Contents lists available at ScienceDirect

Neuropsychologia

Increased parietal activity after training of interference control

Stephan Oelhafen [a,*], Aki Nikolaidis [a,b], Tullia Padovani [a], Daniela Blaser [a],
Thomas Koenig [c], Walter J. Perrig [a,d]

[a] Institute of Psychology, University of Bern, Switzerland
[b] Neuroscience Program, University of Illinois Urbana Champaign, United States
[c] Department of Psychiatric Neurophysiology, University Hospital of Psychiatry, University of Bern, Switzerland
[d] Center for Cognition, Learning and Memory, University of Bern, Switzerland

ARTICLE INFO

ABSTRACT

Recent studies suggest that computerized cognitive training leads to improved performance in related but untrained tasks (i.e. transfer effects). However, most study designs prevent disentangling which of the task components are necessary for transfer. In the current study, we examined whether training on two variants of the adaptive dual n-back task would affect untrained task performance and the corresponding electrophysiological event-related potentials (ERPs). Forty three healthy young adults were trained for three weeks with a high or low interference training variant of the dual n-back task, or they were assigned to a passive control group. While n-back training with high interference led to partial improvements in the Attention Network Test (ANT), we did not find transfer to measures of working memory and fluid intelligence. ERP analysis in the n-back task and the ANT indicated overlapping processes in the P3 time range. Moreover, in the ANT, we detected increased parietal activity for the interference training group alone. In contrast, we did not find electrophysiological differences between the low interference training and the control group. These findings suggest that training on an interference control task leads to higher electrophysiological activity in the parietal cortex, which may be related to improvements in processing speed, attentional control, or both.

1. Introduction

A fundamental question in the study of learning is whether acquired knowledge or skills transfer to new but similar tasks or situations. Formerly, many researchers agreed that in general, little transfer occurs (Detterman, 1993); recently, the transfer of learning has received more scientific attention due to encouraging approaches to train core mechanisms of working memory (WM). These forms of WM training have improved performance in related but untrained tasks, such as measures of fluid intelligence (gF; Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Jaeggi, Buschkuehl, Jonides, & Shah, 2011; Klingberg et al., 2005). However, inconsistencies in methodology and results have raised questions about the efficacy of WM interventions for the enhancement of fluid intelligence (cf. Morrison & Chein, 2011; Redick et al., 2012). While the factors promoting transfer are still unknown, many have argued that transfer of learning is possible to the degree that the training and transfer tasks involve overlapping neural networks or share cognitive mechanisms (Dahlin,

Stigsdotter Neely, Larsson, Bäckman, & Nyberg, 2008; Jonides, 2004; Klingberg, 2010; Perrig, Hollenstein, & Oelhafen, 2009; Persson & Reuter-Lorenz, 2008).

Methodological issues in several WM training studies have presented challenges to the replication and generalization of their findings (cf. Shipstead, Redick, & Engle, 2010). First, since cognitive training studies typically involve complex training paradigms or batteries of training tasks, isolating the cognitive processes critical for transfer proves difficult (Jaeggi et al., 2008; Klingberg et al., 2005; Olesen, Westerberg, & Klingberg, 2004; Thorell, Lindqvist, Bergman, Bohlin, & Klingberg, 2009). A second, related issue concerns the *adaptive* adjustment of difficulty in many training regimens. Since these are often compared with *non-adaptive* training tasks, the intended differences may be confounded with decreased motivation and training compliance in active control groups. These unintended differences between training and active control groups impair internal validity (cf. Shipstead et al., 2010; Shipstead, Redick, & Engle, 2012). Therefore, contrasting the training intervention with a similar active control group increases the validity of conclusions about the training program. Ideally, researchers should compare training groups that differ only in the demands of critical cognitive processes in order to disentangle relevant components of a training intervention (cf. Schneiders, Opitz, Krick, & Mecklinger, 2011). Likewise, if the training

* Correspondence to: University of Bern, Department of Psychology, Division of Experimental Psychology and Neuropsychology, Fabrikstrasse 8, 3012 Bern, Switzerland. Tel.: +41 316314734.
E-mail address: stephan.oelhafen@psy.unibe.ch (S. Oelhafen).

and transfer tasks share this critical process, the training is more likely to produce learning transfer.

Previous research has demonstrated that the ability to control interference is critical for WM and gF (Burgess, Gray, Conway, & Braver, 2011; Gray, Chabris, & Braver, 2003; Kane, 2003). Interference control involves the detection and resolution of conflicts that arise when irrelevant stimulus dimensions refer to an incorrect stimulus or response pattern (Carter & van Veen, 2007; Friedman & Miyake, 2004). In the n-back task, participants must continuously track a stream of information and decide whether a probe matches the stimulus presented n-steps back. In this task, the process of interference control becomes crucial in *lure* trials (Burgess et al., 2011; Gray et al., 2003; Kane, 2003), i.e. non-target trials with a match in non-n positions. For instance, in a 2-back task, the probe 'P' in the sequence 'P–T–W–P' is an $n+1$ lure trial, because the actual stimulus only corresponds with the stimulus $n+1$ trials before. Thus, $n+$ lures refer to items presented before the nth item, whereas $n-$ lures match with more recent stimuli. These lures induce a familiarity signal which conflicts with explicit recollection of the sequence (Oberauer, 2005; Szmalec, Verbruggen, Vandierendonck, & Kemps, 2011). As a result, the additional demand in interference control leads to increased reaction times and false alarms rates. Specifically, lures that are presented close to the critical lag ($n+1$ or $n-1$) elicit the strongest conflict (Kane, Conway, Miura, & Colflesh, 2007; McCabe & Hartman, 2008; Szmalec et al., 2011). Thus, interference control is a critical process in the n-back task, and training of this component may lead to transfer in related cognitive domains.

Similar to isolating cognitive processes necessary for transfer, mapping neural activity allows one not only to test whether trained and untrained tasks overlap in specific brain regions but also to estimate training induced activity changes (Dahlin et al., 2008). A few studies have examined the brain regions that mediate the relationship between working memory, interference control and gF. In a functional magnetic resonance imaging (fMRI) study using a visual 3-back task with $n-1$ and $n+1/+2$ lures, Gray et al. (2003) determined that the blood oxygenation level dependent (BOLD) signal in the left lateral prefrontal and bilateral parietal regions explained 99.9% of the relationship between gF and accuracy in lure trials. In trials with increased demand for interference control, participants with high gF showed increased signal in these regions. In another fMRI study, Burgess et al. (2011) found that BOLD signal in the bilateral dorsolateral prefrontal cortex (middle frontal gyrus) and the inferior parietal cortex mediated the relationship between gF and WM. Thus, regions in the lateral prefrontal and parietal cortices play a crucial role in interference control (cf. Jonides & Nee, 2006; Kane, 2003; Nee, Wager, & Jonides, 2007).

While fMRI allows for assessing the brain regions that contribute to transfer, event-related potentials (ERP) indicate the time frames in which task-critical cognitive processes overlap. ERP research on interference control has often focused on two components. The N2, a fronto-central negative component typically elicited between 200 and 350 ms post-stimulus, is associated with conflict monitoring or the resolution of conflict (Carter & van Veen, 2007; Kopp, Rist, & Mattler, 1996; Silton et al., 2010). The second component, the parietal P3, which typically appears between 300 and 600 ms post-stimulus, is thought to reflect general processes of attentional control, stimulus categorization, and the effort to choose between competing stimuli or responses (Bledowski, Prvulovic, Goebel, Zanella, & Linden, 2004; Neuhaus, Trempler et al., 2010; Neuhaus, Urbanek et al., 2010; Rueda, Posner, Rothbart, & Davis-Stober, 2004).

