

Disturbing visual working memory:  
Electrophysiological evidence for a role of prefrontal cortex in recovery from interference

Klaus Kessler<sup>1</sup> and Markus Kiefer<sup>2</sup>

<sup>1</sup>Department of Neurology, Duesseldorf University, Germany

<sup>2</sup>Department of Psychiatry, University of Ulm, Germany

Citation:

Kessler, K., & Kiefer, M. (2005). Disturbing visual working memory: electrophysiological evidence for a role of the prefrontal cortex in recovery from interference. *Cerebral Cortex*, 15(7), 1075-1087.

Corresponding author

Prof Klaus Kessler

School of Life and Health Sciences

Aston Brain Centre

Aston University

Aston Triangle, Birmingham B4 7ET

e-mail: k.kessler@aston.ac.uk

phone: +44 (0)1212043187

## Abstract

Single cell recordings in monkeys support the notion of lateral prefrontal cortex (PFC) controlling reactivation of visual working memory representations when rehearsal is disrupted. In contrast, recent fMRI findings yielded a double-dissociation for PFC and the medial temporal lobe (MTL) in a letter working memory task. PFC was engaged in interference protection during reactivation while MTL was prominently involved in the retrieval of the letter representations. We present event-related potential data (ERP) that support PFC involvement in the top-down control of reactivation during a visual working memory task with endogenously triggered recovery after visual interference. A differentiating view is proposed for the role of PFC in working memory with respect to endogenous/exogenous control and to stimulus type. General implications for binding and retention mechanisms are discussed.

Key words: Binding, ERP, medial temporal lobe, memory reactivation, occipito-temporal cortex, retention.

## Introduction

An important feature of biological information processing systems in general is that all operations are implemented in a massively parallel and distributed architecture, which works by means of highly complex cooperation among a huge number of neurons as simple processing elements. The question therefore arises of how chaotic cross-talk in such architecture can be avoided, yet allow for the activity of billions of neurons to lead to meaningful mental representations. Controlled processing in the human brain seems to be primarily achieved by a severely restricted number of high-level representations that can be maintained and processed in parallel. Storage and manipulation of information held in working memory is one prototypical example for the computational requirements of a situation, in which several chunks of information have to be maintained online, sometimes even in the presence of interfering stimuli and mental operations. It is a crucial research question in this field, how maintenance of information in working memory is still accomplished when interference impedes rehearsal mechanisms to keep information online. In the present research, these working memory mechanisms were investigated by means of event-related potentials (ERPs) in the EEG.

A growing body of evidence has been reported so far that supports the notion of working memory being implemented by prefrontal circuits in interaction with other cortical areas (see Duncan, 2001; Fuster, 1989; Goldman-Rakic, 1996; and Miller and Cohen, 2001, for reviews). Separate networks have been revealed for the so-called “phonological loop” and the “visuo-spatial sketch pad” (e.g. Gruber and von Cramon, 2003), which denote the fragmentation of working memory into a language-based and a visuo-spatial subsystem respectively (Baddeley, 1986). However, both physiological networks involve PFC (e.g. Gruber and von Cramon, 2003).

For visual pattern working memory (“visual cache”, cf. Logie, 1995) the interaction

between lateral prefrontal cortex (PFC) and inferior temporal cortex (IT) has proven to be crucial (e.g. Miller et al., 1996; Ranganath et al., 2004; Tomita et al., 1999). Similar to PFC neurons, IT neurons fire during the delay period of a working memory task, yet firing is disrupted by visual distractors (Miller et al. 1996). This is not the case for PFC neurons (Miller et al. 1996; Goldman-Rakic, 1996). What is more, PFC seems to be able to activate representations in IT as reported by Tomita et al. (1999). With specific lesions of the transcallosal interconnections between left and right IT of monkeys the authors showed that in a visual association paradigm a visual stimulus was nevertheless able to activate the contralateral IT, although information could not access IT via bottom-up or transcallosal connections. The authors conclude that information was transmitted by means of inter-hemispheric interactions within PFC and then by top-down intra-hemispheric projections to the corresponding neurons in IT. Support comes from the observation that PFC lesions disrupted performance and activation of contralateral IT. A possible conclusion regarding processing during external interference would be that PFC maintains information during interference and then re-activates information in IT when interference is over (e.g., Miller and Cohen, 2001). This would suggest that what persist in PFC are not representations per se but the “links” to where this information can be retrieved.

The essential role of PFC in maintenance and recovery has been recently questioned by research that showed an even stronger involvement of the medial temporal lobe (MTL) in the reactivation of representations in long- and short-term memory (Nyberg et al., 1996; Nyberg et al., 2000; Ranganath and D’Esposito, 2001; Sakai et al. 2002) whereas lateral PFC activation seemed rather unaffected by the need for, or by the success of reactivation (Buckner and Wheeler, 2001; Maril et al., 2001; Kikyo et al., 2002; Sakai et al. 2002). Sakai and Passingham (2004) propose – in concordance with previous findings (Gruber and von Cramon, 2003; Incisa della Rocchetta and Milner, 1993; Shimamura, 1995) - that PFC might be primarily involved in interference protection (for a discussion, see also Kiefer, Ahlegian,

and Spitzer, 2004). Accordingly, Sakai and Passingham (2004) report a double-dissociation between MTL and lateral PFC regarding reactivation of representations and interference protection in working memory. That is, lateral PFC showed stronger activation with letter distractors than with number distractors if letters had to be maintained in a specific order in working memory. This suggests higher activation of lateral PFC with higher interference during recall (the distractors were presented just before the memory probe was given). MTL, in contrast, was equally activated in both tasks. Rehearsal was equally disrupted with both kinds of distractors (MTL), whereas interference was higher with letter distractors (PFC). In a second study (Sakai and Passingham, 2004) interference was kept constant (number distractors only) but the need for reactivation was enhanced in half of the trials by arithmetical operations that had to be performed on the number distractors. Here, lateral PFC revealed no differential activation while MTL showed a stronger activation if recovery of representations became necessary (with arithmetic ops) as compared to the condition where rehearsal was not disrupted (no arithmetic ops). The authors conclude that this double-dissociation shows that MTL mediates reactivation from long- and short-term memory after disrupted rehearsal while lateral PFC assures protection from interference.

However, it is premature to refute a role of PFC for re-activating working memory representations on the basis of the Sakai and Passingham findings for several reasons: First, only material was used that could be encoded into the phonological loop (a specific sequence of letters). This leaves the question unresolved as to which extent the reported findings would generalize to visual working memory. Gruber and von Cramon (2003), for example, found modality-specific interference effects for visuo-spatial and verbal working memory (see also Zurovski et al., 2002; Smith et al., 1995). In fact, the above mentioned studies by Miller et al. (1996) and Tomita et al. (1999) suggest a role for PFC during re-activation of IT during visual tasks. Second and closely related to the first issue, letters are highly over-learned stimuli that can be easily matched to long-term concepts and, hence, can be easily encoded into, and

retrieved from associative memory (e.g. Chalmers and Humphreys, 2003; Guttentag, 1995; Westerman, 2001), prominently mediated by hippocampal and parahippocampal structures (e.g. Martin, 1999; Ranganath and D'Esposito, 2001; Ranganath et al. 2004; Schacter and Wagner, 1999; Stark, Bayley and Squire, 2002; Sutherland and Rudy, 1988). This might be because letters achieve deeper processing than purely perceptual features as suggested by Craik and Lockhart's levels of processing framework (1972; Craik and Tulving, 1975; also Cleary, 2002; Hyde & Jenkins, 1969). It is possible that within a visual task, where low-level perceptual information has to be maintained, recovery of working memory representations depends more on PFC activity.

Third, and most importantly, in the Sakai and Passingham studies interference was administered just before the memory probe was displayed. In such a procedure, recovery of representations could be strongly biased by the displayed probe. The probe provides a contextual cue that might directly access associative memory of the encoded letters without necessarily engaging lateral PFC (see also the concept of long-term working memory by Ericsson and Kintsch, 1995). Alternatively, activation of PFC related to recovery and not to interference protection might be earlier, but very transient and subtle, hence, hard to trace by fMRI, a method with poor temporal resolution. It would therefore be necessary to incorporate a longer gap between end of interference and probe. With such a procedure we predict that PFC would trigger recovery of information as without an external retrieval cue information in associative memory would have to be accessed via traces in lateral PFC. This is in concordance with recent findings by Ranganath et al. (2004) who showed stronger involvement of lateral PFC in working memory and stronger involvement of hippocampus in associative learning within the same study that employed faces and houses as items. Concluding, we are postulating an account where the involvement of lateral PFC in maintenance and recovery increases, the less is known about the items that have to be encoded and, most importantly, the more retrieval is self-initiated.

In the present study, we investigated the mechanisms underlying the recovery of visual working memory representations from interference by means of EEG. ERPs were used to track the time course of cortical activation with high temporal resolution in a visual working memory task. This was, however, traded for spatial resolution, especially regarding hippocampal regions. Our research therefore focused on the temporal orchestration of lateral prefrontal and occipito-temporal activity. We designed a task that would allow testing for PFC involvement in active recovery without external retrieval cue after interference. We adapted a paradigm initially developed by Mecklinger (1999), where geometrical shapes have to be encoded, but where the probe at the end of the trial was varied in size and not in identity. This was to prevent participants to encode the shapes solely via a linguistic code into the phonological loop. Instead, the participants were forced to maintain low-level perceptual information about visual details of the shapes to be able to solve the task. Yet, the availability of a cognitive concept and a linguistic label for each geometrical shape would increase the stability of the representations (Miyake & Shah, 1999, for a review) in the visuo-spatial sketch-pad, the capacity of which is known to be highly dependent on the complexity and the similarity of the items<sup>1</sup>.

Difficulty of the working memory task was varied on two dimensions: working memory load and interference. Working memory load was manipulated by presenting two or four shapes to be encoded. Load should increase the effort of maintenance (cf. Baddeley, 1986), but not the need for recovery (cf. Sakai and Passingham, 2004). The need for recovery was implemented in our experiment by administering an interference task within the retention interval. A simple visual discrimination task was employed, which should be sufficient to induce low-level interference with the encoded information about shape size, hence, disrupt sustained firing. Occipito-temporal cortex has been shown to be involved in early visual

---

<sup>1</sup> Logie (1995) reviews findings that show capacity of only one item with complex and very similar square matrix patterns. Broadbent and Broadbent (1981), however, report capacity of three items with more dissimilar wallpaper textures.

stimulus discrimination (see Hopf et al, 2002 for a review). It is therefore a likely candidate for a module where shape integration is accomplished to an extent so that a match with simple target-templates can be performed. We therefore expected occipito-temporal cortex to be the site where maintenance of task-relevant size information is strongly threatened by interference (cf. Miller et al. 1996).

