Title:
Motor conflict in Stroop tasks: direct evidence from single-trial electro-myography and electro-encephalography

Authors:
Dénes Szűcs[1,2] ; Fruzsina Soltész[1], Sonia White[1]

Affiliations of authors:
[1] Centre for Neuroscience, University of Cambridge, Faculty of Education
     184 Hills Road, CB2 8PQ, United Kingdom
[2] Research Institute for Psychology of the Hungarian Academy of Sciences, Department of Psycho-physiology, Budapest, Hungary

Corresponding author: Dénes Szűcs
Email: ds377@cam.ac.uk
Phone: +44-01223-767600
Fax: +44-01223-767602

Abstract

Several brain imaging studies have assumed that response conflict is present in Stroop tasks. However, this has not been demonstrated directly. We examined the time-course of stimulus and response conflict resolution in a numerical Stroop task by combining single-trial electro-myography (EMG) and event-related brain potentials (ERP). EMG enabled the direct tracking of response conflict and the peak latency of the P300 ERP wave was used to index stimulus conflict. In correctly responded trials of the incongruent condition EMG detected robust incorrect response hand activation which appeared consistently in single trials. In 50-80% of trials correct and incorrect response hand activation coincided temporally, while in 20-50% of trials incorrect hand activation preceded correct hand activation. EMG data provides robust direct evidence for response conflict. However, congruency effects also appeared in the peak latency of the P300 wave which suggests that stimulus conflict also played a role in the Stroop paradigm. Findings are explained by the continuous flow model of information processing: Partially processed task-irrelevant stimulus information can result in stimulus conflict and can prepare incorrect response activity. A robust congruency effect appeared in the amplitude of incongruent vs. congruent ERPs between 330-400 ms, this effect may be related to the activity of the anterior cingulate cortex.

Introduction

It is a major question whether conflict is resolved at the level of stimulus representations or at the level of motor response preparation (Hock and Egeth, 1970; Morton and Chambers, 1973; Posner and Snyder, 1975). The Stroop paradigm is the most popular experimental model of conflict situations (Stroop, 1935; MacLeod, 1991). In this paradigm both task-relevant and task-irrelevant stimulus dimensions deliver information. Participants have to respond according to the task-relevant dimension while neglecting the task-irrelevant dimension. Generally, it is assumed that the Stroop paradigm evokes both stimulus and response conflict. Accordingly, several functional magnetic resonance imaging (fMRI) studies attempted to separate brain correlates of stimulus and response conflict in various Stroop paradigms (for reviews see van Veen and Carter 2005; Liston et al. 2006). However, no Stroop studies measured response conflict directly at the level of effectors. Hence, conclusions about response conflict rest on assumptions about mental processes involved in paradigms. Here we have filled this gap by measuring response conflict directly, by using electro-myography (EMG). Furthermore, we used EMG to guide the interpretation of event-related brain potential (ERP) data, in order to relate ERP findings to response conflict effects.

Stimulus conflict is due to a mismatch at the level of the representations of stimulus dimensions (Hock and Egeth, 1970). Response conflict appears when the relevant and irrelevant stimulus dimensions are processed, up to response initiation, in parallel. In the congruent condition of a Stroop task both the task-relevant and the task-irrelevant stimulus dimensions prepare similar responses. In contrast, in the incongruent condition the task-relevant and the task-irrelevant stimulus dimensions may prepare

opposing response tendencies. It is this motor conflict that may hinder performance. Several ERP studies have investigated conflict processing in Stroop tasks by exploiting the high temporal resolution of ERPs. One line of ERP studies attempted to separate stimulus and response conflict by examining the latency of ERP effects (for reviews see Szűcs and Soltész, 2007; 2008). Some studies assumed that relatively early (at around 200 ms) congruency effects modulating ERP amplitude precede response preparation. Hence, early congruency effects in amplitude were interpreted as expressions of stimulus conflict (Aine and Harter, 1984; Atkinson et al. 2003). Other studies tested the latency of the P300 ERP wave, whose latency is usually thought to correlate with the completion of stimulus analysis but not with response preparation (Kutas et al. 1977; McCarthy and Donchin, 1981; Donchin, 1981). Duncan-Johnson and Koppel (1981) and Ilan and Polich (1999) could not demonstrate congruency effects in the latency of the P300 ERP wave in the classical color-word Stroop task (Stroop, 1935). Hence, it was concluded that congruency effects were not related to stimulus conflict; rather, they were assumed to be related to response conflict. In contrast, we demonstrated congruency effects on P300 latency in a numerical Stroop task (Szűcs, Soltész, Jármi and Csépe, 2007; adult group). This finding suggests that stimulus conflict could, in fact, play a role in Stroop tasks.

Another line of ERP Stroop studies has focused on the analysis of topographical features of ERPs. The most prominent topographic effect identified by this research is a negativity appearing at around 350-500 ms in incongruent minus congruent difference potentials (most previous ERP studies did not use a neutral condition when examining the topography of congruency effects; for a review see Szűcs and Soltész, 2007). This negativity is usually called the N450 (West 2003) and has been interpreted in two ways:

On the one hand, some studies (Liotti et al. 2000; West, 2003) considered it as a correlate of the activity of the Anterior Cingulate Cortex (ACC) because several fMRI studies have demonstrated ACC activation in Stroop tasks (Milham et al. 2001; van Veen and Carter 2005; Liston et al. 2006). The ACC has been implicated in response conflict detection and resolution (Botvinick et al. 1999; van Veen et al. 2001; Milham et al. 2001; Kerns et al. 2004; Botvinick et al. 2004). Hence, an ACC source of the effect would suggest that the N450 is related to response conflict detection and resolution. On the other hand, some studies (Rebai et al. 1997) suggested that the congruency effect was related to the N400 wave which is thought to index semantic integration (Kutas and Hillyard, 1980) and appears at semantically unexpected events even when these are represented by single digits (Szűcs, Soltész, Czigler and Csépe, 2007). This explanation would suggest that the congruency effect is more related to conflict at the level of stimulus representations. Currently data cannot clearly discriminate between the above two hypotheses.

@ **Figure 1**

In this study the numerical version of the Stroop paradigm was used which we used in previous similar studies (**Fig. 1.**). In this paradigm participants judge which of two simultaneously presented digits is larger in physical size (Henik and Tzelgov, 1982). In the neutral condition the two digits are of equal numerical value. That is, the task-irrelevant stimulus dimension (numerical magnitude) delivers neutral information relative to the task-relevant stimulus dimension (physical size). In the congruent condition the task-relevant and task-irrelevant dimensions are in coherence with each other. In the incongruent condition the task-relevant and task-irrelevant dimensions are in conflict with each other. Performance is worse in the incongruent (conflict) condition in

comparison to the congruent and neutral conditions. An advantage of the numerical Stroop paradigm over the classical color word Stroop paradigm (Stroop, 1935) is that the numerical paradigm easily enables the parametric manipulation of task-difficulty. This can be done by varying the physical size difference between digits. In the easy condition the physical size difference between digits is large. In the difficult condition the physical size difference between digits is small (the size difference between digits is harder to discriminate). Manipulating task difficulty is advantageous because it is thought to modulate the amount of response conflict (Posner 1978; MacLeod, 1991 for review).