Based on the relationship between WM, gF, and interference control, the present study was designed to investigate the assumption that transfer effects result from partially shared neural and cognitive mechanisms between the training and transfer tasks. We tested this assumption on a behavioral and electrophysiological level with training and transfer tasks that tax interference control heavily. We used the adaptive dual n-back as a training task and manipulated the occurrence of lure trials between training groups. To test for changes in interference control, we used the Attention Network Test (ANT), a cued variant of the flanker task (Fan, McCandliss, Sommer, Raz, & Posner, 2002). Others have used this task (Fan, McCandliss, Fossella, Flombaum, & Posner, 2005), both in healthy children and adults (Fan et al., 2002; Rueda et al., 2004) and patients with various mental disorders (Fernandez-Duque & Black, 2006; Neuhaus et al., 2010a, 2010b). In the ANT, participants must quickly indicate the direction of a central arrow, which is flanked by congruent or incongruent arrows. Distinct cue-flanker combinations allow for assessing the efficiency of three attentional networks. C*onflict resolution* is manipulated by congruent and incongruent flanker arrows, and the efficiency of *alerting* and *orienting* are derived from several types of valid cues (see method section). Since we expected the lure n-back training to increase the efficiency of conflict resolution, which is thought to be closely related to interference resolution (cf. Miyake et al., 2000; Szmalec et al., 2011), we focused on this attention network for our theoretical and empirical analysis.

Our primary objective was to test whether training of shared cognitive processes in the training and transfer tasks would lead to improved performance in the ANT and corresponding electrophysiological changes. Researchers have found both increased and decreased activations after task practice, and these changes may arise from either implementing initial strategies more effectively, learning new strategies during training or the reorganization of cognitive functions (Dahlin et al., 2008; Jolles, Grol, Van Buchem, Rombouts, & Crone, 2010). Unfortunately, the heterogeneity of methods and tasks used in the field of cognitive training makes it difficult to predict changes in neuronal activity precisely (cf. Buschkuehl, Jaeggi, & Jonides, 2012). Nevertheless, based on the aforementioned fMRI and ERP studies, we expected ERP effects in the time range of the N2 and P3 components, with activation focused in the fronto-parietal network. Specifically, we predicted that we would detect changes in the time range where the training and transfer tasks overlap in a topographic map of whole brain activity. Additionally, we assessed WM capacity and gF to evaluate transfer effects after training with the n-back task (Jaeggi et al., 2008; Jaeggi et al., 2011; Studer-Luethi, Jaeggi, Buschkuehl, & Perrig, 2012). Given the reported relationship between WM, interference control, and gF, we predicted the largest transfer gains after training with lure n-back.

2. Methods

2.1. Participants

Forty eight healthy, young adults, were recruited from an academic environment and randomly assigned to the lure training group, the non-lure training group (i.e. the active control), or the passive control group. In each group, one participant dropped out between pre- and post-testing and two dropped out during pretest EEG recording, leaving a final sample of 43 participants (mean age: 25.2 years; $SD = 4.1$; range: 18–34). Both training groups consisted of 14 participants (6 female and 1 left-handed each), and 15 people were assigned to the control group (8 female and 2 left-handed). Each individual submitted a written informed consent before the experimental procedure began. After completion of the study, all participants were paid 50 Swiss Francs for participation. The study was of the type approved by the local ethics committee.

2.2. Materials

All training and transfer tasks were conducted with E-prime (Psychology Software Tools, Inc., Pittsburgh, PA). We used a simplified version of the ANT in pre- and post-testing (Fan et al., 2005), and modified the adaptive dual n-back task

(Jaeggi et al., 2008) for the purpose of lure- and non-lure differentiation. Accuracies and reaction times were recorded through a standard computer keyboard for the *n*-back and a button box for the ANT. Offline ERP analyses were conducted using BrainVision Analyzer (Brain Products, Munich, Germany) and Ragu (Koenig, Kottlow, Stein, & Melie-García, 2011) for subsequent statistical ERP analyses.

2.3. Training and transfer tasks

2.3.1. Dual n-back (training)

Each training session consisted of 20 blocks ($\approx$ 25 min), and each block consisted of 20+*n* trials, e.g. a 3-back block consisted of 23 trials. Each trial presented a blue square at one of eight possible locations around a central fixation cross. Meanwhile participants heard one of eight spoken letters (c, g, h, k, p, q, t or w) via headphones. These visual and auditory stimuli were presented simultaneously for 500 ms with an interstimulus interval (ISI) of 2500 ms. In both modalities, participants had to press a key when the currently presented stimulus matched the one presented previously. For example, in the 2-back, participants needed to press 'A' with the left index finger when the letter spoken two stimuli previously was the same as the current spoken letter. On visual match trials they needed to press the 'L' key with the right index finger. Participants did not need to press a key on non-target trials. After completion of one block, the program presented feedback about the performance in both modalities and adjusted the *n*-back difficulty level in an adaptive manner. If participants made more than 5 mistakes (i.e. misses and false alarms) in at least one modality, the *n*-level dropped in the subsequent block. If they made fewer than 3 mistakes in both modalities, the *n*-level increased by 1. Otherwise, they proceeded on the same *n*-level in the next block. In sessions 1–3, daily training sessions began with a 1-back block, and in sessions 4–14 with a 2-back block.

Irrespective of the current *n*-level, each block consisted of 6 targets per modality. Out of 14 non-target trials, the lure *n*-back training contained an average of 3 lure trials per block and modality, and we defined lures as probes that matched the stimuli with a lag of $n \pm 1$ or $n \pm 2$. For instance, in a 3-back block, probes that matched 1, 2, 4 or 5 trials back were lures. On the other hand, the non-lure *n*-back training did not have any lures.

2.3.2. Pre- and post-assessments

2.3.2.1. Dual n-back.
During pre- and posttest EEG recording, a total of 16 blocks were administered. We used the lure *n*-back for the pre- and posttesting; however unlike the training task, this task was non-adaptive and increased in difficulty from 2- to 5-back, with 4 blocks at each level of difficulty.

2.3.2.2. Attention Network Test.
We used a simplified version of the ANT that contains 144 trials with equal numbers of three cue types and two flanker conditions (Fan et al., 2005). Instructions were translated into German, and we adjusted the display size of stimuli for EEG recordings to ensure comparable visual angles. The program presented cues for 100 ms, which were displayed above, below (*spatial* cues), or in place of the fixation cross (*central* cue), or else the cue was absent. After a fixation period of 400 ms, an arrow pointing to the left or right was displayed for 1700 ms or until button press. The arrow appeared either above or below the fixation cross, and this arrow was flanked by two arrows on each side, pointing either in the same (*congruent*) or the opposite direction (*incongruent*). Using their right and left thumbs, participants aimed to quickly and accurately indicate the direction of the central arrow while neglecting the four flankers. The interstimulus interval (ISI) was variable such that ISI and RT summed up to 3500 ms (Fan et al., 2002).

2.3.2.3. Reading Span Task.
To assess WM capacity, we used the Reading Span Task with an automated procedure as described by Unsworth, Redick, Heitz, Broadway, and Engle (2009). This task presented an alternating sequence of sentences and capital letters, and participants first indicated whether the sentence made sense and then were presented a letter to commit to memory. Once a sequence of 3 to 7 sentence-letter pairs has been displayed, participants had to reproduce the sequence of letters by clicking in the correct order on letters shown on screen. The dependent variable was the sum of all correctly recalled letters (i.e. *reading span total*). The total test duration was approximately 15–20 min.

2.3.2.4. Bochumer Matrizentest.
The short version of the Bochumer Matrizentest – advanced (BOMAT) was used to determine gF (Hossiep, Turck, & Hasella, 1999). Similar to Raven's progressive matrices (Raven, 1990), this standardized test consists of 29 increasingly difficult visual analogy problems. We used parallel forms for pre- and post-tests with the original time limit of 45 min. The total number of correctly solved problems was the dependent variable.

2.4. Procedure

All participants performed the same tasks before and after an interval of 21 days to measure training induced effects. At the pretest session, we informed participants about the experimental procedure and asked them to complete

BOMAT and RST in a separate testing room. During EEG recording, they first completed the ANT and then the *n*-back. At the end of the pretest, we informed participants if they belonged to a training or the passive control group. Therefore, participants in the passive control group knew that they belonged to the control group of a training study, but participants in the two training groups were not informed about the lure/non-lure manipulation. Both training groups trained with the dual *n*-back for 14 self-administered sessions at home, distributed over 3 weeks, with a weekly schedule of 5 consecutive training days and 2 days off. Individuals in the control group were not contacted during the training interval. After this interval, participants of all three groups were tested again with the tasks in the same order (posttest).