External interference ended 2 sec before the memory probe was administered. This time point marked the beginning of the recovery interval, which lasted until onset of the probe presentation. During the recovery interval, subjects were released from interference and were instructed to prepare for the probe, thus allowing for endogenously controlled reactivation. After release from interference, PFC should therefore trigger recovery in occipito-temporal cortex. However, as compared to representations, where rehearsal was undisturbed, recovered representations are expected to be weaker (less rich), as only task-relevant information could be incorporated into PFC representations. In turn, the system would have to engage in a higher effort to maintain these representations. Therefore, we expected qualitative differences during the recovery interval at occipito-temporal sites, i.e. at occipito-temporal electrodes, depending on the administration of interference.

Concluding, ERP differences between trials without interference and trials with interference at lateral prefrontal electrodes should be fast and transient since they should reflect the initiation of a recovery process in the latter case. Subsequently, strong and stable ERP differences at occipito-temporal electrodes should emerge that would reflect the qualitative differences between recovered, effortfully maintained representations (after interference) compared to undisturbed, rich representations. We also expected an interaction between load and interference. According to the account we are postulating, prefrontal involvement should reflect to some extent general cognitive effort, namely interference (“with” vs. “without”) and load (“high” vs. “low”), while posterior effects (temporo-occipital) should primarily reflect the qualitative differences between recovered (with interference) and



maintained (without interference) representations. This latter prediction is based on the assumption (in concordance with Sakai and Passingham, 2004) that enhanced load would increase the effort of maintenance but not the need for re-activation as rehearsal is not disrupted. Interference, in contrast, disrupts rehearsal, and recovery becomes necessary. Especially at occipito-temporal sites, recovered representations should qualitatively differ from continuously maintained (i.e., rehearsed) representations. As high load does not disrupt rehearsal, representations in occipito-temporal brain regions should not differ qualitatively between “high” and “low” load and, hence, occipito-temporal ERP effects of load should be smaller than effects of interference.

## Methods

### Participants, experimental stimuli and apparatus

9 female and 6 male students of Ulm University with a mean age of 25.3 were employed as subjects and received payment for their participation in the experiment. All were right-handed and had normal or corrected to normal vision. The study has been approved by the local Ethical committee and is in accordance with the declaration of Helsinki.

The stimuli in this experiment were basic geometrical or symbolic shapes as shown in Figure 1 (Panel A). The experiment was performed on an IBM-compatible personal computer and each trial in the experiment was divided in four phases (Figure 2): First, “encoding” of shapes, second, “retention” with either a blank screen or with two visual discrimination tasks, third, 2 sec “recovery” of representations until the probe (fourth phase) appeared on the screen.

### Figure1 about here

Encoding displays could consist either of two or four (unique) shapes in order to vary

the load imposed onto the participants' working memory. The position of each shape was randomized across a 4x4 matrix of possible positions on the screen that covered an area of 6 cm width ( $3.8^\circ$  visual angle at a viewing distance of 90 cm) and of 4 cm height ( $2.5^\circ$  visual angle). The over-all size of the relevant display was kept as small as possible to reduce the need for eye-movements. The geometrical shapes were 1 cm wide and 0.75 cm high in their basic form. As pointed out, the experiment was designed to force the participants to maintain perceptual information of the stimuli. Therefore, either the width or the height of a stimulus could vary between prime and probe by 25% (Figure 1, Panel B). Thus, we expected that although all shapes could be named, the task could not be solved by phonological encoding alone. Instead, maintenance of visual information about the actual width or height would be indispensable to solve the task, while the long-term concept associated with each figure could help to efficiently generate a working memory representation. All probes shared the concept with one of the primes, but their size varied on 50% of the trials (25% height and 25% width changes).

Figure 2 about here

As described, prime and probe were separated by a 5.1 sec ISI (interstimulus interval). During the 3.1 sec of the "retention" interval either a blank screen with a gray fixation cross of  $0.64^\circ$  visual angle was shown in the middle of the display or either the subject had to solve two visual discrimination tasks (Figure 1, Panel C). The latter condition was used to induce interference in visual areas that would also be involved in the processing of the encoded geometrical shapes. Each discrimination task consisted of a black random shape that was either "edged" or "round" embedded into another random, but gray shape (again, "edged" or "round"). The subjects had to decide in a "go"/"no-go" fashion if the inner black form was "round". The size of each double-shape was  $1.9^\circ$  horizontal and vertical visual angle. Finally,

during the last 2 sec before the onset of the probe a blue fixation cross was displayed in all conditions that was of the same size as the gray cross during the baseline “retention” condition. The blue fixation cross signaled subjects the recovery interval.

### Procedure and Design

Participants sat before the monitor at 90 cm viewing distance, with a button box in front of them, under dim lighting conditions. Each participant completed a practice session of 15 trials. Testing of the subsequent 320 experimental trials lasted approximately 70 minutes. Subjects had a self-paced break after the first half of the trials. The timing within a trial is shown in Figure 2. After self-initiating each trial, a gray fixation cross was displayed for 1 sec. Next, two (50%) or four (50%) shapes were displayed for encoding. The subjects had been instructed to memorize all shapes within 4 sec. In the following 3.1 sec “retention” interval in half of the trials the display just turned black with a gray fixation cross in the middle. In the other half of the trials two of the described discrimination tasks had to be solved within 1500 ms each. Next, in all conditions the display turned black and a blue fixation cross appeared in the center of the screen exactly 3100 ms after the start of the retention phase to indicate the onset of the recovery interval. Subjects had been instructed that 2 sec were left until the probe would be displayed. They were also told to retrieve the encoded information in preparation for the probe display. This recovery interval was of major importance to our investigation because recovery processes after interference would possibly become observable without any concurrent external stimulation. Finally, the probe appeared either in the same or in a different location as the corresponding prime and could be of same or of different size as the prime. Subjects were instructed to ignore changes in position<sup>2</sup> but to report changes in size by pressing a key for “different size” with the right index finger or a

---

<sup>2</sup> The position change was introduced to investigate position-identity binding in a follow-up study with exactly the same stimuli. Only by changing the instruction (“Is the probe in the same location as the prime?”) - while keeping the stimuli and the procedure absolutely identical - it would be possible to test recovery of position-identity representations after interference. However, this was not incorporated in the experiment presented here.

second key respectively for “same size” with the right middle finger. The instruction strongly emphasized to react as accurately and as quickly as possible. Participants could therefore make two kinds of errors: falsely reporting same size or falsely reporting different size. Concluding, we employed a 2x2x2 experimental design that included the factors “load” (low vs. high), “interference” (with vs. without), and “probe size” (“same” vs. “different”) with 40 trials in each condition. As dependent behavioral measures, percentage of correct reports and reaction times were calculated individually for each condition.

## EEG

Scalp potentials were collected using a equidistant montage of 64 sintered Ag/AgCl-electrodes mounted in an elastic cap (Easy Cap, Falk Minow Systems). An electrode between Fpz and Fz was connected to the ground, and an electrode between Cz and FCz was used as recording reference. Eye movements were monitored with supra- and infra-orbital electrodes and with electrodes on the external canthi. Electrode impedance was kept below 5 k $\Omega$ . Electrical signals were amplified with Synamps amplifiers (low-pass filter = 70 Hz, 24 dB/octave attenuation; 50 Hz notch filter) and continuously recorded (digitization rate = 250 Hz), digitally band-pass filtered (high cut-off: 16 Hz, 24 dB/octave attenuation; low cut-off: 0.2 Hz, 12 dB/octave attenuation) and segmented (200 ms before to 9100 ms after the onset of the encoding displays). EEG data were corrected to a 200 ms baseline prior to the onset of the encoding displays. ERPs were extracted by averaging artifact-free EEG segments synchronous to the onset of the stimulus separately for each experimental condition. In order to obtain a reference independent estimation of scalp voltage, the average-reference transformation was applied to the ERP data (Bertrand, Perrin and Pernier, 1985; Kiefer, Marzinzik, Weisbrod, Scherg and Spitzer, 1998a).

The aim of our investigation was primarily focused on the recovery interval (7100 - 9100 ms after the onset of the encoding display) in order to observe recovery processes after

release from interference. However, we could not ignore the rest of the trial because of two reasons. First, artifact rejection had to be performed throughout the trial length to ensure good data quality for all intervals. Second, it was important to calculate the baseline correction before the actual begin of the trial and not preceding the recovery interval (cf. Figure 3), because we expected processes during retention to differ substantially among conditions (i.e. “with” vs. “without” interference). Accordingly, different levels of activation preceding the recovery interval might reflect true differences in cognitive processing. Setting these differences to zero might therefore obliterate early cognitive differences yet causing artificial effects to emerge later in the recovery interval. For that reason, the baseline correction was performed on the 200 ms interval preceding the onset of each encoding display, i.e. 7300 ms earlier than the start of the recovery interval (Figure 3).

Figure 3 about here

Due to the limited amount of available trials data were averaged across false and correct reports in order to secure a sufficient signal-to-noise ratio (minimum of 30 trials per condition). Note that the analysis was restricted to the recovery interval. Hence, the probe was not yet visible and could not initiate matching and decision processes. We therefore expected similar processes to occur regardless of the behavioral outcome, although processes might differ in strength and efficiency depending on (subsequent) correct and false reports. In the worst case, however, condition effects were more likely to get deteriorated with this strategy. That is because frontally-driven recovery processes should be best to observe on trials where they are most efficiently engaged, i.e., where they lead to correct results. Therefore, by including incorrect trials as well, frontal effects (and their occipito-temporal consequences) are more likely to get attenuated than artificially enhanced.

