neuronal mechanisms of WM training: Our study tested the hypothesis that WM training would lead to more effective maintenance of the to-be-remembered stimuli, as reflected by changes in slow waves occurring during active maintenance of WM contents (Bailey et al. 2016, Vilà-Balló et al. 2018). In accordance to our expectations, NSW was indeed affected by training. Both behavioral studies (Ecker et al. 2010, Ecker et al. 2013, Botto et al. 2014) as well as neurophysiological recordings (Bailey et al. 2016, Vilà-Balló et al. 2018) have suggested that there are more resources available for the active maintenance of relevant information when the executive demands are low. By examining the slow waves elicited in-between stimuli, we were able to probe how training influences maintenance of WM information. It should be noted that another late slow response, namely the CNV, would be expected to be increase in relation to response anticipation (Walter et al. 1964) that can be improved mostly in the easier 1-back task. This supports our expectations that our late slow waves were not explained by anticipatory responses.

Accumulating evidence suggests that WM training modulates activity in the fronto-parieto-striatal networks (Salmi et al. 2018). Decreased task-related brain activity after practice is likely to reflect more efficient neuronal processing due to automatization of particular cognitive processes (Constantinidis and Klingberg 2016, see also Chein and Schneider 2005). In our study, decreased slow wave in the low-load condition may reflect change from controlled processing to a partly automatized, procedural processing mode. Reduced slow wave amplitudes could also reflect a redistribution of neuronal resources. While direct evidence from EEG studies is still lacking, existing interpretation is mostly based in data accumulated across various human fMRI studies (see Salmi et al. 2018). Decrease in task-related brain activity following WM training

583 has been systematically reported in the occipitoparietal areas (Salmi et al. 2018). There
584 is also evidence of the enhanced selection of information after WM training (Kundu et
585 al. 2013). Similarly to Kundu et al. (2013), we observed decreased brain responses in
586 the posterior electrode sites. However, there is also another neuronal mechanism for
587 automatization that is repeatedly reported in WM training studies, namely the decrease
588 in the dorsolateral prefrontal activity (Dahlin et al. 2008, for a review see Bäckman and
589 Nyberg 2013). Due to the limited spatial resolution of our study, we cannot reliably
590 specify the source location of the slow wave in the 1-back condition. In addition to
591 partial automatization of WM processing components (see von Bastian and Oberauer
592 2014), decreased brain activity could reflect better exploitation of individual capacity
593 via the development of task-specific strategies (Dunning and Holmes 2013, De Simoni
594 and von Bastian 2018). Such strategies that might decrease brain activity (cf. Klingberg
595 2010) start to develop already after very short practice (Laine et al. 2018).
596
597 Training-related activation increases, taking place mainly in the frontal eye fields,
598 supplementary motor cortex and ventral prefrontal cortex, have been reported as
599 systematically as activation decreases (Salmi et al. 2018). Although it has been
600 suggested that activation increases could reflect enhanced capacity to utilize attentional
601 resources in the trained tasks (Olesen et al. 2004, Klingberg 2010), direct evidence of
602 the functional role of these activation increases resulting from WM updating training
603 has been scarce. Training-related modulations of slow wave activity also corresponds to
604 studies in non-human primates where in the course of practice, the amount of neurons
605 activated during the maintenance period increases (Qi et al. 2011). While the increased
606 ERPs could potentially also reflect more focused neural sources, we can only rely on
607 fMRI studies by noting that this is unlikely to be the case (Salmi et al. 2018). Although

608 our findings highlight a specific neuronal mechanism that is affected by training, two
609 alternative theories explaining how the learning occurs at the behavioral level remain. It
610 could either be that enhanced strategy use results to increased activity in the
611 maintenance phase in a demanding task (Cole et al. 2010), or that the increased brain
612 responses reflect enhanced ability to allocate attention in a task that is structurally
613 similar to the trained one (Klingberg 2010).
614
615 ***Limitations of the study:*** While our findings related to the behavioral transfer are
616 consistent with numerous other studies (see Soveri et al. 2017), the lack of systematic
617 behavioral transfer effects also restrict the interpretation of the present findings. It
618 should also be noted that single n-back tasks were conducted only during the EEG
619 recording, which may have contributed to weak near transfer effects (Bäckman et al.
620 2017). Alternatively, also the differential inter-stimulus-intervals in the dual and single
621 n-back tasks may have influenced the training effects. Nevertheless, due to the high
622 similarity between the trained task and the single n-back task, our findings are likely to
623 reflect learning related to the trained task, rather than some general capacity change.
624 Indeed, training-related improvements in the n-back tasks could be largely explained by
625 adoption of task-specific strategies (Laine et al. 2018). It should also be noted that the
626 sample size and other sample-specific features could affect the generalizability of the
627 findings. Regarding our experimental design, accuracy in the 1-back task was relatively
628 high already prior to training, leaving limited room for improvement. Finally, more
629 evidence of the functional roles of the slow waves associated with WM and their
630 responsiveness to training is clearly needed. High-resolution MEEG (combined MEG
631 and EEG) or combined EEG and fMRI might help in detailing the sources of the slow
632 wave components and in further clarifying the functional roles of these components.