In a series of ERP studies we have identified robust ERP amplitude markers of conflict in the numerical Stroop paradigm (Szűcs Soltész, Jármi and Csépe, 2007; Szűcs and Soltész, 2007; 2008). In these studies we analyzed amplitude effects in relation to the timing of the Lateralized Readiness Potential (LRP). The LRP is computed from the voltage recorded at electrodes positioned over the motor cortex, and it is thought to index movement preparation (Gratton et al. 1988; De Jong et al. 1988). We assumed that effects preceding LRP onset were related to stimulus conflict, while the timing and topography of effects following LRP onset may have been related to both stimulus and response conflict. In our Stroop studies we have identified ERP amplitude effects in the incongruent relative to the neutral condition both before and after LRP onset, including the previously mentioned N450 effect. Overall it was suggested that our findings provided evidence for both stimulus and motor conflict in Stroop tasks.

Our previous findings were explained within the framework of the continuous flow model of Eriksen and Schultz (1979). This model assumes that perceptual stimulus processing and response activation/organization cannot be separated clearly. Rather than

following each other in a serial fashion, the stages overlap and perceptual processing influences response preparation in a continuous fashion. That is, perceptual information accumulates gradually and the outcome of this analysis is continuously fed into the response system where it continuously affects response activity. Consequently, even partially processed perceptual stimulus attributes will result in motor activation even before the completion of final perceptual analysis. According to the model, a motor response is triggered when the activation level of a response exceeds a criterion (Eriksen and Schultz, 1985; Eriksen et al. 1985; Coles et al. 1985; Gratton et al. 1988; Smid et al. 1990). The perceptual flow model can explain the presence of both stimulus and response conflict in Stroop tasks. However, a shortcoming of our previous studies was that similarly to other studies of the Stroop effect we could not directly measure motor-conflict. Hence, data interpretation relied on assumptions about the temporal relations of ERP effects and the LRP onset.

Currently the strongest neuroimaging evidence supporting the response-conflict theory of the Stroop effect comes from a single functional near-infrared imaging study (fNIRS) (DeSoto et al. 2001). This study sampled brain activity at 25Hz temporal resolution in a manual spatial Stroop task. The study measured lateralized fNRIS activity over motor cortices, exploiting the fact that fNIRS data is not subject to volume conduction like ERPs. Hence, unlike the LRP, lateralized fNRIS activity is able to characterize response preparation in motor cortices in a hand-specific manner. The average fNIRS phase value during an 80 ms period preceding the response was computed and brain activity was characterized by measuring the phase of the optical signal. This technique was based on the fact that increased phase-delay of the optical signal

(increased average time of flight of photons through the head) can serve as a marker of increased physiological activity of the brain (Gratton and Limkeman, 1983). Results showed increased phase-delay over the motor cortex ipsilateral to response hand in incongruent but not in the congruent trials, and increased phase-delay over the motor cortex contralateral to response hand in both congruent and incongruent trials. Put otherwise, the motor cortex associated with the correct response hand was activated in both the congruent/incongruent conditions, while the motor cortex associated with the incorrect response was activated only in the incongruent condition. This data was interpreted as demonstrating selective incorrect response activation in the incongruent condition, supporting the response competition theory of the Stroop effect.

Further direct evidence for response competition in other tasks have been found in seminal studies using EMG to represent correct and incorrect response activation. These studies have argued that according to the motor response competition theory effectors involved in preparing and executing a potential incorrect response may be subliminally activated, even in correctly responded trials. Hence, EMG may have the potential to observe incorrect effector activation in a direct manner. Such data can provide direct evidence for response conflict and motor response competition. This proved to be a fruitful approach in the Eriksen flanker task (Eriksen et al. 1985; Coles et al. 1985; Gratton et al. 1988) and the Simon task (Hasbroucq et al. 1999; Burle et al. 2002), where EMG was successfully used to demonstrate latent muscle activity in incorrect response hands.

Building on the above studies, here we have not only recorded ERPs but also recorded EMG synchronized with ERPs in a numerical Stroop task. First, we exploited

previously used ERP measures: peak latency measures (timing), amplitude, topography and the measurement of the onset of the LRP (Duncan-Johnson and Koppel, 1981; Ilan and Polich 1999; Liotti et al. 2000; West 2003; West and Alain, 1999; Szűcs Soltész, Jármi and Csépe, 2007; Szűcs and Soltész, 2007; 2008). Second, EMG was also measured. We hypothesized that if response conflict plays a crucial role in the numerical Stroop task than the activation of the incorrect response hand should be detected by EMG, even in correctly responded trials of the incongruent condition. One caveat may have been that incorrect response hand EMG activity may have appeared in all slowly responded trials, irrespective of experimental condition. In order to exclude this alternative explanation we conducted separate analyses of trials that were deemed as having either a quick or slow response. This way we examined whether incorrect hand EMG activity was related to performance and whether it was specific to the incongruent condition. That is, true incorrect response activation must be specific to the incongruent condition both in slow and quick response trials. In order to investigate the consistency of the relationship of EMG amplitude and performance we correlated single-trial EMG amplitude with reaction time (RT). If incorrect response activation is closely related to behavioral congruency effects, then the correlation of EMG amplitude and RT can be expected.

Apart from the potential co-activation of correct and incorrect response hands in the incongruent condition it is a further question whether the timing of correct and incorrect hand activity is co-incident or whether correct hand activation follows incorrect hand activation (DeSoto et al. 2001). In order to study this question, we also compared the timing of EMG activity in correct and incorrect response hands in single trials.

Furthermore, EMG was used to guide the analysis of high-resolution 129-channel electro-encephalography data. This enabled effective examination of the temporal relationship of the N450 ERP effect and motor conflict. If the N450 is related to ACC motor conflict detection/resolution its timing must be closely related to the activity of the incorrect response hand. The effect of numerical (semantic) distance on ERP data was also tested, because such effects have been related to the classical N400. Hence, semantic distance effects in the N450 time range would link the N450 to the classical N400 (Niedeggen & Rösler 1999; Szucs & Csépe 2005).

Materials and Methods

Participants

18 adults' data (mean age 25±5 years; 5 males, 3 left-handed) were analyzed. Participants were graduate and undergraduate students at the University of Cambridge. The study received ethical approval from the Psychology Research Ethics Committee of the University of Cambridge.