Figure 4 about here

We were especially interested in the different processing between anterior and posterior parts of the human brain. For that reason, we selected for statistical analysis four bilateral pairs of electrodes over frontal (AF7/AF8, F9/F10, FT7/FT8, T1/T2) and occipito-temporal cortex (P9/P10, P7/P8, O9/O10, O1/O2) as shown in Figure 4. ERP effects at corresponding frontal electrodes have been related previously to activity in lateral prefrontal cortices using dipole source analysis (Snyder et al., 1995; Kiefer et al., 1998a; Posner et al., 1988). Corresponding occipito-temporal electrodes have been employed by Ruchkin et al. (1992, 1995) who found negative slow wave effects during retention intervals. Although the localizational value of ERPs has to be viewed with caution (Nunez, 1981), the topography of such occipito-temporal effects is well in line with reports from animal neurophysiology demonstrating activity in occipito-temporal cortex during visual working memory tasks. Scalp voltages were submitted to four-way repeated-measures ANOVAs with the within-subjects factors “hemisphere” (left/ right), “horizontal alignment” (frontal/ occipito-temporal), “load” (low/ high) and “interference” (yes/ no).

Results

Behavioral Results

Performance accuracy (percentage of correct reports) was calculated separately for all conditions. Accuracy and reaction times were then included into two separate 2x2x2 ANOVAs with the factors “load” (low/ high), “interference” (yes/ no), and “probe size” (“same”/ “different”). The ANOVA for accuracy yielded three significant main effects (all means and standard deviations for reaction times and accuracy are provided in Table 1). None of the interactions reached significance ( $F(1, 14) < 2.023$ ;  $p > 0.177$ ). It was harder for the participants to correctly judge probes of a different size than the prime. If the probe was of the

same size as the prime, significantly more correct reactions occurred ( $F(1, 14) = 20.605$ ;  $p < 0.0005$ ). This general tendency to report more often that prime and probe were of the same size could reflect a response bias induced by the common shape for prime and probe (cf. Figure 1, Panel B). Furthermore, as expected, it was significantly easier to respond correctly if memory load was low ( $F(1, 14) = 23.706$ ;  $p < 0.0002$ ) and, most importantly, if no interfering tasks had to be performed during retention ( $F(14) = 5.7024$ ;  $p < 0.042$ ).

Table 1 about here

The ANOVA for reaction times was calculated for correct responses only and revealed exactly the same three significant main effects as the accuracy analysis. Participants responded faster to probes of the same size as the prime ( $F(1, 14) = 7.758$ ;  $p < 0.015$ ) replicating the bias observed with the accuracy data. In addition, participants were significantly faster with low memory load ( $F(1, 14) = 24.124$ ;  $p < 0.0002$ ), and without interference during retention ( $F(1, 14) = 6.546$ ;  $p < 0.023$ ). Hence, the reaction time results exactly mirror the accuracy results in that cognitive effort generally attenuates performance.

ERP Results

The goal of this study was to investigate whether lateral prefrontal cortex is involved in the recovery of visual working memory representations after interference. We therefore expected early ERP differences between trials with vs. without interference at frontal electrodes and subsequent sustained effects at occipito-temporal electrodes that would reflect qualitative differences between maintained (without interference) vs. recovered representations (after release from interference). We aimed at answering our research questions by applying an analysis that started with a very coarse temporal resolution (500 ms averaged per electrode cluster) which was then more refined (sample by sample t-tests per

electrode for the first 1000 ms of the recovery period). As described, distinct recovery sub-processes are thought to be quite different with respect to their time scale: An initial top-down trigger of recovery might be very transient and best to identify with a high temporal resolution, while top-down maintenance-control as well as maintenance of recovered perceptual features might occur later and lead to longer-lasting effects – best to be captured with a coarse temporal resolution.

Figure 5 about here

Figure 5 shows the waveforms of all 16 electrodes in the four clusters (bilateral frontal and occipito-temporal clusters). As can be seen, the major expectations were met: At frontal electrodes interference causes stronger negative and stronger positive evoked responses, thus reflecting fast and transient as well as slower but stable effects. At occipito-temporal electrodes early negative components are stronger without preceding interference. These are, however, followed by stronger sustained negativity in the interference conditions, hence suggesting transient as well as stable differences in occipito-temporal cortex. This first impression is supported by our statistical analysis.

*Low Temporal Resolution*

Analysis with a coarse temporal resolution was based on mean voltages in four 500 ms time windows in the recovery interval (2000 ms length in total) starting with the onset of the blue fixation cross. Separate ANOVAs were performed for each time window with the factors “hemisphere” (left/ right), “horizontal alignment” (frontal/ occipito-temporal), “load” (low/ high) and “interference” (with/ without). Signals were averaged across all four electrode sites within each cluster. This analysis revealed a significant interaction between “interference” (with/ without) and “horizontal alignment” (frontal/ occipito-temporal) as shown in Figure 6



(Panels A, C). This effect was significant in two of the four time windows ( $F(1, 14) > 4.64$ ;  $p < .049$ ) and marginally significant in a third window. In the first time window (0-500 ms) both interference conditions (with vs. without) showed a significant difference between frontal and occipito-temporal sites ( $F(1, 14) > 8.52$ ;  $p < .011$ ). This difference was, however, even stronger for the condition without interference, which is suggested by the simple effects at each site: At occipito-temporal electrodes a significantly stronger negative component was observed without interference ( $F(1, 14) = 9.36$ ;  $p < .008$ ) while at frontal sites a numerically stronger negative component was observed with interference. In the second time window (500-1000 ms) this interaction did not reach significance, possibly because of a change in processing stages, as can be seen in Figure 5 by the crossing of the waveforms in the interval between 600 and 800 ms at occipito-temporal (earlier) and frontal (later) electrodes. In the third time window (1000-1500 ms) the interaction between “interference” (yes/ no) and “horizontal alignment” (frontal/ occipito-temporal) reached significance again ( $F(1, 14) = 7.14$ ;  $p < .018$ ), while showing a reversed picture. Conform to Figure 5, the condition with interference showed the stronger effects at both sites: ERPs in trials with interference were significantly more negative at occipito-temporal ( $F(1, 14) = 13.9$ ;  $p < .002$ ) and numerically more positive at frontal sites than the “no interference” condition (Figure 6, Panel C). The same picture emerged in the third time window were the interaction almost reached significance ( $F(1, 14) = 4.47$ ;  $p < .053$ ).

Figure 6 about here

The only other ANOVA effect that reached significance ( $F(1, 14) = 6.06$ ;  $p < .027$ ) in one of the four 500 ms windows was the interaction between “interference” (yes/ no), “horizontal alignment” (frontal/ occipito-temporal), and “load” (low/ high) shown in Figure 6 (Panel B). This 3way interaction reached significance in the first time window (7100-7600

ms). Basically, “load” adds the following fact to the previously discussed interaction between “interference” and “horizontal alignment”: This interaction was even more pronounced for the “low load” condition. The condition “high load without interference” yielded a similar pattern as the two interference conditions (low and high load), while “low load without interference” showed a more pronounced difference to all the other three conditions that reached significance at occipito-temporal sites ( $F(1, 14) > 16.96$ ;  $p < .001$ ).

### *High Temporal Resolution*

To better understand the fast and transient processes at the beginning of the recovery interval continuous t-tests (“without” vs. “with” interference) were performed per electrode for each and every sample of the first second of the recovery interval. In this analysis we were primarily interested in where and when differences between the conditions with interference versus without interference would emerge. Figure 7 shows continuous t-tests (without interference vs. with interference and high vs. low load respectively) for two pairs of frontal/occipito-temporal electrodes (AF7/O1 and AF8/O2 respectively). Note that the t-tests compared the condition without interference to the condition with interference (the direction of the comparison strongly affects the valence of the resulting t-values), hence, a positive t-test value could either indicate a stronger negative or a weaker positive ERP component in the condition with interference. Corresponding reflections apply to negative t-test values. To resolve this ambiguity the time courses of activation in each condition have to be considered (see Figure 5).

### Figure 7 about here

Significant ( $p < .05$ ) t-test differences (as indicated by the dotted horizontal lines in Figure 7) between the interference conditions within the first 400 ms of the recovery interval

were observed at frontal as well as occipito-temporal electrodes: Positive t-test differences at frontal and negative t-test differences at occipito-temporal sites respectively. This is due to stronger negative ERP responses being observed after interference at frontal electrodes while stronger negative ERP responses are revealed at occipito-temporal electrodes without preceding interference (compare Figure 5). However, the early differences at frontal sites seem to be bi-phasic in nature in that an even earlier stronger negative t-test difference between the interference conditions (in Figure 7 only significant for electrode AF8 in the high load condition) is immediately followed by the discussed significant positive t-test difference. Two further frontal electrodes (T1, T2) replicate this effect by yielding a significant negative t-test difference around 200 ms (Figure 8). This early negative difference corresponds to the first positive ERP component in the recovery interval as shown in Figure 5 (first black arrow around 200 ms), which is faster after interference.

Figure 8 about here

These differences clearly show that at a very early processing stage (< 600 ms) lateral PFC is activated stronger in the conditions with preceding interference while temporo-occipital cortex is more involved in the conditions without preceding interference during this stage (conform to the ANOVA results of the 1<sup>st</sup> time window reported in the previous section). It takes as long as 600-800 ms until stronger sustained negativity in occipito-temporal cortex emerges in the conditions where interference was administered during retention (conform to the ANOVA results of the 3<sup>rd</sup> time window reported in the previous section). Concluding, significantly stronger cortical responses for the interference conditions are observed very early at frontal electrodes while stronger occipito-temporal responses in these conditions emerge only after 600-800 ms but then remain stable until response.

## Discussion

The present study was set up in order to investigate the mechanisms underlying recovery of visual working memory representations after release from interference. Within a visual working memory task, ERPs were used to track the time course of activity in prefrontal and occipito-temporal regions within a recovery interval, during which subjects were instructed to prepare themselves for the response to a probe display. The behavioral results show that our experimental manipulation of interference and load was successful: Error rates as well as reaction times support the notion that enhanced cognitive effort (high load, interference) deteriorates performance. In addition, participants showed an effect in both dependent measures to process probes of the same size as the primes more efficiently. Possibly, subjects were biased by the fact that the prime and the probe shared the same long-term memory concept.