2.5. ERP recording

The EEG measurement was conducted in an electrically shielded and air-conditioned room at the Institute of Psychology, University of Bern. The viewing distance was approximately 120 cm. We recorded from 64 Ag/AgCl electrodes mounted in the 64 channel EasyCap electrode system (EASYCAP GmbH, Herrsching-Breitbrunn, Germany) in the extended 10–20 system. Two additional electrodes placed below the left and the right eye recorded vertical eye movements, and experimenters ensured impedances stayed below 20 kΩ. The recording reference was at electrode position Fz, and two 32-channel BrainAmp MR plus amplifiers (Brain Products, Munich, Germany) sampled the activity at 500 Hz and bandpass-filtered the boosted signal between .016 and 250 Hz.

In order to correct for eye movements, we used EEGLAB runica algorithms (Delorme & Makeig, 2004) to perform an independent component analysis. Components with a distinct temporal and spatial signature associated with eye movements or blinks were then subtracted from the data. Then, in periods with remaining artifacts, we used visual inspection to remove these time periods and did not interpolate channels. The EEG data was then re-referenced to average reference, downsampled to 250 Hz, and bandpass filtered at 0.3 Hz (24 db/oct)–12.0 Hz (24 db/oct).

2.6. ERP analysis

The ERP analysis followed a similar procedure as described by Koenig and Melie-García (2009). First, we computed grand mean ERPs of correct responses by participant, session and condition (target locked, 0–800 ms, no baseline correction). Then, we applied topographic consistency tests for all time points to check if grand averages across participants reflected a partially similar configuration of generators (Koenig & Melie-García, 2010). Time points where the topographic consistency test indicated that the probability of a consistent topography across subjects was larger than .05 were excluded from further analyses. For the main analysis, we used Topographic ANOVA (TANOVA) which is a non-parametric randomization test based on global dissimilarities between electric fields (Lehmann & Skrandies, 1980; Murray, Brunet, & Michel, 2008). As opposed to more standard channel-wise statistics, a TANOVA results in a precise estimation of the probability that the variance of global maps is related to the assumed structure in the data (i.e. group, session, or condition). We first computed TANOVAs for each consecutive time point, and in order to circumvent the problem of multiple testing, we set a global duration cutoff (Koenig & Melie-García, 2009). The purpose of this cutoff was to determine if the duration of a significant time-period exceeded chance (Koenig et al., 2011).

Through our ERP analyses, we aimed to evaluate task overlap and training induced effects. First, we conducted a TANOVA of the activity during the *n*-back and the ANT to detect time frames that showed a conflict effect regardless of task and testing session. Second, in a separate TANOVA, we assessed training induced effects specific to the ANT. After observing periods above duration threshold, we computed TANOVAs on the average electrical fields in order to assure that global maps differed not just over single time points but actually over the whole period. Post-hoc channel-wise *t*-tests (*t*-maps) and standardized low resolution brain electromagnetic tomography (sLORETA) furthered our understanding of the topographic distribution and location of the neuronal generators of the map dissimilarities. We applied sLORETA based on 6239 equally spaced cortical gray matter voxels. The forward solution was obtained using the boundary element method applied to the MNI152 template (cf. Pascual-Marqui, 2002). A signal-to-noise ratio (SNR) of 100 was assumed for the computation of the pseudoinverse matrix, but pilot analyses with other SNRs yielded similar results.

2.7. Behavioral data analysis

An analysis of *n*-back training data was conducted to test for both improvements during training and differences between lure- and non-lure training. Moreover, we investigated if the lure training group improved their ability to correctly reject lures. We also tested if lure training led to improvements in the ANT, RST, and BOMAT.

Behavioral data were analyzed with analyses of variance (ANOVA), and for pairwise comparisons, we conducted Tukey HSD (within subject) and Games-Howell

(between subject) corrected tests (Howell, 2009). All reported p values resulted from two-sided testing with an alpha of 0.05. Effect-sizes are reported with partial eta squared (η_p^2). For the sake of brevity, we focus only on significant results relevant to our research questions.

3. Results

3.1. Behavioral results

3.1.1. n-back (training)

A total of four training sessions (1% of data) were missing and thus linearly interpolated for subsequent training gain analysis. One session was not registered due to technical problems and three subjects completed 13/14 training sessions due to scheduling issues. The n-back training groups reached an average n-level of $M = 2.16$ ($SD = 0.38$) in the first session and an average of $M = 3.78$ ($SD = 0.65$) in the final session (Fig. 1). As expected, participants in the lure training group reached a slightly lower n-level in their final session. A repeated measures ANOVA on mean n-level with the within-subjects factor session (average of 1st/2nd or 13th/14th) and the between-subjects factor training group (lure n-back or non-lure n-back) revealed a main effect of session, $F(1, 26) = 216.56$, $p < 0.001$, $\eta_p^2 = 0.89$. There were no significant differences between the two training groups, irrespective of session, although a trend was evident, $F(1, 26) = 3.76$, $p = 0.06$, $\eta_p^2 = 0.13$. The session by training group interaction was not significant, $F(1, 26) < 1$. Taken together, the training groups improved in mean performance during training, but they did not differ significantly in terms of training gain.

In order to test if participants in the lure training group improved in rejecting lures, we performed a second analysis based on false alarm rates (FA; i.e. average accuracies of both modalities in non-target trials). First, a repeated measures ANOVA with the within-subjects factors of trial type (lure or non-lure) and training session (average of 1st/2nd or 13th/14th) revealed a significant main effect of trial type indicating more lure trial FAs ($F(1, 13) = 66.93$, $p < 0.001$, $\eta_p^2 = 0.84$; Fig. 2). The main effect of session approached significance ($F(1, 13) = 4.08$, $p = 0.06$, $\eta_p^2 = 0.24$). Moreover, a significant type by session interaction ($F(1, 13) = 22.79$, $p < 0.001$, $\eta_p^2 = 0.64$) indicated a significant reduction in lure trial FAs ($t(13) = 3.21$, $p < 0.01$); however, in non-lure trials, we noted a non-significant trend of FA increase from initial to final training sessions ($t(13) = -1.84$, $p = 0.09$). Tukey HSD corrected pairwise comparisons further indicated that in the initial training sessions, $n-2$ lures had a significantly reduced rate of FAs compared to $n-1$, $n+1$ ($ps < 0.05$) and $n+2$ lures ($p < 0.01$), indicating a processing advantage for recently displayed items. At the end of the training, lure trials closer to the critical lag of n (i.e. $n-1$ and $n+1$) still yielded higher FA rates than $n-2$ lures, $ps < 0.01$. Interestingly, the improvement in $n+$ lures was significantly higher than in $n-$ lures, $t(13) = 3.16$, $p < 0.01$.

These results indicate that participants in the lure training group improved their ability to reject lures, especially in the
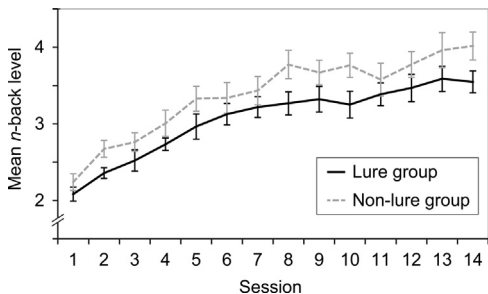

Fig. 1. Average n-back training level by group (lure or non-lure training) as a function of session. Error bars represent standard errors of the mean.
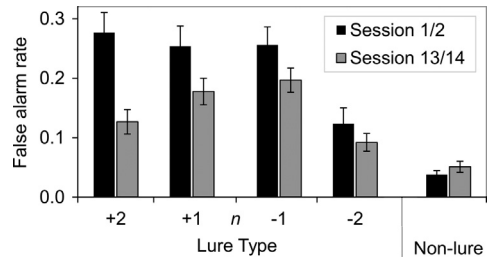

Fig. 2. n-back type of trial. Bars represent false alarm rates in non-target trials of both modalities for non-lure and four different kinds of lure trials in sessions 1/2 and 13/14 of the lure n-back training. $n-2$ lures are closer to the current probe and show a reduced false alarm rate, while $n+1$ lures close to the critical lag of n show an increased rate irrespective of training session (see text). Error bars represent the standard error of the mean.

Table 1
Descriptive data of pre-/post training assessment: Mean (SD).