Stimuli and procedure

A stimulus consisted of two Arabic digits shown simultaneously in the middle of a 19-inch computer screen. Stimuli were white digits on black background. A trial started with a fixation sign shown for 300 ms (the drawing of an eye). This was followed by a delay for about 1000 ms (a random interval between ± 50 ms was added to 1000 ms in order to suppress alpha activity which may time-lock to stimulus-presentation). After this the stimulus appeared for 3000 ms, or until the participant gave a behavioral response. The inter-trial interval was 1000 ms.

The following number pairs were used as stimuli in the congruent and incongruent conditions: 1-2, 2-1, 8-9, 9-8, 1-8, 8-1, 2-9 and 9-2. The following number pairs were used in the neutral condition: 1-1, 2-2, 8-8 and 9-9. The two digits in a pair were of different physical font size. In one half of trials the physical size of digits was 40 and 45 points. In the other half the physical size of digits was 40 and 50 points. This manipulation defined the task difficulty factor: the physical size difference between digits was either small (5 points: difficult condition) or large (10 points: easy condition). The numerical distance between digits was either 1 or 7 in the congruent and incongruent conditions (numerical distance factor), and 0 in the neutral condition. Exactly the same digits were used for both the numerical distance 1 and 7 conditions. By using two numerical distances we kept the number of levels of the task-irrelevant factor at the same value of the number of levels of the task-relevant factor. In the congruent condition the physical larger digit was also numerically larger than the other one. In the incongruent condition the physically larger digit was numerically smaller than the other one. In the neutral condition the two digits were of the same numerical value. In half of the trials the physically larger number appeared on the right, in the other half, on the left. The same held for the position of the numerically larger number as well. Congruency, size difference, numerical distance, and the side of the response (left or right hand) were manipulated orthogonally.

Participants' task was to indicate with a button press whether the physically larger number appeared on the left or on the right. Participants pressed response buttons with their thumbs. Stimuli in stimulus sequences were pseudo-randomized in a way that controlled for the number and distribution of response side (left or right) combined with

congruency in one stimulus sequence. All 36 possible combinations of response side (4 combinations: right after right, left after right, left after left and right after left) and congruency (9 possible pairs of congruent, incongruent and neutral) were controlled for and evenly distributed within a stimulus sequence in order to avoid any response preparation bias. Each participant received a unique pseudo-randomized sequence of stimuli, equating response preparation effects both within- and across subjects. There were 6 experimental blocks with 96 trials in each block (576 stimuli). The experiment was preceded by 24 practice trials. Stimuli were delivered by Presentation 11 (Neuro-behavioral systems).

<u>Behavioral data analysis</u>

Accuracy and RT were analyzed by Congruency (neutral, congruent and incongruent) × Difficulty (small vs. large size difference) × Numerical distance (1 vs. 7) × Side of response (left vs. right hand) repeated measures ANOVAs. By definition numerical distance was zero in the neutral condition. Therefore, in the above ANOVA half of neutral trials were randomly assigned to the numerical distance 1 condition, and the other half was assigned to the numerical distance 7 condition. Tukey-HSD tests were used for post-hoc analyses. Because the effect of numerical distance may be suppressed in the above ANOVA, another Congruency (congruent and incongruent only) × Difficulty × Numerical distance × Side of response ANOVA was run merely to confirm findings relating to the numerical distance factor.

In a third analysis, individual trials in each subject in each congruency condition were sorted into slow and fast response trials. Fast trials were defined as trials responded faster than the mean RT of a particular participant in the congruent condition. Slow trials

were defined as trials responded slower than the mean RT of a particular participant in the incongruent condition. In order to compare RT in slow and fast trials a Speed (Fast vs. Slow) × Congruency × Difficulty × Numerical Distance × Side of response ANOVA was run. Tukey tests for Speed × Congruency contrasts were used to evaluate congruency effects in fast and slow trials. In all behavioral and physiological ANOVAs Greenhouse-Geisser epsilon (ε) correction was used when necessary. Original df values and corrected p values are reported. Behavioural data was analyzed in Statistica 7.0.

Electro-myography (EMG) recording and pre-processing

EMG was measured by EMG110C amplifiers using an MP150 data acquisition unit (Biopac Inc.). Two disposable cloth-based hypoallergenic Ag-AgCl EL504 recording disc electrodes were connected by 110S shielded touch-proof leads. Active electrodes were placed along the left and right flexors of the thumb (flexor pollicis brevis). An electrode on the left elbow served as ground. Before electrode application the skin was washed with soap, gently abraded and washed with alcohol. The electrodes were attached by adhesive solid gel. EMG was sampled at 1000Hz, band-pass filtered between 10-250 Hz (Fridlund and Cacioppo, 1986), rectified and scaled relative to the maximum amplitude measured in each individual (Lehman and McGill, 1999). Hence, EMG is expressed as percent of the maximum value measured. EMG was also baseline-corrected relative to the -100 to 0 ms interval preceding stimulus presentation. EMG epochs extended from -100 to 998 ms relative to stimulus presentation.

Electro-myography data analysis

EMG data was analyzed in Matlab 7.1 and Statistica 7.0. The deviation of EMG amplitude from zero was tested by point-by-point one-sample t-tests run against zero for

each Congruency × Difficulty condition. Deviations from zero were considered significant if they reached significance over a minimum of 20 consecutive sampling points at $p<0.025$. A Congruency × Difficulty ANOVA was run on the mean EMG amplitude of the interval found to show significant deviations from zero. Single trial analysis was used to assess the relationship of EMG amplitude and RT. First, EMG trials were sorted as a function of RT.

In a single trial analysis individual trials were pooled across all subjects for all Congruency × Difficulty conditions. The EMG amplitude measured at individual sampling points of each trial was correlated with the RT measured in each particular trial. Point-by-point correlations were considered significant if they appeared for at least 20 consecutive sampling points (20 ms) at $p<0.025$. In order to avoid that outliers drive correlations, trials with RTs longer than the overall mean RT plus 3 standard deviations were removed from correlational analyses. A further single-trial analysis examined whether the activation of the correct response hand temporally followed the activation of the incorrect response hand, or whether the activation of response hands temporally overlapped. This question was analyzed in 1588 trials collected in the difficult incongruent condition where response conflict was the most robust. The activation difference between the correct and incorrect response hands was characterized by the difference of the peak latency of the largest amplitude sampling point in single-trials (measured in trials smoothed by a 50 ms running average window). In the incorrect response hand the peak latency and amplitude were measured between 100-400 ms, during the interval where significant incorrect response hand activation was found. A larger interval was used than the exact 214-378 ms interval with significant incorrect

hand activation so that single trial variability could be captured more successfully. However, analyses were also re-run on data measured when finding peaks was not constrained to the 100-400 ms interval. In the correct response hand the peak latency and amplitude were defined as the latency/amplitude of the most positive EMG peak found during the whole analysis epoch (in the correct hand EMG peak amplitude tightly correlated with RT).