### Working memory and ERPs

Although the electrophysiological correlates of recovery of visual working memory representations have never been investigated before, our results are compatible with earlier ERP studies on retention in visual working memory. Ruchkin et al. (1992) found negative slow waves during the retention interval over occipital and temporal cortex. As also observed in the present study, these posterior negative ERPs increased as a function of cognitive effort (for similar slow wave effects during episodic memory encoding see Heil, Rösler and Hennighausen, 1996; Mecklinger, 1999; Rösler, Heil and Hennighausen, 1995). Moreover, Ruchkin, Canoune, Johnson and Ritter (1995) showed that this occipito-temporal negative ERP deflection does in fact relate to working memory operations and not simply to general preparatory processes. Although the localizational value of ERPs has to be viewed with caution (Nunez, 1981), the topography of these occipito-temporal effects is well in line with

reports from animal neurophysiology demonstrating activity in occipito-temporal cortex during visual working memory tasks. The topography of the frontal ERP effects observed in the present study is similar to that observed in other tasks involving prefrontal cortex such as semantic retrieval (Kiefer et al., 1998b; Snyder et al., 1995) and executive function (Kiefer et al., 1998a; Weisbrod et al., 2000). These frontal ERP effects have been related to activity in lateral prefrontal cortices using dipole source analysis (Snyder et al., 1995; Kiefer et al., 1998a; Posner et al., 1988).

### Two processing stages

Our ERP results suggest two major processing stages during the recovery of working memory representations that are highly compatible with our expectations. During the first 600 ms of the recovery interval, the condition without interference shows stronger responses at occipito-temporal sites while the condition with interference shows stronger responses at frontal sites. The picture is different for the last 1000 ms of the recovery interval: Here, the condition with interference shows enhanced responses at both sites – stronger positivity at frontal and stronger sustained negativity at occipito-temporal sites. Furthermore, “load” only plays a statistically significant role during the first 500 ms of the recovery interval. Hence, two distinct processing stages seem to be separated in time by an interval of transition (600-800 ms).

### The first processing stage

The first stage (0-600 ms) is characterized by stronger negativity at occipito-temporal sites for the conditions without interference (Figures 5, 6, 7). In contrast, at frontal sites the conditions with interference show stronger responses (Figures 5, 6, 8): An initial dip that reflects an early and fast more positive response and a subsequent maximum that reflects a later but stable, stronger negative response. Taken all together this suggests a primarily

bottom-up driven process in the conditions without interference (stronger occipito-temporal responses), yet a more top-down driven process in the conditions with interference (stronger frontal responses). Furthermore, it is also of great importance in this context that the condition without interference but with high working memory load is quite similar to the conditions with interference, while the condition without interference but low load is very different from all the other conditions - especially at posterior sites. This points to a general modulation by the amount of cognitive effort at this early processing stage. Increasing the effort – either by interference or by load - shifts the equilibrium from occipito-temporal towards frontal processing resources. Concluding, the observed cortical responses support the hypothesis that prefrontal areas control maintenance and recovery of posterior representations under conditions of high cognitive effort – especially after posterior representations have been disturbed by new perceptual input. However, recovery could be accompanied by a certain loss, as indicated by an attenuated early negative component in the interference conditions at occipito-temporal electrodes (Figures 5, 6, 7). The data pattern is congruent with the notion of PFC providing a representation in form of “links” to where to retrieve the encoded information. Such a representation would allow for a recovery process initiated by PFC and has important implications on binding and retention mechanisms.

### The second processing stage

The second processing stage is primarily reflected by highly stable occipito-temporal differences during the last 1000 ms of the recovery interval, where the interference conditions show stronger sustained negativity than the conditions without interference (Figures 5, 6, 7). This suggests that recovered representations after interference might differ qualitatively from undisturbed representations, because of a sustained effort to maintain the hardly recovered information. This is supported by stronger positive responses at frontal sites in the conditions with interference (Figures 5, 6, 8), which can be interpreted as top-down support for the

maintenance of fragile recovered representations during this second processing stage. Taken altogether this is strong evidence for a recovery process of representations in occipito-temporal areas that is triggered by early PFC activation in the first processing phase. Hence, these data are in concordance with the notion that PFC provides information about where to retrieve information after disrupted rehearsal and has strong implications on mechanisms of binding and retention that will be discussed shortly.

### The role of PFC in working memory processing

In sum, the major findings from both processing stages reveal a predominance of early prefrontal as well as of later occipito-temporal responses after interference. We interpret these findings in terms of a recovery process initiated by lateral prefrontal representations after disrupted rehearsal, which, subsequently, leads to strong differences at occipito-temporal electrodes that reflect the qualitative differences between recovered, effortfully maintained representations (after interference) compared to undisturbed, automatically rehearsed representations (without interference).

The present pattern of results clearly speaks in favor of lateral PFC being involved in the endogenous control of recovery of visual working memory representations and not only in interference protection. Our data are in line with the above described findings by Miller et al. (1996) and by Tomita et al (1999) showing that prefrontal cortex might reactivate representations in inferior temporal cortex. Accordingly, PFC representations might serve as “links” to where information could be retrieved after interference. This implies quite complex representational structures and mechanisms that subserve reactivation of information in working memory, which go beyond “simply” binding the correct information together in order to generate a coherent mental representation.

### Implications for computational mechanisms of binding and retention in working memory

With respect to binding, Hommel (2004) suggests that three distinct mechanisms can be derived from the existing literature: Synchronization, convergence and indexing. There is a significant body of evidence showing that synchronized oscillatory activity is a basic mechanism for communication among wide-spread areas in the cortex. Singer and colleagues (e.g. Engel et al., 1991; Gray and Singer, 1989) were able to show that synchronous gamma band oscillations subserve binding-processes in visual cortex. Since this seminal work, oscillatory activity in the theta, alpha and beta band has been related to high-level attentional processes (Varela et al., 2001), to consolidation of representations (Gross et al., 2004) and also to (undisturbed) retention of representations in working memory (e.g. Sarnthein et al., 1998; Weiss and Rappelsberger, 2000). From a computational point of view binding by synchronization is a very simple mechanism as no additional binding-structures have to be incorporated into the architecture of a model (e.g. Goebel, 1993; Lisman and Idiart, 1995; von der Marlsburg, 1999).

However, such a simple binding mechanism could not explain retention of working memory representations while rehearsal is disrupted by interfering information. If synchronization alone would hold a representation together single elements would be lost without any chance for recovery in case they get de-synchronized or de-activated. Obviously then, representations in IT could not be reactivated after neural firing was disrupted by distractors (Miller et al., 1996), and performance would drop to chance level. In contrast, a significant body of evidence has been discussed in the present report that suggests that recovery of representations is possible after disrupted rehearsal. “Convergence” as a binding mechanism also fails to account for recovery of representations as convergence is per definition a “one way” process. Activated neurons at a lower level converge onto neurons at a higher level, but cannot be traced back by neurons from the higher level. Thus, a robust mechanism for retention that may explain preserved performance after interference still needs to be found.



“Indexing” refers to a binding mechanism that is based on explicit links, which temporarily connect the correct elements together (e.g. Dehaene, Sergent and Changeux, 2003; De Vries and Dalenoort, 1995; Kessler and Rickheit, 1999, 2000; Smolensky, 1988). Although computationally more complex than synchronization and convergence, temporary links offer an amount of robustness and flexibility of representations that allows for recovery after disrupted rehearsal. Temporary links allow to trace back each element by its temporary connections to other elements: If an element is deactivated, for example as a result of external interference, then it may be reactivated by means of the temporary links. To avoid an exponentially exploding number of temporary links, it is plausible to assume that only task-relevant information might get interconnected. Accordingly, recovered representations would be qualitatively less rich than maintained representations, which is congruent to the observed long-lasting occipito-temporal differences in the present report. Finally, temporary links can be regarded as a sort of “pointer-representation” in the sense that what persists in working memory is not information per se but the links to where this information can be found (e.g. Duncan, 2001; McClelland, McNaughton and O’Reilly, 1995; Teyler and Discenna, 1986).

We propose that all three kinds of binding mechanisms (synchronization, convergence, indexing) might be employed simultaneously by the human brain to achieve “coincidence” of all relevant pieces of information that would allow for the formation of a coherent mental representation. However, we also propose that some sort of indexing is indispensable for robust maintenance if rehearsal is disrupted. In fact, recent physiologically plausible models already incorporate most of these aspects (e.g. Dehaene et al., 2003; also to some extent DeVries and Dalenoort, 1995). Hence, it might depend on the requirements of the situation, which aspect bears more relevance: In a situation where synchronized patterns can be maintained without external disruption, a pointer-representation might not be necessarily engaged. In a situation with interfering events, however, temporary links would allow to ensure maintenance by recovering task-relevant information.

### Neural substrates for “pointer-representations”

In contrast to synchronization, it is somewhat harder to define how temporary links would manifest themselves in neurophysiological measures. It seems plausible that specific brain areas might work as a pointer-representation by means of long-range connections to other areas that code perceptual information. These pointer-areas would remain active during interference and would be able to re-activate patterns in more posterior areas. Therefore, prefrontal cortex is a likely candidate for such representations as suggested by the described findings from Tomita et al. (1999), Miller et al. (1996), Goldman-Rakic (1996) and also by the findings in the present report. Alternatively, others have proposed that hippocampus and related structures might work as pointer-representations in long-term yet also in short-term memory (McClelland, McNaughton and O'Reilly, 1995; Teyler and Discenna, 1986). The Sakai and Passingham (2004) results described above support this notion by demonstrating stronger involvement of MTL during recovery by means of fMRI.

As fMRI and EEG are complementary methods with respect to spatial and temporal resolution the differences between the Sakai and Passingham findings and the ERP results described in the present report have to be considered with care. Scalp ERPs might not be suitable to detect activity in deep brain structures such as the MTL whereas fMRI might be not sensitive enough to capture transient PFC activity. Nevertheless it is also possible that PFC and MTL play different roles in the recovery of working memory representations after interference. In concordance with Sakai and Passingham (2004), who also suggest a functional dissociation between MTL and PFC with respect to automatic vs. controlled processing, we propose that the MTL supports recovery of working memory representations, which is automatically and exogenously triggered by a retrieval cue (e.g., the probe display). PFC, in contrast, might be involved in controlled, endogenously initiated recovery. PFC has been repeatedly shown to be one of the major neural substrates for cognitive control in

general and for working memory control in particular (see Duncan, 2001; Fuster, 1989; Goldman-Rakic, 1996; Miller and Cohen, 2001, for reviews). On the other hand associative memory has been related to MTL (e.g. Ranganath et al. 2004; Schacter and Wagner, 1999; Sakai et al. 2002; Stark, Bayley and Squire, 2002; Sutherland and Rudy, 1988) and retrieval seems to be triggered quite automatically by an external cue (e.g. Kessler and Tipper, 2004; Martin, 1999; Sutherland and McNaughton, 2000).