633 Based on the current findings it remains partially unclear, for instance, to what extent
634 the observed slow wave activity reflects the same underlying neural functions and how
635 these functions are associated with other components, such as CNV, that are also
636 modulated by the WM load. Despite these limitations and limited transfer, there is still
637 hope that in the long run WM training results to some potential applications, as
638 especially some of the clinical studies have provided promising initial results (Owens et
639 al. 2013, Cortese et al. 2015, Motter et al. 2015, Saunders et al. 2015, Leung et al. 2016,
640 Jones et al. 2018). Moreover, it is possible that some of the EEG effects are not directly
641 reflected to task-related responses, but only observed during resting state (e.g., Sari et
642 al. 2016).
643
644 ***Conclusions:*** Despite the extensive research on WM training, its underlying
645 mechanisms have remained unclear. We provide evidence that ERP responses at
646 different latencies and stages of WM processing are differentially affected by training.
647 Our findings provide new insights to the role of task load in the training-related
648 increases and decreases in brain responses (see Salmi et al. 2018 for a meta-analysis).
649 Early responses were affected by practice both in the training group and in the control
650 group, but the modulations were different in the two groups. The precise temporal
651 resolution of EEG and a recently identified marker for an important WM component,
652 active maintenance, enabled us to study WM training effects on this component for
653 which there has been evidence only in non-human primates (Qi et al. 2011, Meyers et
654 al. 2012). The reported training-related changes in a cascade of brain responses shed
655 light on human brain plasticity following prolonged practice with cognitive tasks.
656
657 **Acknowledgements**

The study was supported by funds from the Åbo Akademi University Endowment (BrainTrain project) and the Academy of Finland (grant #260276) awarded to ML. We are grateful to E. Serrat-Sellabona, the Head of the Psychology Department, University of Girona, for the institutional support during the data collection. We also would like to thank D. Cucurell from the Brain Plasticity Group, and D. Espinosa from the University of Girona for their technical advice. Special thanks to all members of BrainTrain research group for their valuable comments.

References

Azizian, A., Freitas, A.L., Parvaz, M.A., Squires, N.K., 2006. Beware misleading cues: perceptual similarity modulates the N2/P3 complex. Psychophysiology 43, 253-60.

Bailey, K., Mlynarczyk, G., West, R., 2016. Slow wave activity related to working memory maintenance in the N-Back task. J. Psychophysiol., 30, 141-154.

Bäckman, L., Nyberg, L., 2013. Dopamine and training-related working-memory improvement. Neurosci. Biobehav. Rev. 37, 2209-2219.

Bäckman, L., Waris, O., Johansson, J., Andersson, M., Rinne, J.O., Alakurtti, K., Soveri, A., Laine, M., Nyberg, L., 2018. Increased dopamine release after working-memory updating training: neurochemical correlates of transfer. Sci. Rep. 7, 7160-017-07577-y.

Baddeley, A.D., 1986. *Working memory.* Oxford: Oxford University Press.

Barriga-Paulino, C.I., Rodríguez-Martínez, E.I., Rojas-Benjumea, M.Á., Gómez, C.M., 2014. Slow wave maturation on a visual working memory task. Brain Cogn. 88, 43-54.

Botto, M., Basso, D., Ferrari, M., Palladino, P., 2014. When working memory updating requires updating: Analysis of serial position in a running memory task. Acta Psychol. (Amsterdam), 148, 123–129.

Buschkuehl, M., Jaeggi, S.M., Jonides, J., 2012. Neuronal effects following working memory training. Dev. Cogn. Neurosci. 2 (Suppl. 1), S167–179.

Bunting, M., Cowan, N., Saults, J.S., 2006. How does running memory span work?. Q. J. Exp. Psychol. (Hove) 59, 1691-1700.

Carretti, B., Cornoldi, C., Pelegrina, S.L., 2007. Which factors influence number updating in working memory? The effects of size distance and suppression. Br J Psychol. 98, 45-60.

Chein, J.M., Schneider, W., 2005. Neuroimaging studies of practice-related change: fMRI and meta-analytic evidence of a domain-general control network for learning. Brain Res. Cogn. Brain Res. 25, 607-623.

Chein, J.M., Schneider, W., 2012. The brain's learning and control architecture. Curr. Dir. Psychol. Sci. 21, 78–84.

Cohen, J., 1992. A power primer. Psychol. Bull., 112, 155-159.

Cohen, M., 2014. Analyzing neural time series data. Cambridge, Massachusetts: MIT Press.

700 Constantinidis, C., Klingberg, T., 2016. The neuroscience of working memory capacity
701 and training. Nat. Rev. Neurosci. 17, 438-449.

702 Cole, M. W., Bagic, A., Kass, R., Schneider, W., 2010. Prefrontal dynamics underlying
703 rapid instructed task learning reverse with practice. J. Neurosci., 30, 14245–14254.

704 Crowley, K. E., Colrain, I. M., 2004. A review of the evidence for P2 being an
705 independent component process: Age, sleep and modality. Clin. Neurophysiol., 115,
706 732–744.

707 Covey, T.J., Shucard, J.L., Benedict, R.H., Weinstock-Guttman, B., Shucard, D.W.,
708 2018. Improved cognitive performance and event-related potential changes following
709 working memory training in patients with multiple sclerosis. Mult. Scler. J. Exp. Transl.
710 Clin. 11:2055217317747626.

711 Covey, T. J., Shucard, J. L., & Shucard, D. W. (2019). Working memory training and
712 perceptual discrimination training impact overlapping and distinct neurocognitive
713 processes: Evidence from event-related potentials and transfer of training gains.
714 Cognition, 182, 50-72.

715 Gómez, C.M., Barriga-Paulino, C.I., Rodríguez-Martínez, E.I., Rojas-Benjumea, M.Á.,
716 Arjona, A., Gómez-González, J., 2018. The neurophysiology of working memory
717 development: from childhood to adolescence and young adulthood. Rev. Neurosci. 28,
718 261-282.

719 Daffner, K. R., Chong, H., Sun, X., Tarbi, E. C., Riis, J. L., McGinnis, S. M., Holcomb,
720 P.J., 2011. Mechanisms underlying age- and performance-related differences in working
721 memory. J. Cogn. Neurosci. 23, 1298–1314.

722 Dahlin, E., Neely, A.S., Larsson, A., Bäckman, L., Nyberg, L., 2008. Transfer of
723 learning after updating training mediated by the striatum. Science 320, 1510-1512.

724 De Simoni, C., von Bastian, C.C., 2018. Working memory updating and binding
725 training: Bayesian evidence supporting the absence of transfer. J. Exp. Psychol. Gen.
726 147, 829-858.

727 Dien, J., Spencer, K. M., Donchin, E., 2004. Parsing the late positive complex: Mental
728 chronometry and the ERP components that inhabit the neighborhood of the P300.
729 Psychophysiology 41, 665–678.