Task	Group	Pre-test	Post-test
n-back			
	Lure	0.33 (0.09)	0.72 (0.11)
	Non-lure	0.31 (0.10)	0.68 (0.06)
	Control	0.33 (0.10)	0.40 (0.12)
ANT			
Congruent	Lure	481 (42)	440 (47)
	Non-lure	469 (38)	452 (49)
	Control	477 (44)	468 (53)
Incongruent	Lure	562 (55)	513 (56)
	Non-lure	569 (59)	522 (64)
	Control	572 (54)	548 (58)
RST			
	Lure	57.5 (10.5)	59.7 (10.8)
	Non-lure	56.5 (9.4)	58.4 (9.1)
	Control	58.7 (8.4)	62.0 (5.3)
BOMAT			
	Lure	15.3 (2.9)	18.5 (2.4)
	Non-lure	15.7 (5.0)	18.9 (5.1)
	Control	16.5 (5.0)	19.9 (4.7)

Note. $n = 43$. n-back: accuracies (rates of hits – false alarms) in 2- to 5-back blocks. ANT: Attention Network Test reaction times (ms) by congruency. RST: Automated Reading Span Task, total sum of correctly recalled letters. BOMAT: Bochumer Matrizen Test, total of correctly solved problems.

distant $n+$ lures. Lure trials led to higher FA rates compared to non-lure trials, even at the end of training. Lures closer to the critical lag ($n-1$ and $n+1$) led to the highest FA rates.

3.1.2. Pre- and post-assessments

Descriptive data of pre- and post-assessment tests are displayed in Table 1. One-way ANOVAs with the between-subjects factor group (lure n-back, non-lure n-back, or no training) did not reveal any significant group differences at pretest, all $Fs < 0.32$, $ps > 0.73$.

3.1.2.1. Non-adaptive dual n-back.
We calculated pr values (hit rate – false alarms rate) to compare performance of the non-adaptive dual n-back in pre- and post-test. An ANOVA with the between-subjects factor group (lure n-back, non-lure n-back, or no training) and the within-subjects factor session (pre- or post-test) revealed a main effect of session, $F(1, 40) = 509.56$, $p < 0.001$, $\eta_p^2 = 0.93$, and a significant session by group interaction, $F(2, 40) = 81.98$, $p < 0.001$, $\eta_p^2 = 0.80$. Importantly, when the analysis was restricted to the two training groups, the critical session by group interaction was not significant, $F(1, 26) < 1$, $p = 0.59$, $\eta_p^2 = 0.01$, and the main effect group did not reach level of significance, $F(1, 26) = 1.02$, $p = 0.32$,

$\eta_p^2 = 0.04$. Therefore, although the non-lure training group tended to reach higher levels during training, performance in the non-adaptive lure n-back did not differ between the two training groups.

3.1.2.2. Attention Network Test. Overall, participants performed with very high accuracy in the ANT and higher in the congruent ($M_{ACC} = 0.99$, $SD_{ACC} = 0.02$) compared to the incongruent condition ($M_{ACC} = 0.94$, $SD_{ACC} = 0.05$). There were no reliable changes associated with group. Therefore, the focus of the ANT analysis was on reaction times (RT). We analyzed RTs within a range of 200–1000 ms (2.3% of data excluded), and we used correct responses for further analysis. RTs were shorter in the congruent condition ($M_{RT} = 465$ ms, $SD_{RT} = 44$) compared to the incongruent condition ($M_{RT} = 548$ ms, $SD_{RT} = 55$; Table 1).

A repeated measures ANOVA on reaction times with the between-subjects factor group (lure n-back, non-lure n-back, or no training) and the within-subjects factors session (pre- or post-test), congruency (congruent or incongruent) and cue (spatial, central, or none) revealed a main effect of session ($F(1, 40) = 53.51$, $p < 0.001$, $\eta_p^2 = 0.57$). This indicated that RTs were significantly shorter in the post-test, and a session by congruency interaction ($F(1, 40) = 29.35$, $p < 0.001$, $\eta_p^2 = 0.42$) showed that these RT reductions were more pronounced in the incongruent condition.[1] Importantly, there was also a significant group by session by congruency interaction, $F(2, 40) = 3.97$, $p < 0.05$, $\eta_p^2 = 0.17$. We decomposed this three-way interaction into two ANOVAs, separated by congruency. In the congruent condition only, the session by group interaction was significant (congruent: $F(2, 40) = 6.00$, $p < 0.01$, $\eta_p^2 = 0.23$; incongruent: $F(2, 40) = 2.66$, $p = 0.08$, $\eta_p^2 = 0.12$). In this condition, the lure training group improved significantly more than both the non-lure ($t(26) = 2.68$, $p < 0.05$) and the control group ($t(27) = 3.57$, $p < 0.01$), whereas the non-lure and the control group did not differ ($t(27) = 0.76$, $p = 0.73$; Games-Howell corrected). Finally, regardless of congruency and cue, a significant session by group interaction ($F(2, 40) = 3.96$, $p < 0.05$, $\eta_p^2 = 0.16$) indicated group-specific RT changes. However, this interaction was mainly driven by the significant session by group interaction found in the congruent condition. Specifically, as indicated by descriptive data (cf. Table 1), there were no reliable differences between the lure and the non-lure training group in the incongruent condition.

In summary, RTs decreased significantly from pre- to post-test, and this decrease was more pronounced in the incongruent condition regardless of group. While there was a tendency towards differential improvements in the incongruent condition, we only found reliable group differences in the congruent condition. In this condition, the lure training group improved significantly more than the other two groups.

3.1.2.3. Reading Span Task and Bochumer Matrizentest. In the RST, the total number of correctly recalled letters was $M = 57.6$ ($SD = 9.3$) in the pretest and $M = 60.1$ ($SD = 8.6$) in the post-test. A repeated measures ANOVA with the between-subjects factor group (lure n-back, non-lure n-back, or no training) and the within-subject factor session (pre- or post-test) revealed no significant main effect of group, no interaction (both $Fs < 1$), and only a statistical trend for session, $F(1, 40) = 3.35$, $p = 0.07$, $\eta_p^2 = 0.08$. In the BOMAT, the number of correctly solved analogy problems increased from $M = 15.9$ ($SD = 4.3$) in the pretest to $M = 19.1$ ($SD = 4.2$) in the post-test, which was confirmed by a significant main effect of session, $F(1, 40) = 38.24$, $p < 0.001$,

$\eta_p^2 = 0.49$. However, the main effect of group and the relevant session by group interaction did not reach significance, $Fs < 1$. Also, combining the two training groups and comparing them to the passive control group did not reveal a main effect of group or a group by session interaction for RST and BOMAT (all $Fs < 1$). Thus, the RST and BOMAT scores were higher in the posttest, but neither the lure training nor the non-lure training group showed a higher pre–post gain compared to the control group.

3.2. ERP results

3.2.1. Overlap analysis

The post-stimulus ERPs are displayed in Fig. 3a. In the n-back task, lure and non-lure trials displayed a strong fronto-central positivity from 150 to 300 ms and a parietal P3 complex from 200 to 600 ms. We also detected a frontal positive component between 150 and 300 ms in the ANT, which was also followed by a pronounced parietal P3 component. In order to investigate the overlap of components in the n-back and ANT, we conducted a repeated measures TANOVA with the within-subject factors task (n-back or ANT) and conflict (lure/incongruent or non-lure/congruent). The main effect of task was significant within the entire time range tested (0–800 ms). Thus, the two tasks clearly differed in topography. Interestingly, we observed a significant main effect of conflict within 332–704 ms and a significant task by conflict interaction within 344–444 and 564–680 ms (Fig. 3b). Although the conflict effect was more pronounced in the ANT, both tasks showed a similar conflict effect within 332–704 ms, with a positive centroid moving from the anterior to posterior electrodes (see the difference maps in Fig. 3c). This overlap was most clear in the range of 444–564 ms. Interestingly, in this time period with a significant conflict effect, there was no interaction of task and conflict.

These results indicate that although the two tasks do differ, there is a substantial overlap in topography of the conflict effect in the P3 range from 444 to 564 ms, and the absence of an interaction indicates that there is no evidence for a difference of this conflict effect between the two tasks. This suggests that from 444 to 564 ms similar activity may underlie both types of interference control, and lure-training has induced changes in this time window that have an effect on conflict processing in the ANT task. Hence, transfer is most likely to be promoted by the processes underlying the P3 component and common for both tasks within this time period.