Event-related brain potential recording and pre-processing

@ Figure 2

EEG was recorded by an Electrical Geodesics system with a 129-channel Hydro-Cell Net. Landmark electrode positions are depicted in **Fig. 2**. (All electrodes are shown in **Supplementary Figure 1.**). The sampling rate was 500 Hz, an on-line band-pass filter of 0.01-70 Hz was used. The data was band-pass filtered between 0.01-30 Hz offline, and was recomputed to average reference. Epochs extended from -100 to 998 ms relative to stimulus presentation. Data was baseline corrected relative to the -100 to 0 ms interval. Epochs containing voltage deviations exceeding ±100 µV relative to baseline at any of the recording electrodes and epochs containing ocular artefacts (visually detected by the experimenters at electrodes below, above and next to the eyes) were rejected.

Event-related potential analysis

The LRP was computed as proposed by Gratton et al. (1988):

[(ER - EL) LEFT HAND response + (EL - ER) RIGHT HAND response] / 2,

where EL denotes the amplitude of the ERP at an electrode placed over the left motor cortex, and ER denotes the amplitude of the ERP at an electrode placed over the right motor cortex. In the traditional 10-20 electrode system electrode C3 is used as EL

and electrode C4 is used as ER. Hydro-Cell Net electrode 36 has equivalent position to electrode C3 and Hydro-Cell Net electrode 104 has equivalent position to C4. Hence, electrode 36 was used as EL, and electrode 104 was used as ER (see **Fig. 2**). According to convention a negative LRP indicates a correct response tendency, and a positive LRP indicates an incorrect response tendency. The deviation of the LRP from baseline was tested by point-by-point two-tailed one-sample t-tests run against zero. Effects were considered significant when they reached significance at $p < 0.025$ over a minimum of 10 consecutive sampling points (20 ms).

The overall temporal course of congruency effects was illustrated by the global field power (GFP). The GFP is computed as the mean potential deviation of all recording electrodes, and it reflects the spatial standard deviation of the data (Lehmann and Skrandies, 1980; Skrandies, 1995). A large GFP is computed when ERPs show high peaks and troughs and steep potential gradients simultaneously on several electrode channels. Hence, the GFP is an excellent method for summarizing robust ERP effects appearing at many electrodes in a single curve. Importantly, the GFP characterizes the latency of robust distributed ERP effects by a single curve.

The peak latency of the P300 wave was determined between 300-600 ms. The peak latency was defined as the sampling point with the most positive amplitude on centro-parietal electrodes 7, 129, 106, 31, 80, 54, 55, 79, 61, 62, 78, 67, 72, and 77. These electrodes were chosen because the maximum amplitude of the P300 happened at these electrodes. The peak latency of the occipital P100 wave (sampling point with the most positive amplitude) was determined between 70-140 ms. The peak latency of the occipital N200 wave (sampling point with the most negative amplitude) was determined between

120-220 ms. The peak latency of the P100 and N200 was measured on electrodes 65, 66, 70, 68, 69, 73, 83, 84, 90, 88, 89 and 94 where these waves showed the largest amplitude. The peak latency of ERP waves was tested by Speed × Congruency × Physical distance × Electrode ANOVAs.

Effects in ERP amplitude were first examined by point-by-point Speed (fast vs. slow trials) × Congruency (congruent vs. incongruent) × Numerical distance (1 vs. 7) repeated-measures ANOVAs. There was no effect of numerical distance, nor were any interactions found. Hence, further analysis continued with a point-by-point Speed (fast vs. slow trials) × Congruency (neutral, congruent and incongruent) × Difficulty (small vs. large) repeated-measures ANOVAs. In order to protect against Type-I errors a conservative significance level of $p<0.005$ was chosen for EEG analysis. Time intervals where statistical effects reached significance ($p<0.005$) over a minimum of 10 consecutive sampling points at least at 6 electrode channels were considered to demonstrate significant effects. Pair-wise comparisons between all congruency conditions were carried out by point-by-point Speed (fast vs. slow trials) × Congruency × Size difference (small vs. large) repeated-measures ANOVAs. These ANOVAs included only two levels of congruency (1: neutral vs. congruent. 2: neutral vs. incongruent. 3: congruent vs. incongruent.). Pair-wise comparisons were restricted to electrodes and time-intervals where the overall ANOVA with three levels of congruency identified congruency effects. Pair-wise comparison effects were considered significant when effects reached significance ($p<0.005$) over a minimum of 10 consecutive sampling points at least at 6 electrode channels. The topography of congruency effects was visualized as congruent minus neutral, incongruent minus neutral, and incongruent minus

17

congruent difference potentials. Speed × Congruency × Difficulty × Electrode ANOVAs were run on the mean amplitude of ERPs in time intervals and at the electrodes where point-by-point ANOVAs discovered interactions. These ANOVAs were run separately for electrodes showing effects of opposite polarity (usually due to phase reversal between anterior vs. posterior electrodes). This avoids that opposite polarity effects cancel out in an overall ANOVA.

Brain electric source analysis was done with the BESA method implemented in EEGLab 6 (Delorme and Makeig, 2004). The analysis was run on the mean topography of the incongruent condition between 330-400 ms. The analysis was run blind, that is, no initial dipole was seeded. Rather, the algorithm found a single dipole best fitting the scalp distribution.

Results

Behavioral results

@ Table 1

Behavioral results are shown in **Table 1**. There were 3.4% more correct responses in the congruent and neutral than in the incongruent conditions ($F(2,34)=20.71$; $\varepsilon=0.569$; $p=0.0001$. Tukey congruency contrasts for congruent vs. incongruent: $p=0.0001$; Neutral vs. incongruent: $p=.0001$). Accuracy was 2.5% worse in the difficult than in the easy condition ($F(1,17)=22.33$; $p<0.0002$). Incongruent trials were responded less accurately in the difficult than in the easy condition (Congruency×Difficulty: $F(2,34)=18.82$; $\varepsilon=0.640$; $p=0.0001$).

RT was 30 ms slower in the difficult than in the easy condition ($F(1,20)=100.7$; $p=0.0001$). There was a congruency effect ($F(2,34)=78.49$; $\varepsilon=0.767$; $p=0.0001$).