730 Donchin, E., Miller, G. A., Farwell, L. A., 1986. The endogenous components of the
731 event-related potential–A diagnostic tool? Prog. Brain Res. 70, 87–102.

732 Dong, S., Reder, L. M., Yao, Y., Liu, Y., Chen, F., 2015. Individual differences in
733 working memory capacity are reflected in different ERP and EEG patterns to task
734 difficulty. Brain Res. 1616, 146–156.

735 Dunning, D.L., Holmes, J., Gathercole, S.E., 2013. Does working memory training lead
736 to generalized improvements in children with lowworking memory? A randomized
737 controlled trial. Dev. Sci. 216, 915-25.

738 Ecker, U. K., Lewandowsky, S., Oberauer, K., Chee, A. E., 2010. The components of
739 working memory updating: An experimental decomposition and individual differences.
740 J. Exp. Psychol: Learn., 36, 170–189.

741 Ecker, U. K., Maybery, M., Zimmer, H. D., 2013. Binding of intrinsic and extrinsic
742 features in working memory. J. Exp. Psychol.: Gen., 142, 218–234.

743 Folstein, J. R., Van Petten, C., 2008. Influence of cognitive control and mismatch on the
744 N2 component of the ERP: a review. Psychophysiology 45, 152–170.

745 Ford, J. M., Roth, W. T., Mohs, R. C., Hopkins, W. F., Kopell, B. S., 1979. Event-
746 related potentials recorded from young and old adults during a memory retrieval task.
747 Electroen. Clin. Neuro., 47, 450–459.

748 Fukuda, K., Awh, E., Vogel, E. K., 2010. Discrete capacity limits in visual working
749 memory. Curr. Opin. Neurobiol., 20, 177–182.

750 Fuster, J. M., 2000. Prefrontal neurons in networks of executive memory. Brain Res.
751 Bull., 52, 331–336.

752 Gevins, A., Smith, M.E., Le, J., Leong, H., Bennett, J., Martin, N., McEvoy, L., Du, R.,
753 Whitfield, S., 1996. High resolution evoked potential imaging of the cortical dynamics
754 of human working memory. Electroencephalogr. Clin. Neurophysiol. 98, 327-348.

755 Jennings, J.R., Wood, C.C., 1976. Letter: The epsilon-adjustment procedure for
756 repeated-measures analyses of variance. Psychophysiology, 13, 277–278.

757 Jones, M.R., Katz, B., Buschkuehl, M., Jaeggi, S.M., Shah, P., 2018. Exploring n-back
758 cognitive training for children with ADHD. J. Atten. Disord., 1087054718779230.

759 Kelly, C., Foxe, J.J., Garavan, H., 2006. Patterns of normal human brain plasticity after
760 practice and their implications for neurorehabilitation. Arch. Phys. Med. Rehabil. 87,
761 S20–S29.

762 Klingberg, T., 2010. Training and plasticity of working memory. Trends Cogn. Sci. 14,
763 317-324.

764 Kundu, B., Sutterer, D.W., Emrich, S.M., Postle, B.R., 2013. Strengthened effective
765 connectivity underlies transfer of working memory training to tests ofshort-term
766 memory and attention. J. Neurosci. 33, 8705-15.

767 Laine, M., Fellman, D., Waris, O., Nyman, T.J., 2018. The early effects of external and
768 internal strategies on working memory updating training. Sci. Rep. 8, 4045.

769 Langer, N., von Bastian, C.C., Wirz, H., Oberauer, K., Jäncke, L., 2013. The effects
770 of working memory training on functional brain network efficiency. Cortex. 49, 2424-
771 38.

772 Lefebvre, C. D., Marchand, Y., Eskes, G. A., Connolly, J. F., 2005. Assessment of
773 working memory abilities using an event-related brain potential (ERP) – Compatible
774 digit span backward task. Clin. Neurophysiol. 116, 1665–1680.

775 Lenartowicz, A., Escobedo-Quiroz, R., Cohen, J. D., 2010. Updating of context in
776 working memory: An event-related potential study. Cogn. Affect. Behav. Neurosci. 10,
777 298–315.

778 Leung, I.H., Walton, C.C., Hallock, H., Lewis, S.J., Valenzuela, M., Lampit, A., 2015.
779 Cognitive training in Parkinson disease: A systematic review and meta-analysis.
780 Neurology 85, 1843-51.

781 Levy, R., Goldman-Rakic, P.S., 2000. Segregation of working memory functions within
782 the dorsolateral prefrontal cortex. Exp. Brain Res. 133, 23-32.

783 Liu, Z.X., Lishak, V., Tannock, R., Woltering, S., 2017. Effects of working memory
784 training on neural correlates of Go/Nogo response control in adults with ADHD: A
785 randomized controlled trial. Neuropsychologia, 95, 54-72.

786 Luck, S., 2014. An introduction to the event-related potential technique. Cambridge,
787 Massachusetts: MIT Press.

788 Luria, R., Balaban, H., Awh, E., Vogel, E.K., 2016. The contralateral delay activity as a
789 neural measure of visual working memory. Neurosci. Biobehav. Rev. 62, 100-8.

790 Luu, P., Caggiano, D. M., Geyer, A., Lewis, J., Cohn, J., Tucker, D. M., 2014.
791 Timecourse of cortical networks involved in working memory. Front. Hum. Neurosci.
792 8, 4.

793 Marchand, Y., Lefebvre, C. D., Connolly, J. F., 2006. Correlating digit span
794 performance and event-related potentials to assess working memory. Int. J.
795 Psychophysiol., 62, 280–289.

796 Marco-Pallares, J., Cucurell, D., Münte, T. F., Strien, N. Rodriguez-Fornells, A., 2011.
797 On the number of trials needed for a stable feedback-related negativity.
798 Psychophysiology 48, 852–860.

799 McEvoy, L. K., Smith, M. E., Gevins, A., 1998. Dynamic cortical networks of verbal
800 and spatial working memory: Effects of memory load and task practice. Cereb. Cortex,
801 8, 563–574.