3.2.2. Transfer analysis

In order to analyze training effects on ERPs in the ANT, we computed a repeated measures TANOVA with the between-subjects factor group (lure n-back, non-lure n-back, or no training) and the within-subjects factors session (pre- or posttest) and congruency (congruent or incongruent). We found several significant effects: an effect of congruency from 260 to 800 ms, a main effect of session within 0–196 ms, 212–488 ms, and 564–764 ms, and a group by session interaction within 300–564 ms. No other main effect or interaction reached significance. Since only the group by session interaction was relevant to our research questions, this became the focus of our subsequent analyses.[2]

We computed a group by session TANOVA on averaged activity in the interval within 300–564 ms, and again, the interaction reached significance, $p < 0.001$. Subsequent pairwise comparisons

---

[1] The main effects for congruency ($F(1, 40) = 440.45$, $p < 0.001$, $\eta_p^2 = 0.92$), cue ($F(2, 80) = 295.23$, $p < 0.001$, $\eta_p^2 = 0.88$) and a significant interaction between these factors ($F(2, 80) = 17.11$, $p < 0.001$, $\eta_p^2 = 0.30$) are task inherent and were not further analyzed. Besides the results reported in the text, no other main effects or interactions were significant.

[2] A separate TANOVA with the factors group, session, and cue (spatial, central, or none) revealed a significant main effect of cue within the full range (0–800 ms) but no significant interactions (session by group or session by group by cue) relevant to the current study.
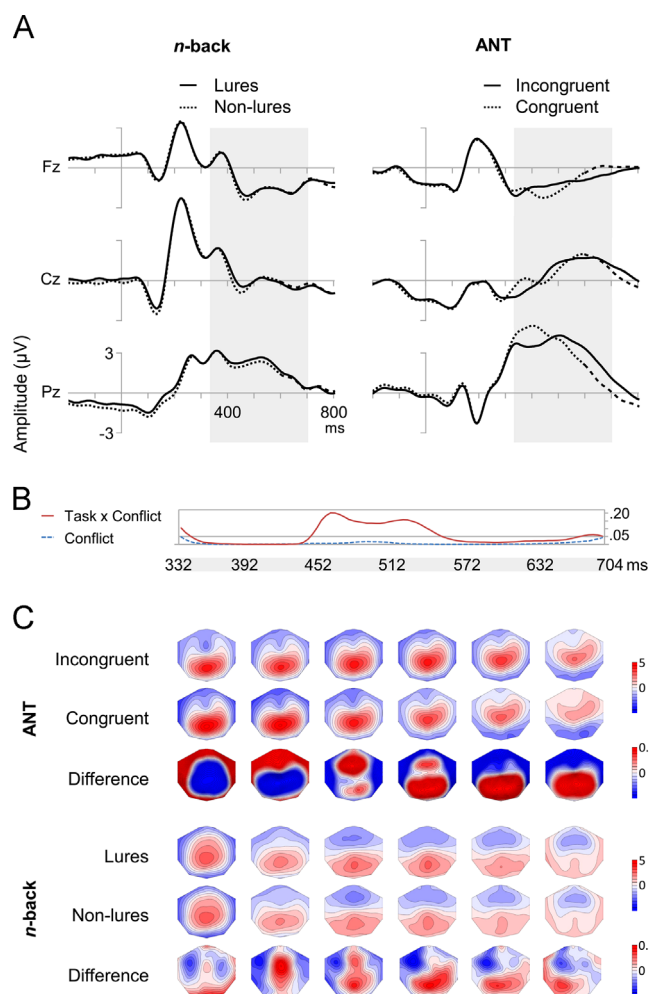
Fig. 3. (A) Post-stimulus ERPs. Grand-averaged ERPs are displayed in selected channels for lure and non-lure trials in the *n*-back task (left) and incongruent and congruent trials in the ANT (right). Gray bars indicate TANOVA time points with a significant main effect of conflict, indicating partially overlapping topography between the two tasks (see text). (B) Results of the TANOVA task by conflict. *p*-values are displayed for main effect of conflict and the task by conflict interaction. The gray line represents a threshold of $p=0.05$. Whereas the main effect of conflict is significant for all time points within this period, the significant interaction within 344–444 ms and 564–680 ms suggests partly differential conflict effects in the two tasks. (C) Maps of electrophysiological activity are shown for incongruent/congruent trials in ANT and lure/non-lure trials in *n*-back. Difference maps reflecting conflict effects in both tasks suggest considerable overlap in the time frame of 444–564 ms.

between groups yielded significant differences between the lure training and the other groups ($ps < .001$), but we did not discover differences between the non-lure training and control group ($p=0.15$). In order to clarify the localization of the pre/post-differences between groups, we calculated post- minus pretest *t*-maps for the lure training group and compared them to the combined non-lure training and control group (Fig. 4). The *t*-map reflecting the group by session interaction revealed increased fronto-central positivity ($t_{max}=2.99$ at FC4) and posterior negativity ($t_{min}=-4.07$ at Iz) for the lure training group, compared to pre-/post-test differences for the control groups. Finally, we computed a LORETA on the posttest data using the pretest data as a baseline. This allowed us to reveal tomographic activity changes after lure training compared to pre-/post-test differences in the non-lure and control group (Fig. 4). The LORETA revealed three clusters with increased activation for the lure training group. 74% of all the activated voxels above threshold ($t \geq 2.02$) were located within a large cluster that compromised the postcentral
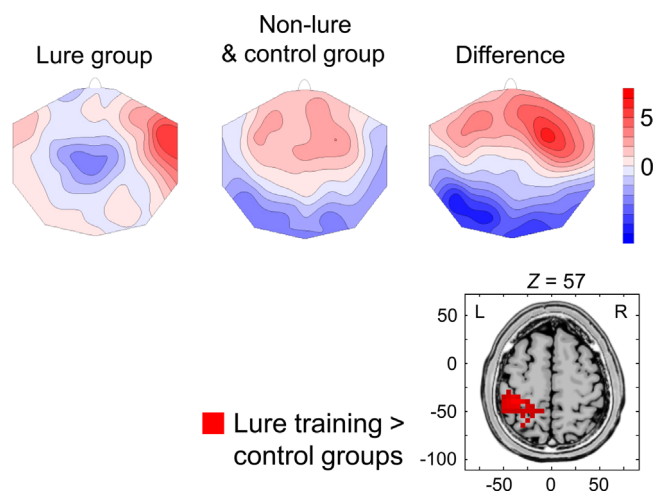

Fig. 4. ANT training effect *t*-Maps/LORETA. Attention Network Test *t*-maps reflecting session effect (post- minus pretest). Session differences are displayed for the lure training and control groups (non-lure training and passive control group combined) within 300–564 ms post-stimulus, as well as the difference between the two, reflecting the interaction of session (post- minus pretest) and group (lure training minus control groups). This difference reflecting the session by group interaction is also reflected in the LORETA, which indicates training induced activity increases for the lure training group. Activation display is thresholded at $|t| \geq 2.02$ ($p < 0.05$, uncorrected).

gyrus (BA 5/7/40), precuneus (BA 7), and the superior (BA 7) and inferior parietal lobule (BA 40), which was also the source of maximal activation (BA 40; $t=2.76$). Additionally, two small clusters centered in the right hemisphere in superior temporal gyrus (BA 41; $t=2.28$) and middle occipital gyrus (BA 19; $t=2.35$) reached significance. There was no indication of significant decreases in activation level.

Interestingly, this session by group interaction (300–564 ms) coincided with the main effect of conflict that we found in the electrophysiological overlap analysis (332–704 ms), which includes the time frame of the conflict effect that did not coincide with the task by conflict interaction (444–564 ms). A separate TANOVA on ANT ERPs indicated that the session by group interaction was also significant within this latter period (444–564 ms). Within these time windows, the lure training group differed significantly from the non-lure training ($p < 0.01$) and control group ($p < 0.05$), yet again, the non-lure training and the control group did not differ from each other ($p=0.39$).