As described above, the type of stimuli (e.g. letters vs. shapes) might also play a role in addition to endogenous/exogenous control in mediating between different memory-subsystems. Concluding, pointer-representations in PFC might be primarily engaged in tasks where no long-term knowledge is available for the task-relevant information and where recovery is triggered endogenously without any external retrieval cue, whereas hippocampus and related structures might be primarily engaged if long-term knowledge allows efficient associative encoding and if external retrieval cues provide direct (automatic) access to these representations. Future work should investigate these factors in more detail by equally employing methods with good spatial (e.g. fMRI, PET) and with good temporal (e.g. EEG, MEG) resolution.

### **Acknowledgements**

K. K. was funded by the German Research Community (CRC 360, project C1). This research was also supported by a grant of the German Research Community (DFG KI 804/1-1) to M.K. We would like to thank two anonymous reviewers for their helpful comments on an earlier version of this manuscript.

Please address correspondence to Klaus Kessler, University of Duesseldorf, Neurology, MEG Lab, Moorenstr. 5, Duesseldorf 40225, Germany. E-mail: klaus.kessler@med.uni-duesseldorf.de

## References

- Bertrand O, Perrin F, Pernier JA (1985) A theoretical justification of the average reference in topographic evoked potential studies. *Electroencephalography and clinical Neurophysiology* 62: 462-464.
- Baddeley A (1986) *Working Memory*. Oxford: University Press.
- Broadbent DE, Broadbent MHP (1981) Recency effects in visual memory. *Quarterly Journal of Experimental Psychology* 33A: 1-15.
- Buckner RL, Wheeler ME (2001) The cognitive neuroscience of remembering. *Nat Rev Neurosci* 2: 624-634
- Cauquil AS, Edmonds GE, Taylor MJ (2000) Is the face-sensitive N170 the only ERP not affected by selective attention? *Neuroreport*, 11: 2167-2171.
- Chalmers KA, Humphreys MS (2003) Experimental manipulation of prior experience: Effects on item and associative recognition. *Memory* 11: 233-46.
- Cleary AM (2002) Recognition with and without identification: dissociative effects of meaningful encoding. *Mem Cognit* 30: 758-67.
- Craik FIM, Lockhart RS (1972) The levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11: 671-684.
- Craik FIM, Tulving E (1975) Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General* 104: 268-94.
- Critchley, H, Daly, E, Phillips, M, Brammer, M, Bullmore, E, Williams, S, Van Amelsvoort, T, Robertson, D, David, A, Murphy, D. (2000). Explicit and implicit neural mechanisms for processing of social information from facial expressions: a functional magnetic resonance imaging study. *Human Brain Mapping*, 9: 93-105.
- Dehaene S, Sergent C, Changeux JP (2003) A neuronal network model linking subjective reports and objective physiological data during conscious perception.

- Proc Natl Acad Sci U S A 100: 8520-8525.
- Downing P, Liu J, Kanwisher N (2001) Testing cognitive models of visual attention with fMRI and MEG. *Neuropsychologia*, 39: 1329-1342.
- Duncan J (2001) An adaptive coding model of neural function in prefrontal cortex. *Nat Rev Neurosci* 2: 820-829.
- De Vries PH, Dalenoort GJ (1995) The binding problem in distributed systems. In: *Neural networks: artificial intelligence and industrial applications* (Kappen B, Gielen S, eds.), pp. 77-80. Berlin: Springer.
- Engel AK, Kreiter AK, König P, Singer W (1991) Synchronization of oscillatory neuronal responses between striate and extrastriate visual cortical areas of the cat. *Proc Natl Acad Sci U S A* 88: 6048-52.
- Ericsson K, Kintsch W (1995) Long-term working memory. *Psychological Review* 102: 211-245.
- Fuster JM (1989) *The prefrontal cortex* (2<sup>nd</sup> ed). New York: Raven.
- Goebel R (1993) Perceiving complex visual scenes: an oscillator neural network model that integrates selective attention perceptual organisation and invariant recognition. In: *Advances in Neural Information Processing Systems 5* (Hanson, SJ Cowan JD, Giles CL, eds.), pp. 903-910. San Mateo (CA): Morgan Kaufmann.
- Goldman-Rakic PS (1996) Memory: recording experience in cells and circuits: diversity in memory research. *Proc Natl Acad Sci U S A* 93: 13435-13437.
- Gray CM, Singer W (1989) Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc Natl Acad Sci U S A* 86: 1698-702.
- Gross J, Schmitz F, Schnitzler I, Kessler K, Shapiro K, Hommel B, Schnitzler A (2004) Long-range neural synchrony predicts temporal limitations of visual attention in humans. *Proc Natl Acad Sci U S A*, in press.
- Gruber O, von Cramon DY (2003) The functional neuroanatomy of human working

- memory revisited: evidence from 3-T fMRI studies using classical domain-specific interference tasks. *Neuroimage* 19: 797-809.
- Guttentag R (1995) Children's associative learning: automatic and deliberate encoding of meaningful associations. *Am J Psychol* 108: 99-114.
- Heil M, Rösler F, Hennighausen E (1996) Topographically distinct cortical activation in episodic long-term memory: the retrieval of spatial versus verbal information. *Memory and Cognition* 24: 777-795.
- Hommel B (2004) Event files: feature binding in and across perception and action. *Trends in Cognitive Science* in press.
- Hopf JM, Vogel E, Woodman G, Heinze HJ, Luck SJ (2002) Localizing visual discrimination processes in time and space. *J Neurophysiol* 88: 2088-2095.
- Hyde TS, Jenkins JJ (1969) Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology* 82: 472-81.
- Incisa della Rocchetta A, Milner B (1993) Strategic search and retrieval inhibition: the role of the frontal lobes. *Neuropsychologia*, 31: 503-524.
- Kessler K, Rickheit G (1999) Dynamic concept generation in connectionist networks: definitions of a model for complex scene reconstruction and experimental results. *Kognitionswissenschaft* 8: 74-96.
- Kessler K, Rickheit G (2000) Dynamic binding and context influences during scene perception: simulations and experimental results. In: *Proceedings of the Third International Conference on Cognitive Modelling* (Taatgen N, Aasman J, eds.), pp. 177-184. Veenendaal: Universal Press.
- Kessler K, Tipper SP (2004) Retrieval of implicit inhibitory processes: the impact of visual field, object-identity and memory dynamics. *Visual Cognition*, in press.
- Kiefer M, Ahlegian M, Spitzer M (2004) Working memory capacity, indirect semantic

- priming and Stroop interference: pattern of interindividual prefrontal performance differences in healthy volunteers. *Neuropsychology*, in press.
- Kiefer M, Marzinzik F, Weisbrod M, Scherg M, Spitzer M (1998a) The time course of brain activations during response inhibition: evidence from event-related potentials in a Go/Nogo task. *NeuroReport* 9: 765-770.
- Kiefer M, Weisbrod M, Kern I, Maier S, Spitzer M (1998b) Right hemisphere activation during indirect semantic priming: Evidence from event-related potentials. *Brain and Language* 64: 377-408
- Kikyo H, Ohki K, Miyashita Y (2002) Neural correlates for feeling-of-knowing: an fMRI parametric analysis. *Neuron* 36: 177-186.
- Lisman JE, Idiart MA (1995 ) Storage of 7 +/- 2 short-term memories in oscillatory subcycles. *Science* 267: 1512-1515.
- Logie RH (1995) *Visuo-spatial working memory*. Hove (UK): Lawrence Erlbaum.
- Maril A, Wagner AD, Schacter DL (2001) On the tip of the tongue: an event-related fMRI study of semantic retrieval failure and cognitive conflict. *Neuron* 31: 653-660.
- Martin A (1999) Automatic activation of the medial temporal lobe during encoding: lateralized influences of meaning and novelty. *Hippocampus* 9: 62-70.
- McClelland JL, McNaughton BL, O'Reilly RC (1995) Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev* 102: 419-457.
- Mecklinger A (1999) *Remembering places and objects*. Göttingen: Hogrefe.
- Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience* 24: 167-202
- Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J Neurosci* 16: 5154-5167.

- Miyake A, Shah P (1999) Toward unified theories of working memory: emerging general consensus, unresolved theoretical issues, and future research directions. In Miyake A, Shah P (Eds.), *Models of Working Memory* (pp. 442-481). Cambridge (UK): University Press.
- Nunez PL (1981) *Electrical fields of the brain: The neurophysics of EEG*. New York: Oxford University Press.
- Nyberg L, McIntosh AR, Cabeza R, Habib R, Houle S, Tulving E (1996) General and specific brain regions involved in encoding and retrieval of events: what where and when. *Proc Natl Acad Sci U S A* 93: 11280-11285.
- Nyberg L, Habib R, McIntosh AR, Tulving E (2000) Reactivation of encoding-related brain activity during memory retrieval. *Proc Natl Acad Sci U S A* 97: 11120-11124.
- Posner MI, Petersen SE, Fox PT, Raichle ME (1988) Localization of cognitive operations in the human brain. *Science* 240: 1627-1631.
- Ranganath C, D'Esposito M (2001) Medial temporal lobe activity associated with active maintenance of novel information. *Neuron* 31: 865-873.
- Ranganath C, Cohen MX, Dam C, D'Esposito M (2004) Inferior temporal prefrontal and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *J Neurosci* 24: 3917-3925.
- Rösler F, Heil M, Hennighausen E (1995) Distinct cortical activation patterns during long-term memory retrieval of verbal, spatial and color information. *Journal of Cognitive Neuroscience* 7: 51-65.)
- Ruchkin DS, Canoune HL, Johnson RJ, Ritter W (1995) Working memory and preparation elicit different patterns of slow wave event-related brain potentials. *Psychophysiology* 32: 399-410.
- Ruchkin DS, Johnson R, Jr, Grafman J, Canoune H, Ritter W (1992) Distinctions and similarities among working memory processes: an event-related potential study.



Cognitive Brain Research 1: 53-66

Sakai K, Rowe JB, Passingham RE (2002) Parahippocampal reactivation signal at retrieval after interruption of rehearsal. *J Neurosci* 22: 6315-63120

Sakai K, Passingham RE (2004) Prefrontal selection and medial temporal lobe reactivation in retrieval of short-term verbal information. *Cereb Cortex* 4: 914-921

Sarnthein J, Petsche H, Rappelsberger P, Shaw GL, von Stein A (1998) Synchronization between prefrontal and posterior association cortex during human working memory. *Proc Natl Acad Sci U S A* 95: 7092-6.