802 Mecklinger, A., Muller, N., 1996. Dissociations in the processing of what and where
803 information in working memory: An event-related potential analysis. J. Cog. Neurosci.,
804 8, 453–473.

805 Mecklinger, A., Pfeifer, E., 1996. Event-related potentials reveal topographical and
806 temporal distinct neuronal activation patterns for spatial and object working memory.
807 Brain Res. Cogn Brain Res., 4, 211–224.

808 Melby-Lervag, M., Hulme, C., 2013. Is working memory training effective? A meta-
809 analytic review. Devel. Psychol. 49, 270-291.

810 Melby-Lervåg, M., Redick, T.S., Hulme, C., 2016. Working memory training does not
811 improve performance on measures of intelligence or other measures of "far transfer":
812 evidence from a meta-analytic review. Perspect. Psychol. Sci. 11, 512–534.

813 Meyers, E.M., Qi, X.L., Constantinidis, C., 2012. Incorporation of new information into
814 prefrontal cortical activity after learning working memory tasks. Proc Natl Acad Sci
815 (USA) 109, 4651-4656.

816 Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., Wager, T.
817 D., 2000. The unity and diversity of executive functions and their contributions to
818 complex Frontal Lobe tasks: A latent variable analysis. Cognitive Psychol., 41, 49–100.

819 Morís, J., Luque, D., & Rodríguez-Fornells, A., 2013. Learning-induced modulations of
820 the stimulus-preceding negativity. Psychophysiology 50(9), 931-939.

821 Morris, N., Jones, D. M., 1990. Memory updating in working memory: The role of the
822 central executive. Brit. J. Psychol. 81, 111–121.

823 Motter, J.N., Pimontel, M.A., Rindskopf, D., Devanand, D.P., Doraiswamy, P.M.,
824 Sneed, J.R., 2016. Computerized cognitive training and functional recovery in major
825 depressive disorder: A meta-analysis. J. Affect. Disord. 189, 184-91.

826 Olesen, P.J., Westerberg, H., Klingberg, T., 2004. Increased prefrontal and parietal
827 activity after training of working memory. Nat. Neurosci. 7, 75-79.

828 Oelhafen, S., Nikolaidis, A., Padovani, T., Blaser, D., Koenig, T., Perrig, W.J., 2013.
829 Increased parietal activity after training of interference control. Neuropsychologia 51,
830 2781-2790.

831 Owen, A.M., McMillan, K.M., Laird, A.R., Bullmore, E., 2005. N-back working
832 memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum
833 Brain Mapp. 25, 46-59.

834 Owens, M,, Koster, E.H., Derakshan, N., 2013. Improving attention control in
835 dysphoria through cognitive training: Transfer effects on working memory capacity and
836 filtering efficiency. Psychophysiology 50, 297-307.

837 Pergher, V., Wittevrongel, B., Tournoy, J., Schoenmakers, B., Van Hulle, M.M. 2018.
838 N-back training and transfer effects revealed by behavioral responses and EEG. Brain
839 Behav. 8(11), e01136.

840 Postle, B.R., Druzgal, T.J., D'Esposito, M., 2003. Seeking the neural substrates of visual
841 working memory storage. Cortex 39, 927-946.

842 Qi, X.L., Meyers, T., Stanford, T.R., Constantinidis, C., 2011. Changes in prefrontal
843 neuronal activity after learning to perform a spatial working memory task. Cereb.
844 Cortex. 21, 2722-2732.

845 Rämä, P., Paavilainen, L., Anourova, I., Alho, K., Reinikainen, K., Sipilä, S., Carlson,
846 S., 2000. Modulation of slow brain potentials by working memory load in spatial and
847 nonspatial auditory tasks. Neuropsychologia, 38, 913–922.

Roth, W. T., Kopell, B. S., Tinklenberg, J. R., Darley, C. F., Sikoda, R., Vesecky, T. B., 1975. The contingent negative variation during a memory retrieval task. Electroencephalograph Clin. Neurophysiol.., 38, 171–174.

Roth, W. T., Rothbart, R. M., Kopell, B. S., 1978. The timing of CNV resolution in a memory retrieval task. Biol. Psychol. 6, 39–49.

Ruchkin, D. S., Johnson, R., Jr., Grafman, J., Canoune, H., Ritter, W., 1992. Distinctions and similarities among working memory processes: An event-related potential study. Brain Res. Cogn. Brain Res., 1, 53–66.

Ruchkin, D. S., Canoune, H. L., Johnson, R., Jr., Ritter, W., 1995. Working memory and preparation elicit different patterns of slow wave event-related brain potentials. Psychophysiology, 32, 399–410.

Salmi. J., Nyberg, L., Laine, M., 2018. Working memory training mostly engages general-purpose large-scale networks for learning. Neurosci. Biobehav. Rev. 93, 108-122

Sari, B.A., Koster, E.H., Pourtois, G., Derakshan, N., 2016. Training working memory to improve attentional control in anxiety: A proof-of-principle study using behavioral and electrophysiological measures. Biol. Psychol., 121, 203-212.

Saunders, N., Downham, R., Turman, B., Kropotov, J., Clark, R., Yumash, R., Szatmary, A., 2014. Working memory training with tDCS improves behavioral and neurophysiological symptoms in pilot group with post-traumatic stress disorder (PTSD) and with poor working memory. Neurocase 21, 271-278..

869 Shiran, A., Breznitz, Z., 2011. The effect of cognitive training on recall range and speed
870 of information processing in the working memory of dyslexic and skilled readers. J
871 Neuroling. 24, 524-537.

872 Soveri, A., Antfolk, J., Karlsson, L., Salo, B., Laine, M., 2017a. Working memory
873 training revisited: A multi-level meta-analysis of n-back training studies. Psychon. Bull.
874 Rev. 24, 1077–1096.

875 Soveri, A., Karlsson, E.P.A., Waris, O., Grönholm-Nyman, P., Laine, M., 2017. Pattern
876 of near transfer effects following working memory training with a dual n-back task.
877 Exp. Psychol. 64, 240-252.