In summary, the first analysis showed overlapping topography between the *n*-back task and the ANT in a time range 332–704 ms, and this overlap was more pronounced within 444–564 ms post-stimulus. This difference of incongruent/lure trials and congruent/non-lure trials showed a positive fronto-central component followed by a parietal positivity. In a second analysis, the session by group interaction indicated that change in ERPs in the lure training group differed significantly from the one in the non-lure training and control group. Using LORETA, we analyzed these differences further and noted an increased activity in left parietal scalp sites after the lure training.

4. Discussion

The present study investigated behavioral and electrophysiological effects following 14 sessions of adaptive dual *n*-back training. To estimate improvements in interference control, we compared performance of two groups training with lure and non-lure variants of the adaptive dual *n*-back with a passive control group. Lure trials were defined as non-targets with a match within

the range of ± 2 items around the critical lag of n. In order to assess the potential benefits of improved interference control, we administered three tasks measuring control of attention, working memory, and fluid intelligence. We observed task-specific improvements after n-back training, yet behavioral measures of transfer revealed unexpected results. In the ANT, we expected conflict resolution to improve most after lure n-back training. However, lure training led to faster responses only in the *congruent* condition, while participants improved more in the *incongruent* condition regardless of group. Interestingly, we did not note any differential effects in the Reading Span Task and BOMAT, i.e. neither type of n-back training led to transfer in these tasks measuring WM and gF. By contrast, ERP analyses revealed unique lure training induced topographical differences in the P3 range, while we observed no differences between the non-lure training and the control group. Using LORETA, we localized these activations to the left parietal lobule, which displayed higher activity after training with lure n-back. In a separate analysis, we determined that conflict induced by the lures in n-back task and the incongruent flankers in the ANT resulted in similar ERPs, although we noted obvious differences between these tasks as well. This suggests that there may exist some similarities in the neural response to lures that produce proactive interference and the neural response to the presence of conflict in the flanker task. This similarity may reflect neural processes in the left parietal cortex common to both conditions, which would provide a basis for transfer between these tasks.

A prerequisite for transfer is improvement on the training task; further, transfer of interference control would specifically require improvement in the rejection of lures in the n-back task. Therefore, we expected increased average n-levels for both training groups and, more specifically, an improved ability to reject lures for the lure n-back training group. Both training groups improved during training, but the average n-level was slightly higher during non-lure training. One could argue that the higher difficulty in the lure training contributed to the observed electrophysiological and behavioral effects. However, as the n-level adjusted to individual performance level in both n-back tasks, we assume that overall difficulty did not differ. This assumption is supported by the results of a post-session questionnaire indicating that participants in both groups perceived the tasks as equally challenging.[3] Therefore, it is rather unlikely that differences in task difficulty can explain such group differences. More importantly, during the training, the lure n-back group decreased their rate of false alarms in lure trials, whereas false alarms of non-lure trials were constantly low. Interestingly, participants' rate of false alarms did decrease more in distant $n+$ than in close $n-$ lures. This might indicate that they learned to distinguish between items in $n-$ positions that are still needed for further processing and $n+$ items that are no longer relevant (Szmalec et al., 2011). This could be interpreted as an increase in efficiency of updating working memory, or may be due to an overall increase in speed of processing in the lure training group. Generally, we observed the highest false alarm rates in lure trials close to the critical lag ($n-1$/$n+1$). This indicates the participants' sensitivity to interference in updating tasks since weak bindings of items with corresponding serial positions may hinder counteracting familiarity-based decisions (Kane et al., 2007; Schmiedek, Li, & Lindenberger, 2009; Szmalec et al., 2011).

The major improvement in rejecting $n+$ lures indicates that participants in the lure training group in fact learned to distinguish between relevant items and no longer relevant but still interfering items. Interestingly, it has been suggested that in tasks where interference trials are *frequent* and therefore can be anticipated, interference control is active prior to conflict onset (i.e. proactive control; Braver, Gray, & Burgess, 2007; Burgess & Braver, 2010; Burgess et al., 2011). Also, Mevorach, Humphreys, and Shalev (2009) showed that the left posterior parietal cortex is involved in *presetting* a system to ignore salient distractors or to preselect a window of attention. Because of the high interference expectancy in the ANT (50%), a shift from *reactive* to *proactive* control would explain the observed improved performance irrespective of interference onset. Moreover, it is interesting to see that the region around the left intraparietal sulcus, which was more active after lure training, has been associated with proactive or preparatory attentional processes.

While improvements on the training task are necessary for transfer to occur, they are not sufficient to prove an enhancement in ability. Improvements of training performance may also reflect newly developed strategies specific to the training task, learned and automated stimulus-response configurations or other unintended task-specific changes (cf. Chooi & Thompson, 2012; Klingberg, 2010; Shipstead et al., 2010). Consequently, although the lure n-back and the ANT do not seem to share many task features, we expected enhanced conflict resolution in the ANT on the assumption that processes engaged in both tasks partially overlap. However, some reports indicate that resolution of proactive interference in n-back and conflict resolution in flanker tasks engage different inhibition processes (cf. Bissett, Nee, & Jonides, 2009; Nee et al., 2007). By contrast, Szmalec et al. (2011) suggest that interference in n-back involves similar mechanisms as in several response-congruency tasks. Our overlap analysis of ERPs in the n-back task and ANT showed similar conflict effects in the P3 range (332–704 ms). One must note that this data-driven ERP analysis does not prove or quantify the degree of electrophysiological overlap, but is instead suggestive of a similarity in neuronal activity and cognitive processes required in both tasks. More importantly, the analysis is primarily informative about the relevant time range where tasks overlap and indicates which processes may promote transfer.

The analysis of training induced electrophysiological changes in the ANT yielded reliable group differences in the P3 time range (300–564 ms post-stimulus). Compared to active and passive control group, the lure training led to increased activity in the inferior and superior parietal lobule in the left hemisphere. By contrast, 14 sessions of n-back training without lures did not reveal any electrophysiological effects compared to the passive control group. Such activation differences can be interpreted based on previous studies, showing that the regions around the intra-parietal sulcus support the voluntary control of attention, including shifts of attention (Hopfinger, Buonocore, & Mangun, 2000; Walsh, Buonocore, Carter, & Mangun, 2011).

Taken together, the analyses on ERPs suggest that overlap and transfer may entail more generic attentional effects, and not interference control specifically. First, the overlap analysis did not include the earlier N2 component, which has been associated with conflict processing and possibly has a neural generator in the anterior cingulate cortex (ACC; Carter & van Veen, 2007; Kopp et al., 1996; Silton et al., 2010). Therefore, although n-back lures and incongruent flankers have been associated with activation in the ACC together with other prefrontal regions (Burgess et al., 2011; Fan et al., 2005; Gray et al., 2003; Nee et al., 2007), the tasks may differ in respect of the time course of conflict monitoring or the resolution of a conflict. Second, given the large time period reflecting overlapping conflict effects that exceeds response times

---

[3] At the beginning of the posttest, subjects in the training groups had to indicate on a scale from 1 to 9 how much effort and concentration the training required. *T*-tests indicated that the mean rating for both items did not differ between groups (effort: lure group: $M=6.14$, $SD=1.83$, non-lure group: $M=6.71$, $SD=1.73$; $t(26)=0.85$, $p=0.40$; concentration: lure group: $M=8.36$, $SD=1.01$, non-lure group: $M=8.57$, $SD=0.65$; $t(26)=0.67$, $p=0.51$).

in the ANT, it is most likely that these similar activations induced by the conflict manipulation reflect more general aspects of cognitive or attentional control and not the detection of a conflict or the action of inhibition. Third, the LORETA we conducted revealed activation differences in the lure group which were restricted to posterior, mainly left parietal sites, and did not include any frontal regions. Therefore, the electrophysiological data do not suggest specific effects of lure n-back training on interference control, but the training may have improved attentional control.