Schacter DL, Wagner AD (1999) Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. *Hippocampus* 9: 7-24.

Shimamura AP (1995) Memory and the prefrontal cortex. *Ann N Y Acad Sci*, 769: 151-159.

Smith EE, Jonides J, Koeppe R.A, Awh E, Schumacher E.H, Minoshima S (1995) Spatial versus object working memory: PET investigations. *Journal of Cognitive Neuroscience* 7: 337-356.

Smolensky P (1988) On the proper treatment of connectionism. *Behavioral and Brain Sciences* 11: 1-74.

Snyder AZ, Abdullaev Y.G, Posner M.I, Raichle M.E (1995) Scalp electrical potentials reflect regional blood flow responses during processing of written words. *Proc Natl Acad Sci U S A* 92: 1689-1693.

Stark CE, Bayley PJ, Squire LR (2002) Recognition memory for single items and for associations is similarly impaired following damage to the hippocampal region. *Learn Mem* 9: 238-42.

Sutherland GR, McNaughton B (2000) Memory trace reactivation in hippocampal and neocortical neural ensembles. *Curr Opin Neurobiol* 10: 180-186.

Sutherland RJ, Rudy JW (1988) Configural association theory: The role of the

- hippocampal formation in learning, memory and amnesia. *Psychobiology* 16: 157-163.
- Taylor TJ, Duncann P (1986) The hippocampal memory indexing theory. *Behavioral Neuroscience* 100: 147-154.
- Tomita H, Ohbayashi M, Nakahara K, Hasegawa I, Miyashita Y (1999) Top-down signal from prefrontal cortex in executive control of memory retrieval. *Nature* 401: 699-703.
- Varela F, Lachaux JP, Rodriguez E, Martinerie J (2001) The brainweb: phase synchronization and large-scale integration. *Nat Rev Neurosci* 2: 229-239.
- von der Malsburg C (1999) The what and why of binding: the modeler's perspective. *Neuron* 24: 95-104, 111-25.
- Weisbrod M, Kiefer M, Marzinzik F, Spitzer M (2000) Executive control is disturbed in schizophrenia: Evidence from event-related potentials in a Go/Nogo task. *Biological Psychiatry* 47: 51-60.
- Weiss S, Rappelsberger P (2000) Long-range EEG synchronization during word encoding correlates with successful memory performance. *Cogn Brain Res* 9: 299-312.
- Westerman DL. (2001) The role of familiarity in item recognition, associative recognition, and plurality recognition on self-paced and speeded tests. *J Exp Psychol Learn Mem Cogn* 27: 723-32.
- White M (1995) Preattentive analysis of facial expressions of emotion. *Cognition & Emotion*, 9: 439-460.
- Zurowski B, Gostomzyk J, Grön G, Weller R, Schirrmeyer H, Neumeier B, Spitzer M, Reske S, Walter H (2002) Dissociating a common working memory network from different neural substrates of phonological and spatial stimulus processing. *NeuroImage* 15: 45-57

Tables

Table 1: Means (avg) and standard deviations (std.dev.) for each experimental condition (with/without interference; high/low load; same/different size of prime and probe). Top: Accuracy (ACC) in terms of percentage correct responses. Bottom: Reaction times (RT) in ms.

	with interference				without interference			
	high		low		high		low	
ACC (%)	same	diff	same	diff	same	diff	same	diff
avg	72.2	57	83.7	62.5	75	61.2	83.8	66
std.dev.	8.2	9.1	12.4	12.5	9.8	13.4	12.4	11.1
RT (ms)								
avg	1604	1676	1439	1551	1532	1686	1329	1503
std.dev.	374	360	331	313	335	362	354	288

Figure captions

Figure 1: Stimulus material employed in the experiment. Panel A shows all shapes that could be displayed for encoding. In Panel B the variation in size between prime and probe is shown: Either the width or the height were increased or decreased by 25%. Panel C shows two examples for the stimuli used in the interference task. Participants were instructed to press a key if the inner black shape was round and not edged (go/no-go paradigm).

Figure 2: The procedure of the experiment. Further explanations in the text.

Figure 3: Time course of activation across the entire trial (9.3 sec). A frontal electrode is shown at the top while an occipital electrode is depicted at the bottom. Baseline was calculated for the 200 ms prior to the onset of the encoding displays. The interval of most relevance for the present report were the 2 sec of recovery after interference and prior to probe administration.

Figure 4: The four electrode clusters analyzed in the present report. Bilateral pairs of electrodes were selected over frontal (AF7=12/ AF8 = 46, F9 = 2/ F10 = 34; FT7 = 11/ FT8 = 45; T1 = 3/ T2 = 35) and occipito-temporal cortex (P9 = 5/ P10 = 37; P7 = 8/ P8 = 42; O9 = 6/ O10 = 38, O1 = 7/ O2 = 41).

Figure 5: Time course of activation in the recovery interval of the analyzed 16 electrodes. As described in the Methods section, epochs have been averaged relative to the onset of the trial (and by using the 200 ms before as a baseline). As the timing of the procedure in all trials (i.e. onset of the recovery period in relation to trial onset) was always identical with an accuracy below 1 ms (warranted by ERTS® experimental stimulation software) the depicted time

course of activation in the recovery interval may be regarded as being averaged relative to the onset of the blue fixation cross (begin of recovery period). This onset is congruent to 0 ms on the x-axes. Gray arrows indicate ERP components that are stronger in the condition with interference. At frontal sites two of these components can be identified across the electrodes in the time interval from 400 to 1000 ms. In addition, around 200 ms an earlier positive response at frontal electrodes is observed for the conditions with interference that is also indicated by a gray arrow. All occipito-temporal electrodes show stronger slow negative waves in the conditions with interference during the last 1000 ms of the recovery interval. Black arrows indicate ERP components that are stronger in the condition without interference. Such a component can only be observed at occipito-temporal electrodes at around 300-400 ms.

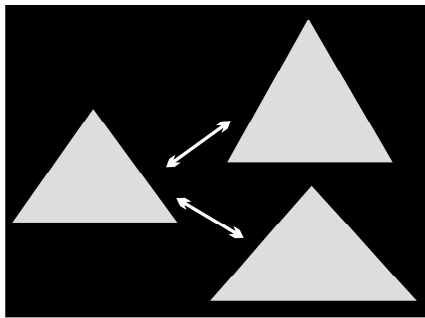
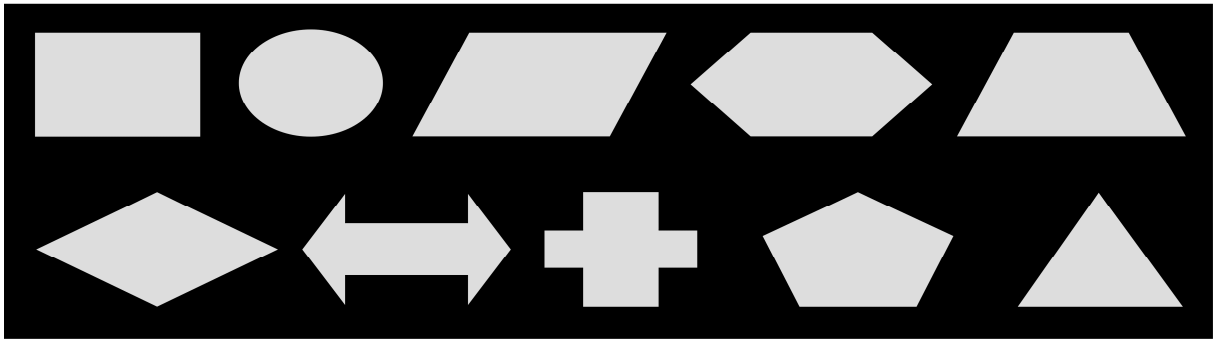
Figure 6: Significant effects in the 1<sup>st</sup> (Panels A and B) and the 3<sup>rd</sup> (Panel C) 500 ms time windows of the recovery interval. Panels A and C show the interaction between “interference” (with/without) and “horizontal alignment” (frontal/occipito-temporal) that reveals a reversed picture in the 1<sup>st</sup> as compared to the 3<sup>rd</sup> time window. Panel B depicts the 3way interaction between “interference” (with/without), “horizontal alignment” (frontal/occipito-temporal), and “load” (low/high) that was significant in the 1<sup>st</sup> time window.

Figure 7: Sample by sample t-tests for the comparison: without interference vs. with interference, separately for high load (top row) and low load (bottom row). Dotted horizontal lines indicate the 95%  $\alpha$ -significance level (two-sided). A left frontal (black line, AF7) and a left occipito-temporal (gray line, O1) electrode are shown on the left, while the corresponding right hemisphere electrodes (AF8 and O2 resp.) are shown on the right. Note that due to polarity of the comparisons (without vs. with interference) a positive t-test value could either indicate a stronger negative or a weaker positive ERP component in the condition with

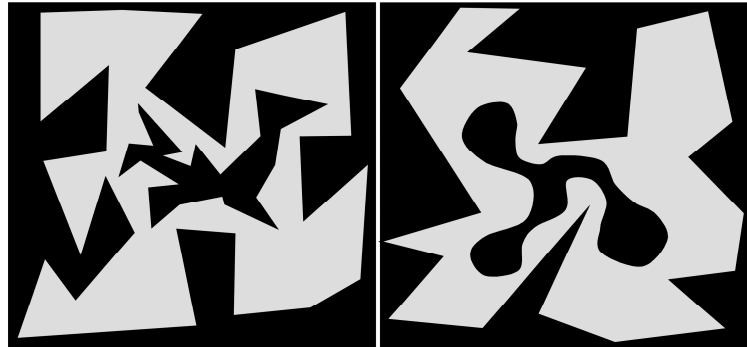
interference. Corresponding reflections apply to negative t-test values. Disambiguation of t-test values may be accomplished by considering the time course of activation in Figure 5. Accordingly, the black arrows denote positive and negative t-test values at frontal sites that reflect stronger negative and stronger positive ERP components in the condition with interference (conform Figure 5). Gray arrows denote positive t-test values that reflect the beginning of the slow negative wave at occipito-temporal sites shown in Figure 5, which is also stronger in the condition with interference.