878 Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., Winter, A. L., 1964.
879 Contingent negative variation: An electric sign of sensori-motor association and
880 expectancy in the human brain. Nature, 203, 380–384.

881 Waris, O., Soveri, A., Laine, M., 2015. Transfer after Working Memory Updating
882 Training. PLoS One. 25, e0138734.

883 Vilà-Balló, A., François, C., Cucurell, D., Miró, J., Falip, M., Juncadella, M., &
884 Rodríguez-Fornells, A., 2017. Auditory target and novelty processing in patients with
885 unilateral hippocampal sclerosis: A current-source density study. Sci Rep., 7(1), 1612.

886 Vilà-Balló, A., Salmi, J., Soveri, A., Rodríguez-Fornells, A., Lehtonen, M., Laine, M.,
887 2018. Neural signatures for active maintenance and interference during working
888 memory updating. Biol. Psychol. 132, 233-243.

889 von Bastian, C.C., Oberauer, K., 2014. Effects and mechanisms of working memory
890 training: a review. Psychol. Res. 78, 803-820.

Table 1. Demographic data on the participants. Numbers of females and males, and means and standard deviations (in parenthesis) for other participant characteristics. See SOM for details of the scales and assessment.

Measure	**Training**	**Control**	***p***
Sex F/M	15/5	19/1	
Age (years)	22.00 (3.31)	21.80 (2.67)	0.837
Education (years)	15.65 (1.53)	16.20 (1.58)	0.270
WAIS Similarities	21.55 (3.85)	19.75 (3.37)	0.124
BDI-II	4.40 (3.72)	3.35 (3.98)	0.394
Motivation			
First meeting	7.90 (1.37)	7.90 (1.48)	1.000
Pretest	8.15 (1.27)	7.85 (1.57)	0.509
During training	8.10 (1.21)	7.95 (1.36)	0.714
After training	8.60 (1.14)	8.55 (1.23)	0.895
Posttest	8.20 (1.06)	8.20 (1.15)	1.000

Table 2. Means and standard deviations (in parenthesis) in each cognitive performance measure included in the analysis of behavioral data.

Measure	**n (training, control)**	**Variable**	**Training pretest**	**Control pretest**	**Training posttest**	**Control posttest**	***p* (group x session)**
Dual n-back	20+20	Max level	2.45 (0.510)	2.50 (0.607)	5.40 (0.821)	2.85 (0.671)	0.001
Single 1-back	20+20	Accuracy	94.30 (3.30)	93.17 (4.03)	93.95 (3.12)	91.62 (5.15)	0.342
Single 3-back	20+20	Accuracy	81.09 (6.76)	78.64 (12.30)	89.56 (6.16)	84.69 (9.70)	0.226
WM Updating	17+19	Z-score	-0.12 (1.39)	0.23 (1.70)	0.29 (1.50)	-0.15 (1.47)	0.107
WM Interference	17+18	Z-score	0.35 (1.79)	0.04 (1.64)	0.86 (1.43)	-0.43 (1.86)	0.177
Passive WM	19+19	Z-score	-0.30 (1.51)	0.18 (1.59)	0.00 (1.58)	0.10 (1.86)	0.445
Active WM	17+18	Z-score	-0.18 (1.65)	0.18 (1.87)	0.19 (1.26)	-0.01 (1.67)	0.221

901

902 **Figure Legends**

903

904 **Figure 1.** The n-back task. **A.** Schematic example of the first part of a 1-back sequence
905 where target, standard non-target, and n+1 lure non-target trials are shown. **B.**
906 Schematic example of the first part of a 3-back sequence where target, standard non-
907 target, n-1 lure non-target, and n+1 lure non-target trials are shown. In each sequence,
908 we presented numbers from 1-9 in the middle of a computer screen. The trial began with
909 a fixation point for 450 ms, followed by the number shown for 1500 ms (1950 ms
910 stimulus onset asynchrony). Participants had to press the 'yes' button (target trials)
911 when the number was the same than the previous number (1-back task) or the number
912 presented three numbers before (3-back task). For the other, non-target numbers that
913 included standard, n-1 lure, and n+1 lure non-target trials, the participants had to press
914 the 'no' button.

915

916 **Figure 2.** Mean (±SEM) performance in the dual n-back tasks at the pretest and posttest
917 sessions for the two groups (A). Training progress across 15 training sessions in the
918 experimental group, including the mean (±SEM) n-back level achieved at each training
919 session (B).

920

921 **Figure 3.** Stimulus-locked ERP responses with a baseline from -200 to 0 ms. Grand
922 average ERPs for the 1-back (black lines) and 3-back (red lines) from nine electrode
923 locations for pre (solid lines) and post (dashed lines) sessions, for both the training (A)
924 and the control (B) group. The P2, N2, P3 and NSW components showed an increased
925 positive activity for the 3-back compared with the 1-back. For illustration purposes,
926 these ERPs were low-pass filtered to 8 Hz.

927

928 **Figure 4.** Stimulus-locked ERP responses with a baseline from -200 to 0 ms. Difference
929 waveform ERPs involving 1-back minus 3-back for pre (solid black lines) and post
930 (dashed black lines) sessions, for both the training (A) and the control (B) group. A
931 cascade effect reflecting the reduction of the load effect was observed. Bottom part:
932 scalp distribution of the P2, N2, P3, and NSW (1-back minus 3-back, −3.5/+3.5μV). For
933 illustration purposes, these ERPs were low-pass filtered to 6 Hz.

934

935 **Figure 5.** Stimulus-locked ERP responses with a baseline from -200 to 0 ms. Difference
936 waveform ERPs involving 1-back post minus 1-back pre (solid black lines), and 3-back
937 post minus 3-back pre (solid red lines), for both the training (A) and the control (B)
938 group. A clear modulation of the ERP signal was observed in a large window at the
939 posttest compared to the pretest for the training group and specially for 1-back, which
940 was not observed in the control group. Bottom part: scalp distribution of the P2, N2, P3,
941 and NSW (1-back post minus 1-back pre, and 3-back post minus 3-back pre,
942 −3.0/+3.0μV). For illustration purposes, these averages were low-pass filtered to 6 Hz.