Incongruent trials were responded 25 ms slower than neutral trials (Tukey test: p<0.0001.) and 30 ms slower than congruent trials (p<0.0001). The incongruent vs. neutral discrepancy was larger in the difficult than in the easy condition (32 vs. 20 ms; Congruency×Difficulty interaction: $F(2,34)=4.47$; $\varepsilon=0.994$; $p=0.0189$). However, the incongruent vs. neutral difference was equally strong in both the difficult and in the easy conditions (Tukey p from Congruency×Difficulty contrasts for both levels of difficulty: $p=0.0001$). The Congruency × Difficulty × Numerical distance × Side of response ANOVA with two levels of congruency replicated results of the ANOVAs with 3 levels of congruency.

@ Table 2

Correctly responded trials in each condition were sorted into slow and fast categories in each subject. The proportion of fast and slow trials is shown in **Table 2**. RT for fast and slow trials is shown in **Table 1C-D.** Fast trials were responded 104 ms faster than slow trials (Speed: $F(1,17)=145.39$; $p<0.0001$). There was a Congruency effect ($F(2,34)=28.14$; $\varepsilon=0.901$; $p<0.0001$) and a Speed × Congruency interaction ($F(2,34)=14.13$; $\varepsilon=0.952$; $p<0.0001$). According to Speed × Congruency post-hoc Tukey contrasts there were no congruency effects in fast trials. In contrast, in slow trials the RT was slower in the incongruent than in the neutral (Tukey $p<0.0001$) and congruent ($p<0.001$) conditions. The effect of Difficulty ($F(1,17)=37.45$; $p<0.0001$) and the Speed × Difficulty interaction ($F(1,17)=23.58$; $p=0.0001$) were significant. Post-hoc tests demonstrated that difficulty had an effect in slow trials only (Tukey p from Speed × Difficulty contrasts: $p=0.0002$).

<u>Group-level electro-myography results</u>

@ Figure 3

@ Table 3

The EMG recorded in the correct and incorrect response hands is shown in **Fig. 3A.** The EMG of the correct response hand robustly deviated from the baseline during the response period (see time intervals in **Table 3A**). In addition, in the incongruent condition there was significant EMG activity in the incorrect response hand between 222-322 ms. In the difficult condition the incorrect response-hand activity appeared between 214-334 ms, in the easy condition it appeared between 229-294 ms. In a secondary analysis the mean EMG amplitude was determined in each Congruency $\times$ Difficulty condition between 222-322 ms (**Fig. 3B.**). There was a congruency effect ($F(2,34)=18.18$; $\varepsilon=0.793$; $p<0.0001$). Post-hoc Tukey tests (congruent vs. incongruent: $p=0.001$; neutral vs. incongruent: $p=0.0001$) revealed that amplitude was larger in incongruent (4.9%) trials than in congruent (1.9%) and neutral trials (0.47%). Difficult trials showed larger amplitude than easy trials (3.74% vs. 1.12%; $F(1,17)=24.98$; $p<0.0001$). The difficulty effect appeared only in incongruent trials (Tukey p from Congruency $\times$ Difficulty contrasts: 0.0001; Congruency $\times$ Difficulty interaction ($F(2,34)=10.17$; $\varepsilon=0.716$; $p<0.002$).

@ Figure 4

EMG in conditions with significant deviations from baseline in slow and fast trials are shown in **Fig. 4.** Time intervals with significant EMG deviations in the correct response hand are summarized in **Table 3B-C.** In fast trials the EMG of the incorrect response hand never deviated significantly from baseline. In slow trials the EMG of the incorrect hand deviated from the baseline in the incongruent condition between 214-352

ms (difficult condition) and between 254-294 ms (easy condition). There were no significant EMG deviations in any other conditions. The mean EMG amplitude of the 254-294 ms interval was entered into a Speed × Congruency × Difficulty ANOVA. This interval was chosen because EMG in both the easy and difficult incongruent conditions significantly deviated from the baseline during this interval. Nevertheless, results were confirmed by other ANOVAs run on the mean EMG amplitude measured between 214-352 ms and 222-322 ms.

@ Figure 5

Mean amplitudes between 254-294 ms and 95% confidence intervals are shown in **Fig. 5.** Slow trials had larger amplitude than fast trials (5.1% vs. 0.45%; $F_{(1,17)}=11.82$; p=0.003). There was a congruency effect ($F_{(2,34)}=8.93$; $\varepsilon=0.937$; p=0.001). Post-hoc Tukey tests revealed that incongruent trials (4.85%) had larger amplitude than neutral (0.67%) and congruent (2.85%) trials (p<0.0006 for both incongruent vs. neutral/congruent contrasts). Difficult trials had larger amplitude than easy trials (3.89% vs. 1.68%; $F_{(1,17)}=16.26$; p=0.008). Post-hoc tests showed that no congruency effects were significant in fast trials. In contrast, amplitude in slow incongruent trials was larger than in slow neutral and slow congruent trials (Tukey p from Speed × Congruency contrasts: 0.0001 for both; Speed × Congruency: $F_{(2,34)}=6.10$; $\varepsilon=0.870$; p=0.008). In addition, the amplitude was larger in the slow congruent condition than in the slow neutral condition (p=0.0346).

The effect of task difficulty was significant only in slow trials (Tukey p from Speed × Difficulty contrasts: 0.0002; Speed × Difficulty: $F_{(1,17)}=18.63$; p=0.005). The difficulty effect was significant only in the incongruent condition but not in other

conditions (Tukey p from Congruency × Difficulty contrasts: 0.0003; Congruency × Difficulty: $F(2,34)=5.51$; $\varepsilon=0.734$; $p=0.0168$). There was a triple, Speed × Congruency × Difficulty, interaction. Post-hoc Speed × Congruency × Difficulty Tukey contrasts revealed that slow incongruent difficult trials had larger amplitude than any other conditions ($p<0.0002$ for all contrasts). In addition, amplitude in the slow congruent difficult condition was larger than amplitude for both levels of difficulty in the neutral condition ($p<0.002$ for both contrasts).

Single-trial electro-myography results

@ Figure 6

In order to assess the relationship of EMG amplitude and RT single trial analysis was used. **Fig. 6.** shows trials in the difficult incongruent condition sorted as a function of RT. Mere visual inspection reveals that slower trials were systematically accompanied with larger EMG amplitude in the incorrect response hand during the time interval where significant incorrect response hand activation was detected at the group-level (surrounded by continuous vertical lines in **Fig. 6.**). Using single-trial data, EMG amplitude was correlated with the RT at each sampling point. In the difficult incongruent condition significant positive correlations were found between incorrect response hand EMG amplitude and RT between 277-378 ms: The larger was the EMG amplitude in the incorrect hand the longer was the RT (surrounded by dashed vertical lines in **Fig. 6.**). There were no reliable correlations between incorrect hand EMG amplitude and RT in other conditions. There were significant correlations between correct response hand EMG amplitude and RT in all conditions (see **Table 4**). Negative correlations between correct response hand EMG amplitude and RT appeared between approximately 200-380 ms.