Figure 8: Sample by sample t-tests for the comparison: without interference vs. with interference, separately for high load (top row) and low load (bottom row). Dotted horizontal lines indicate the 95%  $\alpha$ -significance level (two-sided). A left frontal electrode (T1) is shown on the left and a right frontal electrode (T2) is shown on the right. Black arrows denote the negative t-test values at frontal sites that reflect an earlier, hence stronger positive ERP component at the time of the peak in the condition with interference (conform Figure 5).

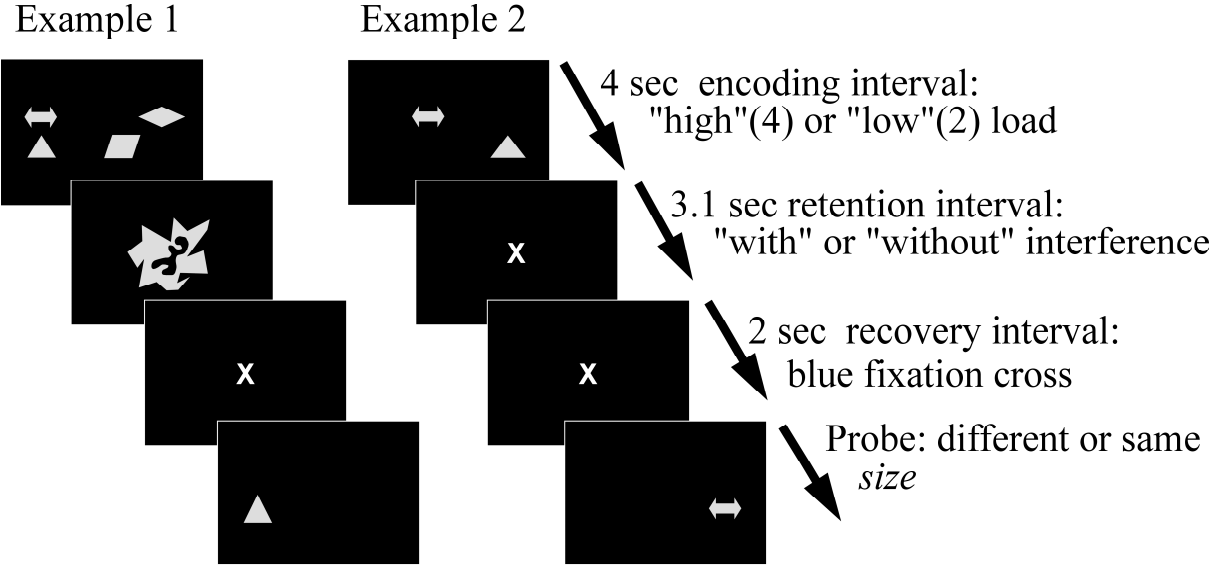