943

944

Figure
Click here to download high resolution image

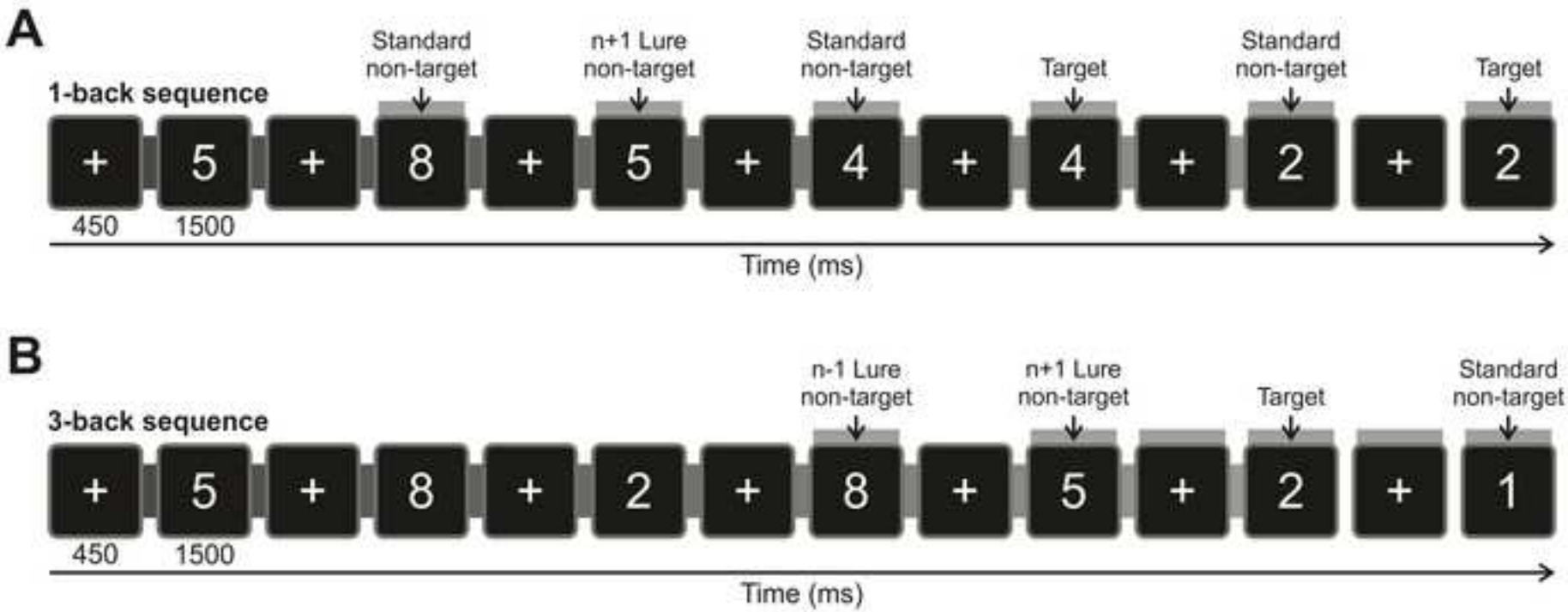