That is, the larger was the EMG amplitude in the correct response hand, the shorter was the RT.

@ Table 4

Further analyses were done on the 1588 trials collected in the difficult incongruent condition in order to see whether the activation of correct and incorrect response hands temporally coincided with each other in individual trials, or rather, the activation of the correct response hand followed the activation of the incorrect response hand. In the incorrect response hand the EMG peak amplitude strongly correlated with the RT ($r=0.4683$; $p<0.0001$). This confirms that the peak amplitude detected in the incorrect response hand provided a measure closely related to performance (in the correct hand there was weaker correlation because peak amplitude values were more stable across trials: $r=0.1781$; $p<0.0001$). The activation difference between correct and incorrect response hands was characterized by the difference of the peak latency of the sampling point with the largest amplitude in EMG epochs. The mean and median temporal difference between the EMG peak latency in correct and incorrect hands was 62.3 ms and 50.5 ms respectively (standard deviation: 76.1 ms).

The distribution of correct hand minus incorrect hand EMG peak latency difference values in single trials is shown in Insert 1 in **Fig. 6.** The cumulative distribution is shown in Insert 2 in **Fig. 6.** In 48.17% of trials the correct hand minus incorrect hand EMG peak latency difference ranged between -60 ms and +50.5 ms (21.09% of trials showed a temporal lag of -60<lag<0 ms). In 29 trials (1.83% of trials) the temporal difference was smaller than -60 ms (the negative value means that correct hand EMG peaked earlier than incorrect hand EMG). In the remaining 50% of trials the

temporal difference ranged between +50.5 and +230 ms except in two trials with the longest RT (843 ms) where the difference was 331 and 451 ms. The above shows that in about half the trials correct and incorrect hand activation very tightly overlapped.

If we further consider that EMG activity was already well noticeable during the interval preceding and following the EMG peak latency, than the proportion of trials with overlapping correct and incorrect hand activation is even larger. For example, a 5% EMG value seems to be a reasonable activation threshold because at this level EMG deflection from baseline was already well noticeable. With regard to this threshold, the amplitude of correct hand EMG exceeded 5% during the time interval ±80 ms around the peak latency in the fastest trials and ±100-150 ms around the peak latency in the slowest trials. When considering a ±100-150 ms range around the peak latency, about 70-80% of trials can be considered to show overlapping activity in the correct and incorrect response hands (73.92% of trials showed a temporal activation difference of -60<lag<100 ms; and 81.35% of trials showed a temporal activation difference of -60<lag<150 ms). When finding the largest peak amplitude was not constrained to 100-400 ms in the incorrect hand, the data changed little, both the mean and the median of the correct-incorrect hand time difference was 51 ms (standard deviation: 149 ms). 46.41% of trials had time difference between -70 ms and +50 ms (68.38%: -70<lag<100 ms; 76.70%: 70<lag<150 ms).

Another question is how the temporal activation difference between correct and incorrect response hands relates to RT. First, the EMG peak latency in the correct hand had practically perfect correlation with RT (r=0.9914; p<0.0001), and the EMG peak latency in the incorrect hand also strongly correlated with RT (r=0.5760; p<0.0001).

Second, correct – incorrect hand peak latency difference values sorted according to the RT are shown in Insert 3 in **Fig. 6.** The correct – incorrect peak latency difference was strongly correlated with RT (r=0.8716; p<0.0001). The 29 trials with smaller difference than -60 ms and the two trials with larger difference than +230 ms were excluded from the above correlation in order to avoid that outliers drive correlations (these trials are visible at the sides of the distributions shown in Insert 3 in **Fig. 6).** Nevertheless, correlations were significant when including these trials as well (r=0.8679; p<0.0001). Further, in order to check whether correlation results were affected by constraining finding the EMG peak amplitude to the 100-400 ms time interval, correlations were also run on data where EMG peaks could be found without any constraints. The correlations remained practically unchanged (r=0.8728; p<0.0001 both when including all trials and when excluding trials with outlier temporal difference values.

Lateralized readiness potential

@ Table 5

As expected, the RT was associated by large negative stimulus and response-locked LRP deflections in all conditions (The stimulus-locked LRP is shown in **Supplementary Figure 2.**, the response-locked LRP is shown in **Supplementary Figure 3**.). The timing of the stimulus-locked LRP was in perfect agreement with correct hand EMG. Intervals where the LRP signaled significant correct response preparation are shown in **Table 5**. The stimulus-locked LRP showed brief incorrect response activation between 228-248 ms (11 sampling points) in the slow difficult incongruent trials only. The response-locked LRP showed brief incorrect response activation between -298 ms to -284 ms (8 sampling points) and shortly thereafter between -262 ms to -248 ms (8

sampling points) in fast difficult incongruent trials. Neither the stimulus-locked, nor the response-locked LRP signaled incorrect response activation in any other conditions.

Latency of event-related brain potentials

@ Figure 7

ERPs on the vertex electrode are shown in **Fig. 7A.** (Further electrodes are shown in **Supplementary Figure 4.**). The time-course of ERPs is summarized by the GFP in **Fig. 7B-D.** The robust GFP peak between 250-600 ms coinciding with the P300 is well visible in these figures. Importantly, the GFP belonging to the congruent condition peaks markedly faster from about 270 ms than the GFP in the congruent and neutral conditions. The mean peak latency of the P300 is shown in **Table 6.** The P300 peaked 45 ms earlier in fast than in slow trials (383 vs. 428 ms; $F(1,17)=22.99$; $p<0.0001$). The P300 peaked 13 ms earlier in the easy than in the difficult condition (399 vs. 412 ms; $F(1,17)=12.23$; $p=0.0027$). There was a congruency effect (congruent: 398 ms; incongruent: 415 ms; neutral: 403 ms; $F(2,34)=7.53$; $\varepsilon=0.689$; $p=0.0065$). Post-hoc Tukey tests revealed that the congruent vs. incongruent ($p=0.0017$) and the incongruent vs. neutral ($p=0.0342$) differences were significant. There were no significant interactions. There were no significant main effects or interactions on the peak latency of the occipito-parietal P100 and N200 waves.