Panel A



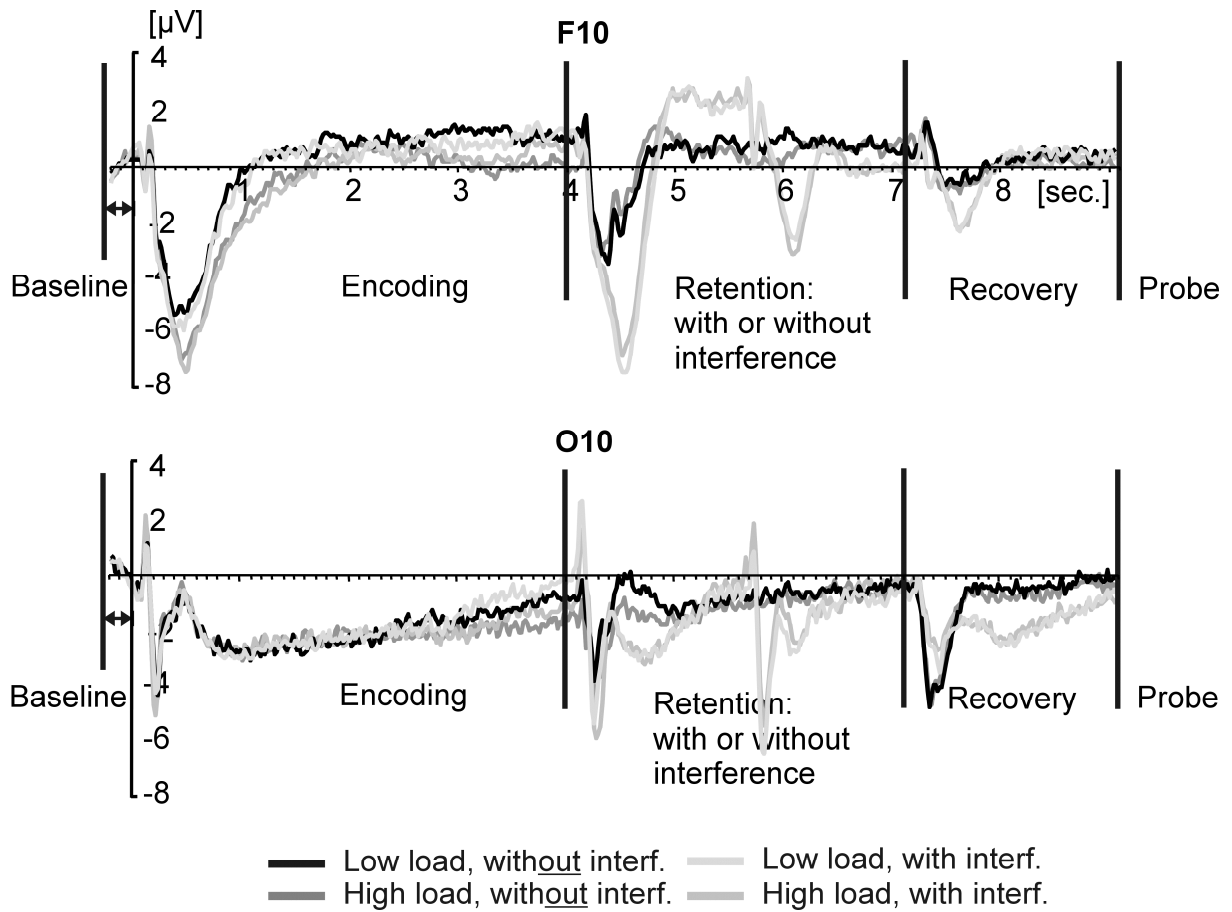
Panel B

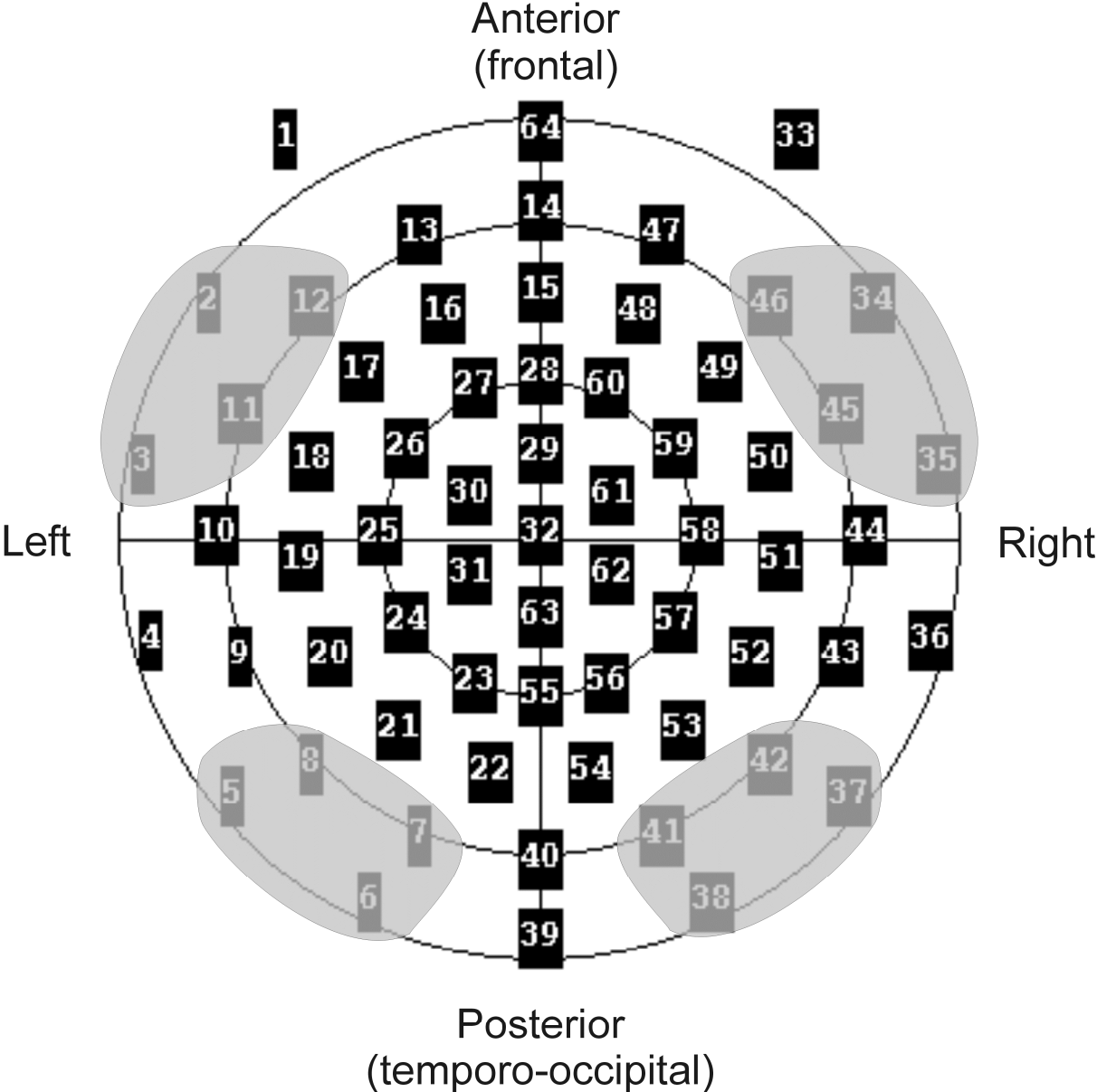


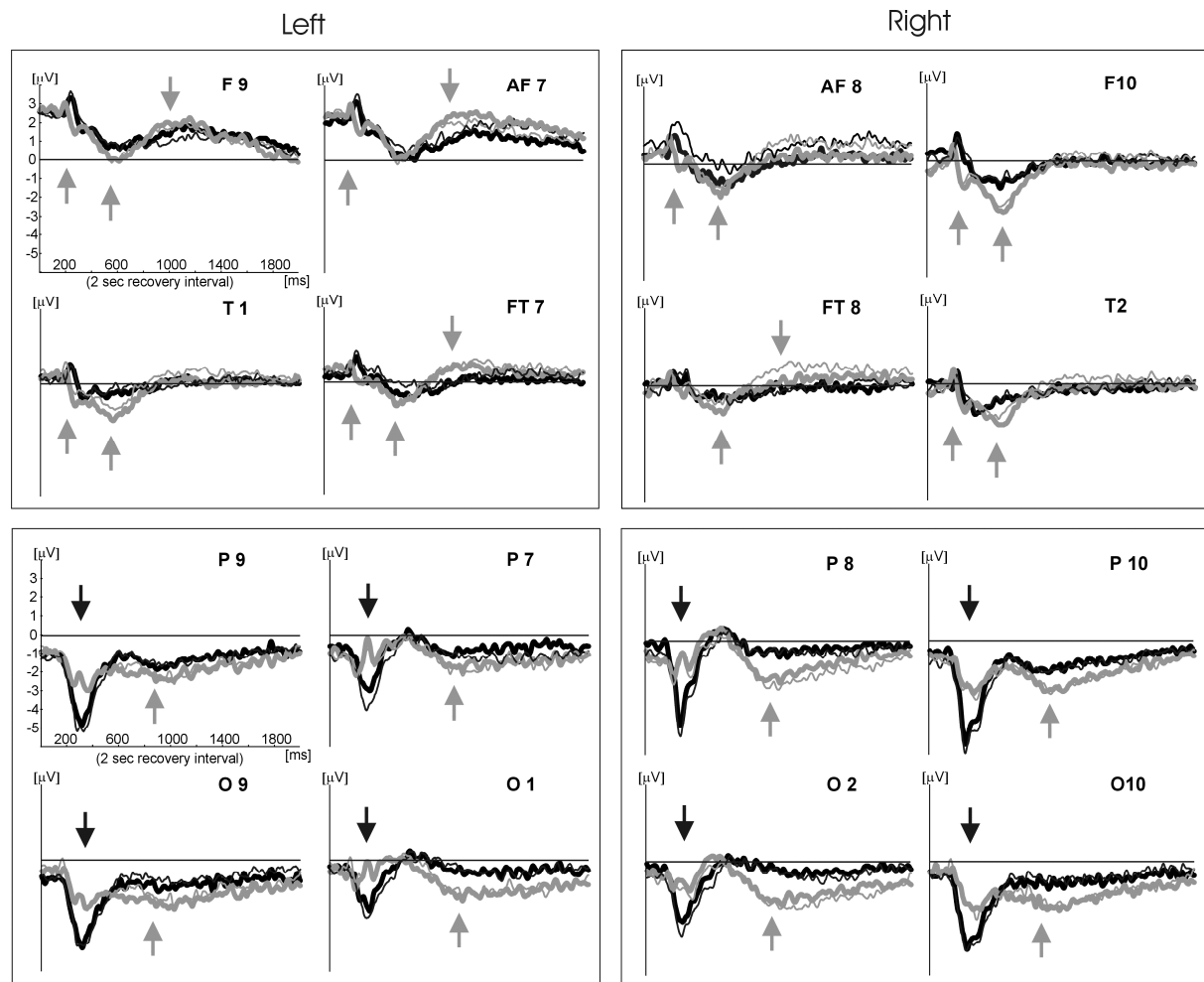
Panel C



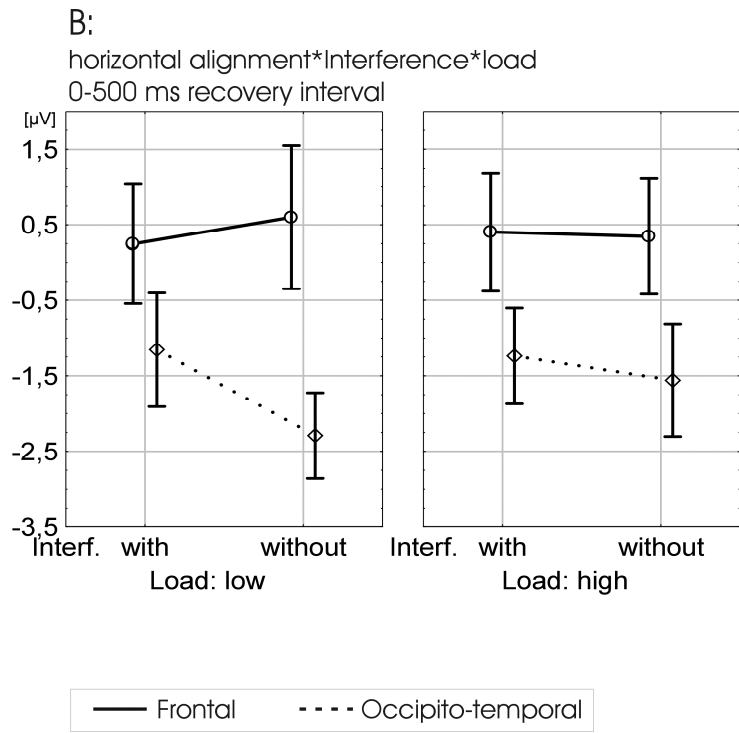
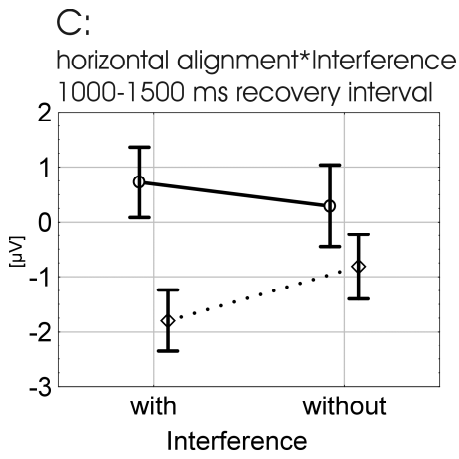
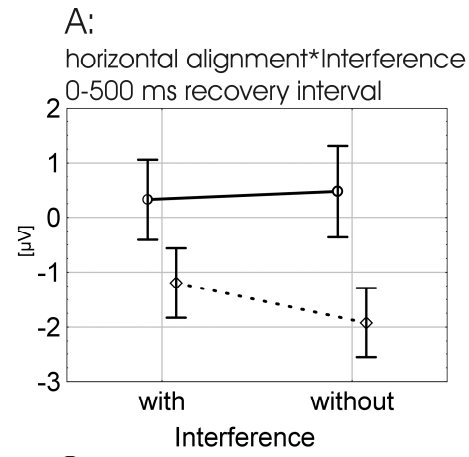






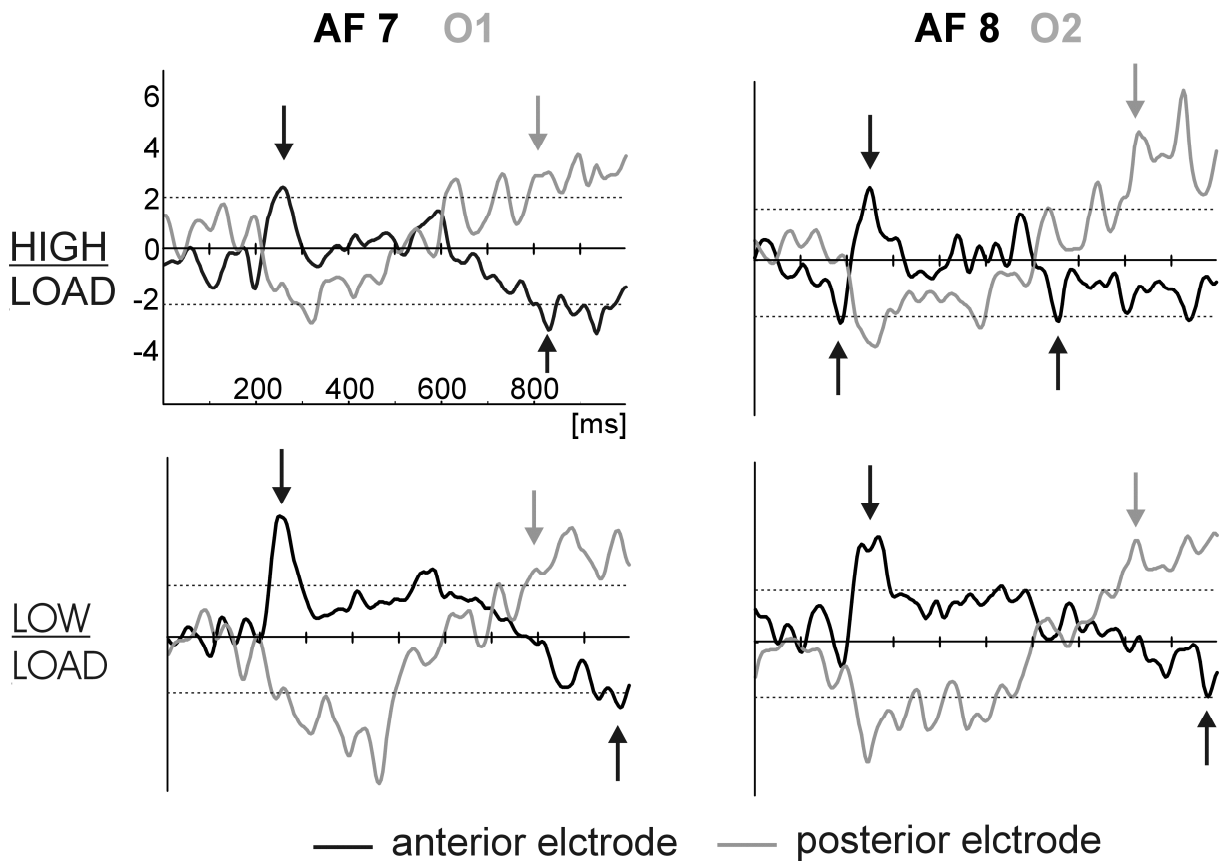


— low load, without interf.    — low load, with interf.  
 — high load, without interf.    — high load, with interf.



— Frontal    - - - - Occipito-temporal

### T-tests: 0-1000 ms of the recovery interval



### T-tests: 0-1000 ms of the recovery interval