While lure training induced electrophysiological changes irrespective of congruency, it only had a behavioral effect on the congruent condition of the ANT. In general, one could argue that the reduced reaction times simply occurred because of an increased speed of processing, similar to that observed after training with action video games (Dye, Green, & Bavelier, 2009). It is assumed that in n-back tasks without lures, performance is mainly based on familiarity matching, requiring rather little cognitive control (Szmalec et al., 2011). By contrast, in the lure n-back task, cognitive control is constantly needed to maintain task goals, prioritize recollection over familiarity and constrain attention only to the relevant items, i.e. the last n presented items. Therefore, these subjects may have learned that they need to make faster decisions after the stimulus presentation. E.g. if they need 1 s to decide whether the current stimulus is a target or a non-target, there are 2 s left till the next stimulus presentation to rehearse n stimuli in each of the two modalities (cf. Barrouillet, Bernardin, & Camos, 2004). This would result in faster reaction times in the ANT irrespective of congruency, but would not lead to improved performance in RST and BOMAT, which are less sensitive for stimulus processing.[4] However, it remains an open question whether the shortened RTs for the lure training group in the n-back task and the ANT reflect an improvement of a cognitive ability or a simple strategy shift. In general, it has been argued that strategies tend to be limited to a particular task or type of information (cf. Morrison & Chein, 2011). On the one hand, it is very likely that subjects in the lure training group developed a different strategy during 14 days of n-back training. On the other hand, the application of a new n-back strategy does not seem to explain shortened RTs in the ANT. At least a simple strategy to make faster decisions would likely lead to decreased reaction times and to decreased accuracies (i.e. speed-accuracy trade-off). However, accuracies were generally high and did not deteriorate more in the lure training group. Therefore, it seems rather unlikely that subjects training with lures specifically transferred a learned strategy from the n-back task to the ANT, but our study provides no definite answer to that question.

Given the reported relationship between WM, interference control, and gF and the earlier reports of far transfer observed after training with the adaptive n-back task (Jaeggi et al., 2008, 2011; Studer-Luethi et al., 2012), we expected training induced improvements in measures of WM and gF. However, we found no reliable group differences in either task. One interpretation of this finding is that the effects were too small to promote reliable far transfer effects. Furthermore, WM and gF may be stable individual characteristics that are not easily modified with a few weeks of cognitive training. A recent report by Redick et al. (2012) seems to confirm this conclusion. In their study, 20 sessions of adaptive dual n-back training did not lead to any substantial performance improvements in several tasks measuring fluid and crystallized

intelligence, WM, multitasking, and perceptual speed. Altogether, we conclude that overlapping cognitive processes may be a requirement for transfer to occur, but the overlap itself is obviously not sufficient. Nevertheless, it is still questionable if working memory and fluid intelligence are trainable at all.

While our study was designed to increase the validity of conclusions about the training program, some important limitations should be noted. We used an n-back paradigm which places heavy demands on inference control and expected to find improved interference control and transfer to related domains. The lure manipulation elicited the expected results with increased false alarm rates around the critical lag of n, and the lure training group was able to improve the lure rejection during training. As mentioned before, the electrophysiological and behavioral data do not suggest that inference control improved in the lure training group, but can be interpreted as either an improvement in attentional control or processing speed. Importantly, we included only one task that measured the cognitive ability we attempted to train (i.e. interference control). In order to minimize task specific variance and reduce the influence of shared task features that are common to training and transfer tasks, transfer should be assessed on the level of latent variables which were derived from several tasks measuring the same underlying construct (Morrison & Chein, 2011; Shipstead, Redick, et al., 2012). Moreover, the design of the current study did not permit to compare lure/non-lure performance between the two training groups. In order to keep the testing sessions reasonably short, the n-back task did not include enough trials to assess reliable accuracy data on the level of lures and non-lures, and this is a clear limitation of the current study. While in our task, key presses were only necessary for n-back matches, future studies should choose a paradigm where subjects need to press a key in mismatch trials as well. This would allow for assessing precisely the effect of the lures on the reaction times in current and subsequent trials (Szmalec et al., 2011). Furthermore, we included a non-adaptive n-back in order to run the same experimental procedure for all subjects and testing sessions. However, without training, the higher n-levels are very difficult and may have led to decreased motivation and lower reliability. Future studies should therefore include reliable neural and behavioral indices of training and transfer that could also be compared between the experimental groups.

With the current study, we addressed a critique by Morrison and Chein (2011), who note that most training studies did not control variables such as training motivation or task difficulty, which may confound with different training regimens. For example, participants training with adaptive as opposed to non-adaptive tasks may be more convinced that this task will result in cognitive improvement (cf. Shipstead, Hicks, & Engle, 2012; Shipstead, Redick, et al., 2012). Therefore, placebo effects cannot be controlled completely with non-adaptive training tasks. Furthermore, we address critiques from Boot, Blakely, and Simons (2011), as the participants in our study had no way of knowing whether they belonged to the experimental or active control group. Since we administered two adaptive dual n-back variants, which differed only in demand of a single executive core process, we believe that the noted differences between the two training groups can clearly be traced back to this manipulation. Also, though we found electrophysiological and behavioral differences in the ANT between the lure n-back training and the control group, there was no indication of any reliable difference between the non-lure training and the control group. We therefore believe that future studies should pursue this methodological approach in order to obtain valid results about the trainability of WM and related abilities.

---

[4] This interpretation is backed up by a post-hoc analysis of average reaction times during lure and non-lure training. Average reaction times of hits in the initial two sessions did not differ between groups (lure group: $M = 1187$ ms, $SD = 179$ ms; non-lure group: $M = 1228$ ms, $SD = 231$ ms; $t(26) = 0.53$, $p = 0.60$), but differed significantly in the last two sessions (lure group: $M = 836$ ms, $SD = 169$ ms; non-lure group: $M = 1000$ ms, $SD = 230$ ms; $t(26) = 2.16$, $p < 0.05$).

Hollenstein and two anonymous reviewers for their helpful comments and suggestions.

References

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*, *133*, 83–100.

Bissett, P. G., Nee, D. E., & Jonides, J. (2009). Dissociating interference-control processes between memory and response. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1306–1316, http://dx.doi.org/10.1037/a0016537.

Bledowski, C., Prvulovic, D., Goebel, R., Zanella, F. E., & Linden, D. E. J. (2004). Attentional systems in target and distractor processing: A combined ERP and fMRI study. *Neuroimage*, *22*, 530–540, http://dx.doi.org/10.1016/j.neuroimage.2003.12.034.

Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do action video games improve perception and cognition? *Frontiers in Psychology*, *2*, 1–6, http://dx.doi.org/10.3389/fpsyg.2011.00226.

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In: A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. Towse (Eds.), *Variation in working memory*. Oxford: Oxford University Press.

Burgess, G. C., & Braver, T. S. (2010). Neural mechanisms of interference control in working memory: Effects of interference expectancy and fluid intelligence. *PLos One*, *5*, e12861, http://dx.doi.org/10.1371/journal.pone.0012861.

Burgess, G. C., Gray, J. R., Conway, A. R., & Braver, T. S. (2011). Neural mechanisms of interference control underlie the relationship between fluid intelligence and working memory span. *Journal of Experimental Psychology: General*, *140*, 674–692, http://dx.doi.org/10.1037/A0024695.

Buschkuehl, M., Jaeggi, S. M., & Jonides, J. (2012). Neuronal effects following working memory training. *Developmental Cognitive Neuroscience*, *2*(Suppl. 1), S167–S179, http://dx.doi.org/10.1016/j.dcn.2011.10.001.

Carter, C. S., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. *Cognitive Affective & Behavioral Neuroscience*, *7*, 367–379.

Chooi, W.-T., & Thompson, L. A. (2012). Working memory training does not improve intelligence in healthy young adults. *Intelligence*, *40*, 531–542, http://dx.doi.org/10.1016/j.intell.2012.07.004.

Dahlin, E., Stigsdotter Neely, A., Larsson, A., Bäckman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science*, *320*, 1510–1512, http://dx.doi.org/10.1126/science.1155466.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21, http://dx.doi.org/10.1016/j.jneumeth.2003.10.009.

Detterman, D. (1993). The case for the prosecution: Transfer as an epiphenomenon. In: D. Detterman & R. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 1–24). Norwood, NJ: Ablex.

Dye, M. W., Green, C. S., & Bavelier, D. (2009). Increasing speed of processing with action video games. *Current Directions in Psychological Science*, *18*, 321–326, http://dx.doi.org/10.1111/j.1467-8721.2009.01660.x.

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *Neuroimage*, *26*, 471–479, http://dx.doi.org/10.1016/j.neuroimage.2005.02.004.

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, *14*, 340–347.