@ Table 6

Amplitude and topography of event-related brain potentials

A first analysis showed no effect of numerical distance in the data. Second, a Speed $\times$ Congruency $\times$ Difficulty ANOVA was run on the amplitude of ERPs. This analysis revealed that there were speed main effects over several electrodes between 140-

170 ms and 280-380 ms and 460-650 ms (see GFP in **Fig. 7C.** and topography in **Supplementary Figure 5.**). Later speed effects can be explained by the latency shift of the P300 in slow relative to fast trials. The main effect of difficulty appeared between 360-420 ms and 560-720 ms (see GFP in **Fig. 7C.** and topography in **Supplementary Figure 6.**). Congruency main effects appeared in four consecutive intervals between 180-210, 270-300, 330-400, and 560-610 ms (see GFP in **Fig. 7B.**). The topography of congruency effects is depicted in **Fig. 8.** Between 180-210 ms ERPs in the incongruent condition were more negative than in the neutral condition over left occipital electrodes. Between 270-300 ms ERPs in the congruent condition were more positive than in the neutral condition over several occipito-parietal electrodes. This congruent vs. neutral difference resulted in more negative ERPs in the incongruent than in the congruent conditions. Between 330-400 ms ERPs in the incongruent condition were more negative than in the neutral condition over central electrodes. There was a similar effect in the congruent condition with an additional effect of left occipital ERPs being more positive in the congruent than in the neutral condition. Between 560-610 ms central ERPs were more positive in the incongruent than in the neutral condition.

@ Figure 8

Congruency × Difficulty interactions appeared over several electrodes between 410-440 ms (see GFP in **Fig. 7D.,** topography and graphs in **Supplementary Figure 7.**). The mean amplitude of this interval on electrodes showing significant effects was entered into an overall Speed × Congruency × Difficulty × Electrode ANOVA. The ANOVA for frontal electrodes showed a Congruency × Difficulty interaction ($F(2,34)=15.03$; $\varepsilon=0.947$; $p<0.0001$). Post-hoc Congruency × Difficulty Tukey contrasts revealed that the

difficulty effect was significant in the congruent (p=0.0001) and incongruent (p=0.001) conditions but not in the neutral condition (p=0.7). The ANOVA for posterior electrodes showed a Congruency × Difficulty interaction (F(2,34)=9.88; ε=0.858; p=0.0009). Post-hoc Congruency × Difficulty Tukey contrasts revealed that the Difficulty effect was significant in the congruent (p=0.0067) but not in the incongruent (p=0.2) and neutral conditions (p=0.3). There were no other significant interactions on the amplitude of ERPs.

The congruency effect over central electrodes between 330-400 ms was of special interest because previous studies (Liotti et al. 2000; West et al. 2003) attributed an incongruent vs. congruent ERP difference with similar timing, topography and polarity to the conflict resolution activity of the ACC. Liotti et al. (2000) used source analysis. Therefore, for the sake of comparability across studies and to check the plausibility of the ACC as the neural generator of the congruency effect we ran blind dipole source analysis on our high spatial density 129-channel ERP data. The mean topography (330-400 ms) of the incongruent condition was fed into the BESA algorithm implemented by EEGLab (**Fig. 9.**). All dipoles were found in the ACC. When using all trials a dipole was found in the dorsal? ACC in Brodmann area 24 (MNI coordinates: x=-2; y=-3; z=33; residual variance: 12.34%). When using only quickly responded trials a dipole was found in the ACC in Brodmann area 33 (x=-7; y=14; z=25; residual variance: 18.77%). When using only slowly responded trials a dipole was again found in the ACC in Brodmann area 24 (x=2; y=-11; z=36; residual variance: 23.18%).

@ Figure 9

Discussion

In this study our objective was to demonstrate response conflict directly in the effectors and to disentangle ERP markers of stimulus and response conflict. To this end we combined EMG and ERP in a numerical Stroop task. EMG enabled us to measure response hand activation and response conflict, directly. EMG findings were in line with the predictions of the motor response conflict theory (Morton and Chambers, 1973; Posner and Snyder, 1975). Correct and incorrect response hands were activated in parallel and robust incorrect response hand activation happened in the incongruent condition. Incorrect response activation happened only in trials that had a slow response. That is, incorrect response hand activation was related to a reduction in performance. Significant incorrect response hand activation happened only in incongruent difficult (between 214-352 ms) and incongruent easy (254-294 ms) trials. This indicated that incorrect response hand activation was not due to a general non-specific EMG activity increase in all slow trials. Rather, it was related to a mechanism highly specific to the incongruent condition.

In theory, group-level EMG effects may have been the consequence of increased EMG activity in a relatively small number of outlier trials. However, single-trial analysis (from which trials with outlier RT were excluded) demonstrated that group-level EMG effects appeared consistently across all trials. By sorting trials according to RT it became visible that the onset/offset and duration of EMG activity was closely related to RT in the difficult incongruent condition. This consistent EMG activity resulted in the significant group-average EMG findings. Further, single-trial EMG amplitude of the incorrect response hand positively correlated with single-trial RT between 200-380 ms: when incorrect response hand EMG activity increased, the RT became longer. Similarly, the peak amplitude of EMG in the incorrect response hand strongly correlated with RT.

These observations confirm the direct relationship between incorrect response hand activation and performance.

The parallel activation of correct and incorrect response hands attests two points; first, both task-relevant and task-irrelevant stimulus dimensions were processed up to the level of response preparation. This fits the predictions of the continuous flow model of information processing (Eriksen and Schultz, 1979). Second, in consequence to the above point, when the task-irrelevant dimension delivered conflicting information with the task-relevant dimension, incorrect motor activity was prepared. As suggested by the motor response conflict theory, the preparation of the incorrect response hand could, in principle, explain the performance decrease in the incongruent condition (Morton and Chambers, 1973; Posner and Snyder, 1975). Our data is in strong agreement with a fNIRS (DeSoto et al. 2001) and a fMRI (Cohen-Kadosh et al. 2007) study[1] testing for concurrent activation of motor cortices related to correct and incorrect responses. Similarly, some of our previous LRP studies also demonstrated substantial incorrect response hand activation in Stroop tasks (Szűcs, Soltész, Jármi and Csépe, 2007; Szűcs et al. 2009).

Besides establishing that partial processing of perceptual information was able to prepare incorrect response activity, a further question regards the dynamics of this process (DeSoto et al. 2001). Is correct and incorrect response hand activation temporally coincident, or rather, correct hand activation follows incorrect hand activation? Our single trial data provides robust evidence for concurrent correct and incorrect response activation in about 50-70% of trials. In another 30-50% of trials

---

[1] It is to note that the course temporal resolution of fMRI studies does not allow discriminating between motor cortex activities appearing before and after the button press response; and also does not allow deciding whether correct/incorrect motor activity is concurrent (De Soto et al. 2001).

correct hand activation followed incorrect hand activation. Concurrent response activation happened in trials with fast RT, while correct hand activation followed incorrect hand activation by increasingly longer time-lags in trials with longer RT. At the same time incorrect hand activation reached significance at group-level only in slow, but not in fast trials. This suggests that there may be two alternative scenarios of incorrect response activation. First, in slow trials with strong incorrect response activation (large EMG amplitude in the incorrect hand) the correct response may be strongly inhibited and therefore activated relatively late (large temporal difference between correct and incorrect hand EMG peak latency). In contrast, in trials with weaker incorrect response activation the incorrect response cannot inhibit the correct response effectively which results in concurrent correct and incorrect response activation.