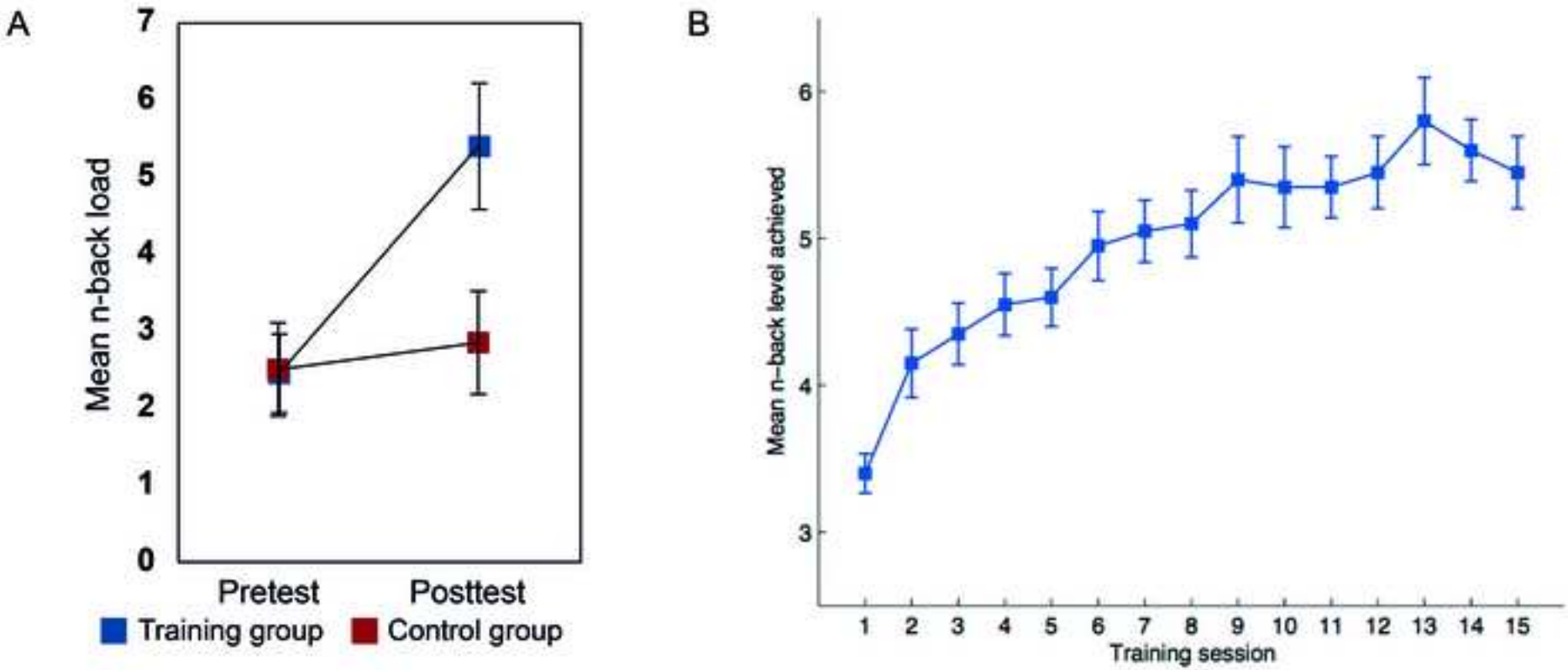
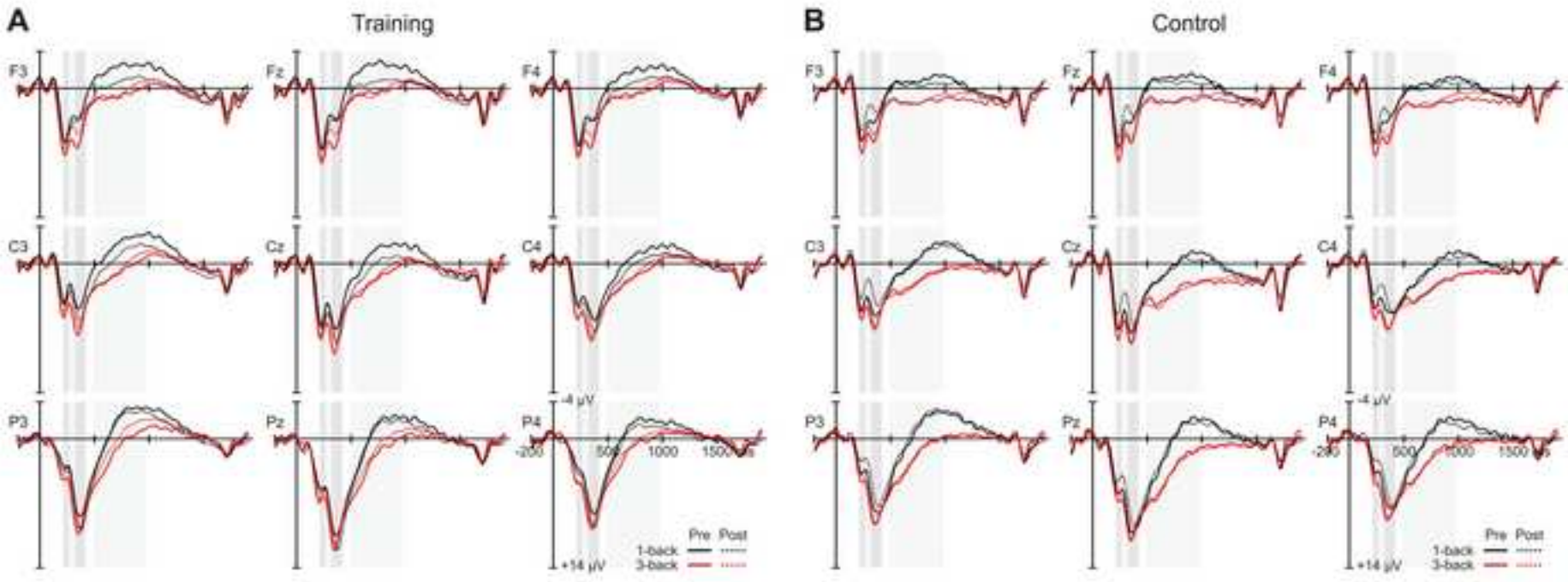