Fernandez-Duque, D., & Black, S. E. (2006). Attentional networks in normal aging and Alzheimer's disease. *Neuropsychology*, *20*, 133–143, http://dx.doi.org/10.1037/0894-4105.20.2.133.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *Journal of Experimental Psychology: General*, *133*, 101–135.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, *6*, 316–322.

Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nature Neuroscience*, *3*, 284–291.

Hossiep, R., Turck, D., & Hasella, M. (1999). *Bochumer Matrizentest. BOMAT – advanced – short version*. Hogrefe: Göttingen.

Howell, D. C. (2009). *Statistical methods for psychology* (7th edition). Belmont, CA: Wadsworth.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, USA*, *105*, 6829–6833.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences, USA*, *108*, 10081–10086, http://dx.doi.org/10.1073/pnas.1103228108.

Jolles, D. D., Grol, M. J., Van Buchem, M. A., Rombouts, S. A. R. B., & Crone, E. A. (2010). Practice effects in the brain: Changes in cerebral activation after working memory practice depend on task demands. *Neuroimage*, *52*, 658–668, http://dx.doi.org/10.1016/j.neuroimage.2010.04.028.

Jonides, J. (2004). How does practice makes perfect? *Nature Neuroscience*, *7*, 10–11.

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience*, *139*, 181–193, http://dx.doi.org/10.1016/j.neuroscience.2005.06.042.

Kane, M. J. (2003). The intelligent brain in conflict. *Trends in Cognitive Sciences*, *7*, 375–377, http://dx.doi.org/10.1016/S1364-6613(03)00193-1.

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the N-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *33*, 615–622, http://dx.doi.org/10.1037/0278-7393.33.3.615.

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, *14*, 317–324, http://dx.doi.org/10.1016/j.tics.2010.05.002.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlstrom, K., et al. (2005). Computerized training of working memory in children with ADHD-A randomized, controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*, 177–186.

Koenig, T., Kottlow, M., Stein, M., & Melie-García, L. (2011). Ragu: A free tool for the analysis of EEG and MEG event-related scalp field data using global randomization statistics. *Computational Intelligence and Neuroscience*, *2011*, 938925, http://dx.doi.org/10.1155/2011/938925.

Koenig, T., & Melie-García, L. (2009). Statistical analysis of multichannel scalp field data. In: C. M. Michel, T. Koenig, D. Brandeis, L. R.R. Gianotti, & J. Wackermann (Eds.), *Electrical Neuroimaging* (pp. 169–190). Cambridge: Cambridge University Press.

Koenig, T., & Melie-García, L. (2010). A method to determine the presence of averaged event-related fields using randomization tests. *Brain Topography*, *23*, 233–242, http://dx.doi.org/10.1007/s10548-010-0142-1.

Kopp, B., Rist, F., & Mattler, U. (1996). N200 in the flanker task as a neurobehavioral tool for investigating executive control. *Psychophysiology*, *33*, 282–294.

Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, *48*, 609–621.

McCabe, J., & Hartman, M. (2008). Working memory for item and temporal information in younger and older adults. *Aging, Neuropsychology and Cognition*, *15*, 574–600, http://dx.doi.org/10.1080/13825580801956217.

Mevorach, C., Humphreys, G. W., & Shalev, L. (2009). Reflexive and preparatory selection and suppression of salient information in the right and left posterior parietal cortex. *Journal of Cognitive Neuroscience*, *21*, 1204–1214, http://dx.doi.org/10.1162/jocn.2009.21088.

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.

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review*, *18*, 46–60, http://dx.doi.org/10.3758/s13423-010-0034-0.

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: A step-by-step tutorial review. *Brain Topography*, *20*, 249–264, http://dx.doi.org/10.1007/s10548-008-0054-5.

Nee, D. E., Wager, T. D., & Jonides, J. (2007). Interference resolution: Insights from a meta-analysis of neuroimaging tasks. *Cognitive Affective & Behavioral Neuroscience*, *7*, 1–17.

Neuhaus, A. H., Trempler, N. R., Hahn, E., Luborzewski, A., Karl, C., Hahn, C., et al. (2010a). Evidence of specificity of a visual P3 amplitude modulation deficit in schizophrenia. *Schizophrenia Research*, *124*, 119–126, http://dx.doi.org/10.1016/j.schres.2010.08.014.

Neuhaus, A. H., Urbanek, C., Opgen-Rhein, C., Hahn, E., Ta, T. M. T., Koehler, S., et al. (2010b). Event-related potentials associated with Attention Network Test. *International Journal of Psychophysiology*, *76*, 72–79, http://dx.doi.org/10.1016/j.ijpsycho.2010.02.005.

Oberauer, K. (2005). Binding and inhibition in working memory: Individual and age differences in short-term recognition. *Journal of Experimental Psychology: General*, *134*, 368–387.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, *7*, 75–79, http://dx.doi.org/10.1038/nn1165.

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details. *Methods and Findings in Experimental and Clinical Pharmacology*, *24*, 5–12.

Perrig, W. J., Hollenstein, M., & Oelhafen, S. (2009). Can we improve fluid intelligence with training on working memory in persons with intellectual disabilities? *Journal of Cognitive Education and Psychology*, *8*, 148–164, http://dx.doi.org/10.1891/1945-8959.8.2.148.

Persson, J., & Reuter-Lorenz, P. A. (2008). Gaining control: Training executive function and far transfer of the ability to resolve interference. *Psychological Science*, *19*, 881–888.

Raven, J. C. (1990). *Advanced progressive matrices. Sets I, II*. Oxford: Oxford University Press.

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., et al. (2012). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *Journal of Experimental Psychology: General*, *142*, 359–379, http://dx.doi.org/10.1037/a0029082.

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004). Development of attentional networks in childhood. *Neuropsychologia*, *42*, 1029–1040, http://dx.doi.org/10.1016/j.neuropsychologia.2003.12.012.

Rueda, M. R., Posner, M. I., Rothbart, M. K., & Davis-Stober, C. P. (2004). Development of the time course for processing conflict: an event-related potentials

study with 4 year olds and adults. *BMC Neuroscience*, *5*, 39, http://dx.doi.org/10.1186/1471-2202-5-39.

Schmiedek, F., Li, S. C., & Lindenberger, U. (2009). Interference and facilitation in spatial working memory: Age-associated differences in lure effects in the *n*-back paradigm. *Psychology and Aging*, *24*, 203–210, http://dx.doi.org/10.1037/A0014685.

Schneiders, J. A., Opitz, B., Krick, C. M., & Mecklinger, A. (2011). Separating intra-modal and across-modal training effects in visual working memory: An fMRI investigation. *Cerebral Cortex*, *21*, 2555–2564, http://dx.doi.org/10.1093/cercor/bhr037.

Shipstead, Z., Hicks, K. L., & Engle, R. W. (2012). Cogmed Working Memory Training: Does the Evidence Support the Claims? *Journal of Applied Research in Memory and Cognition*, *1*, 185–193, http://dx.doi.org/10.1016/j.jarmac.2012.06.003.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica*, *50*, 245–276.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*, 628–654, http://dx.doi.org/10.1037/a0027473.22409508.

Silton, R. L., Heller, W., Towers, D. N., Engels, A. S., Spielberg, J. M., Edgar, J. C., et al. (2010). The time course of activity in dorsolateral prefrontal cortex and anterior cingulate cortex during top-down attentional control. *Neuroimage*, *50*, 1292–1302, http://dx.doi.org/10.1016/j.neuroimage.2009.12.061.

Studer-Luethi, B., Jaeggi, S. M., Buschkuehl, M., & Perrig, W. J. (2012). Influence of neuroticism and conscientiousness on working memory training outcome. *Personality and Individual Differences*, *53*, 44–49, http://dx.doi.org/10.1016/j.paid.2012.02.012.

Szmalec, A., Verbruggen, F., Vandierendonck, A., & Kemps, E. (2011). Control of interference during working memory updating. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 137–151, http://dx.doi.org/10.1037/a0020365.

Thorell, L. B., Lindqvist, S., Bergman, S., Bohlin, N. G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, *12*, 106–113.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, *17*, 635–654, http://dx.doi.org/10.1080/09658210902998047.

Walsh, B. J., Buonocore, M. H., Carter, C. S., & Mangun, G. R. (2011). Integrating conflict detection and attentional control mechanisms. *Journal of Cognitive Neuroscience*, *23*, 2211–2221.