Figure
Click here to download high resolution image



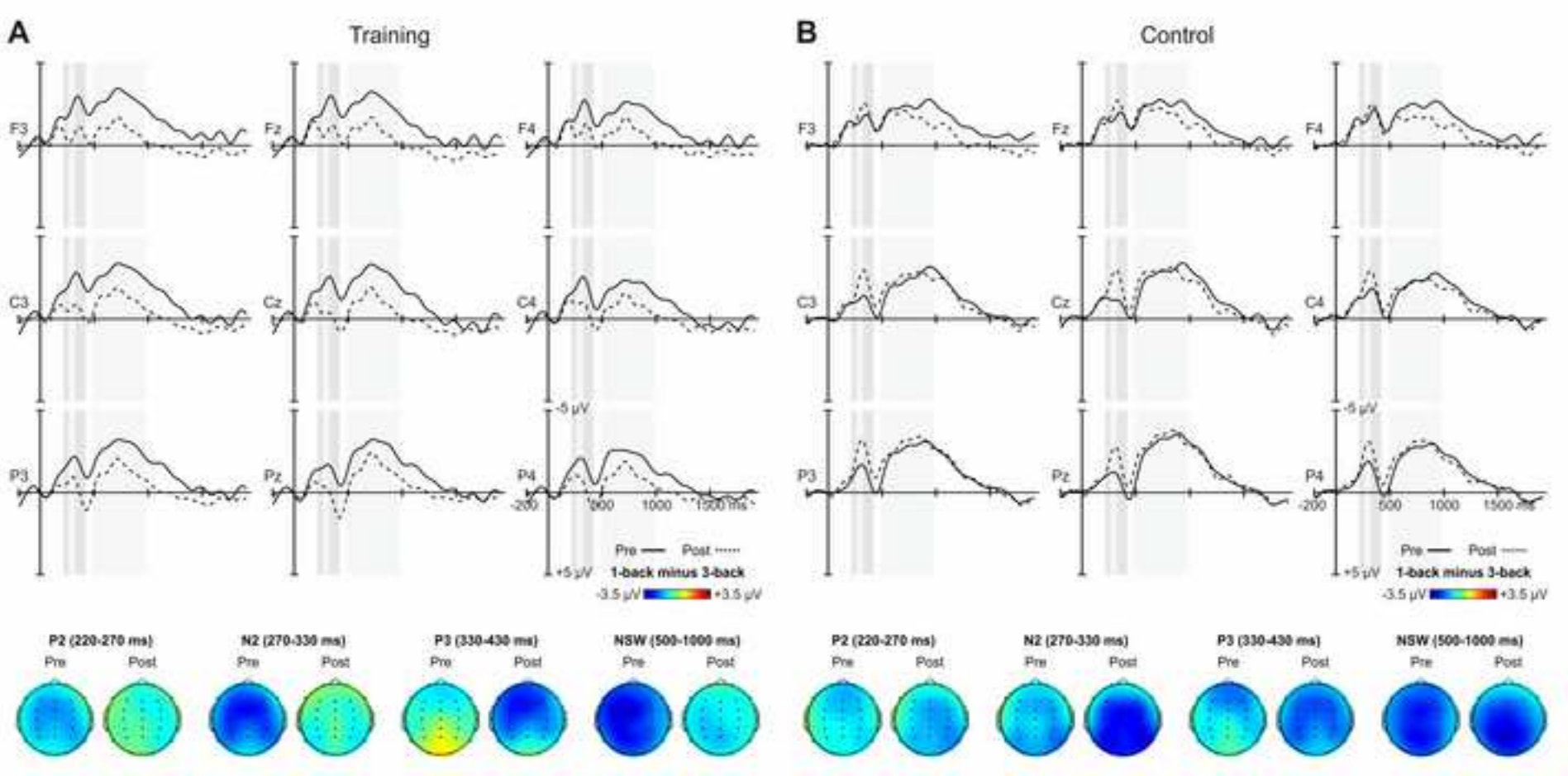
Figure

Click here to download high resolution image



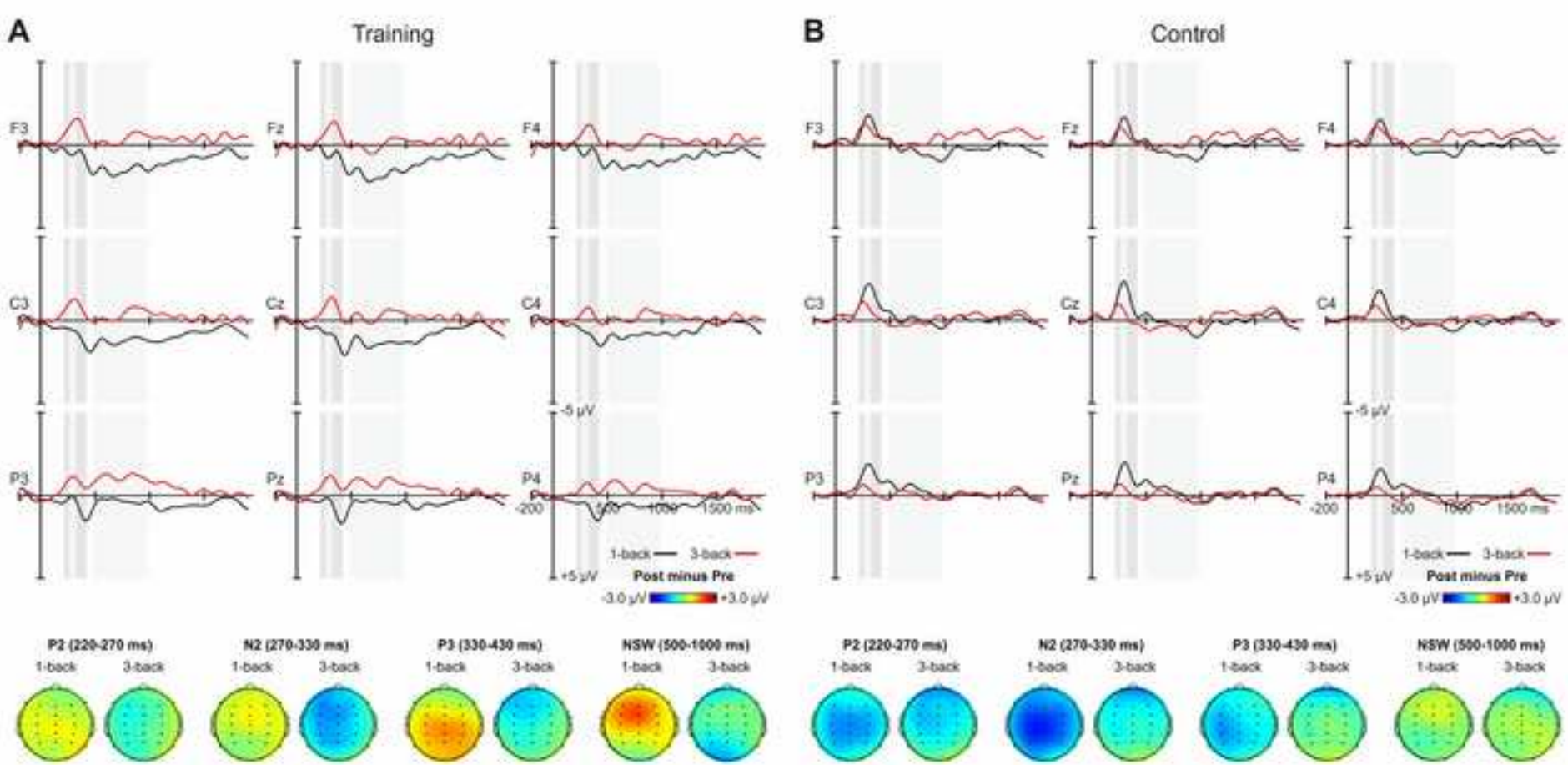
Figure

Click here to download high resolution image



Figure

Click here to download high resolution image



Supplementary Material

Click here to download Supplementary Material: Supplementary_Fig1.jpg

Supplementary Material

Click here to download Supplementary Material: Supplementary_Fig2.jpg

Supplementary Material

Click here to download Supplementary Material: Supplementary_Fig3.jpg

Supplementary Material

Click here to download Supplementary Material: Supplementary_Fig4.pdf