As expected, there was stronger response conflict in the difficult than in the easy incongruent condition. This appeared in the form of Congruency $\times$ Difficulty and Speed $\times$ Congruency $\times$ Difficulty interactions in EMG data. First, there was no significant incorrect hand activation and there were no difficulty effects in fast trials. Second, incorrect hand activation was longer and the amplitude was larger in slow difficult incongruent trials than in slow easy incongruent trials. Correspondingly, correlations between RT and incorrect hand EMG amplitude appeared only in the difficult incongruent but not in the easy incongruent condition, nor in any other conditions. Third, while the deviation in EMG amplitude, from baseline, did not reach significance for any time interval in the congruent condition, the EMG amplitude was larger in slow difficult congruent trials than in slow easy and difficult neutral trials. There was no similar difference between the easy congruent and the neutral conditions. This, again, suggests

that increasing task-difficulty is likely to increase incorrect response tendencies, even in the congruent condition. Overall task-difficulty seems to have resulted in increased incorrect response tendency in the incongruent and congruent conditions but not in the neutral condition.

We suggest that the increasing perceptual difficulty resulted in enhanced incorrect hand activity because the slower processing of the relevant stimulus dimension provided more time for the processing of the task-irrelevant dimension to advance and to influence motor preparation (Szűcs and Soltész, 2007). Further, we speculate that increased task-difficulty may have resulted in the appearance of incorrect response tendencies in the congruent condition, because the mere presence of discrepant numerical values may have activated potential alternative, but incorrect response tendencies in some congruent trials. In contrast, in the neutral condition, the numerical dimension did not deliver differential information. Hence no incorrect response tendencies appeared in the neutral condition. Previously it has been suggested that response conflict appears only when task-difficulty is low (Cohen-Kadosh et al. 2007). Besides pointing out at several difficulties with this conclusion earlier (Szűcs and Soltész, 2008), here we also provide direct EMG evidence for increased response conflict when task-difficulty is high, relative to when it is low.

Whereas our EMG data provides evidence for response conflict, the P300 data also provides evidence for stimulus conflict in the numerical Stroop task. Robust congruency effects were detected not only in RT but in P300 peak latency as well. These P300 congruency effects replicate our previous findings (Szűcs, Soltész, Jármi and Csépe, 2007, adult group). In contrast to present results two studies could not find congruency effects in the latency of the P300 in vocal and manual versions of the classic

color-word Stroop paradigm (Duncan-Johnson and Koppel, 1981; Ilan and Polich, 1999). Importantly, the power of previous studies seems to be adequate when considering the number of participants and the number of trials in each congruency condition (Duncan-Johnson and Koppel, 1981: approximately 106 trials per condition; 12 subjects; vocal task; Ilan and Polich: 64 trials per condition; 32 subjects, manual task.). On the one hand this suggests that P300 effects may differ across different Stroop tasks. For example, the size/number incongruent condition may evoke stronger stimulus conflict than the color/word incongruent condition, because physical size may be a more salient property than color. This speculation is supported by the fact that unlike in the classic color-word Stroop task, congruency effects are bidirectional in the numerical Stroop task (Henik and Tzelgov, 1982; for review see Szűcs and Soltész, 2007). On the other hand, our EMG findings demonstrate that not only stimulus but also response conflict played a role in our Stroop task. This is in line with the conclusions of Duncan-Johnson and Koppel (1981) and Ilan and Polich (1999).

Our EMG and P300 peak latency data confirms the frequent assumption that both stimulus and response conflict may contribute to the Stroop effect (van Veen and Carter 2005; Liston et al. 2006). We suggest that the continuous flow model of information processing (Eriksen and Schultz, 1979) is able to explain the findings. Partially processed task-irrelevant stimulus information can result in stimulus conflict and at the same time can also prepare incorrect motor activity. In the incongruent condition, the parallel processing of the task-irrelevant stimulus dimension slowed down perceptual processing (congruency effects on P300 latency) and the partially processed irrelevant stimulus information also resulted in significant incorrect response hand activation as predicted by

the continuous flow model (Eriksen and Schultz, 1979). Alternatively, in the congruent condition the parallel processing of the task-irrelevant stimulus dimension resulted in faster stimulus processing (P300 latency) and probably contributed to the enhanced amplitude of correct response hand EMG. This suggests that congruency effects appearing in brain imaging parameters may indeed contain a mixture of effects related to both stimulus and response-conflict.

The amplitude and topography of ERPs showed a series of congruency effects. The topography and timing of effects were aligned with our previous study using a similar Stroop task (physical comparison task in Szűcs and Soltész, 2007). Importantly, no congruency effects appeared in the latency of the occipital P1 and N2 waves which suggests that congruency effects in amplitude were not due to early latency differences between conditions. In contrast to this, the latency of the P300 wave clearly contributed to congruency effects. Notably, the earlier peak latency of the P300 can explain the enhanced positivity in the congruent vs. neutral condition between 270-300 ms. This is exactly the time interval when the faster onset of the P300 in the congruent, than in the neutral condition, is most visible (see **Fig. 7B.**). Similarly, the faster offset of the P300 in the congruent, than in the neutral condition, can also explain why central electrodes showed more negative voltage in the congruent than in the neutral condition between 330-400 ms.

However, between 330-400 ms there was a very clear amplitude decrease (negativity) in the incongruent relative to the congruent condition over central electrode sites which cannot be explained by latency shifts of the P300. This negativity appeared in both fast and slow trials right after the offset of significant incorrect hand EMG activity.

The negativity closely resembled the N450 effect in timing, polarity, amplitude and topography (Liotti et al. 2000; West, 2003; Szűcs and Soltész, 2007; 2008). On the one hand, considering our data, at least two arguments go against a classical N400 explanation of the effect. First, the effect seems to offset earlier than a typical N400 response, even when comparing to single-digit identity matching tasks (Szűcs, Soltész, Czigler and Csépe, 2007). Secondly, there was no numerical distance effect in this time range in our data. This suggests that as long as ERP data could detect there was no evidence that the effect was sensitive to semantic relationships. This does not fit a N400 explanation (Kutas and Hillyard, 1980). However, the alternative ACC explanation offers a clear prediction about the dominant neural source of the N450.

Previously Liotti et al. (2000) localized the source of the N450 in the ACC using 64-channel EEG data. In order to check consistency with the results of Liotti et al., we have carried out blind single dipole source localization on our data. It is to emphasize that the accuracy of EEG source localization results is unreliable because of the inverse problem of EEG, and therefore must be interpreted cautiously. In our data the blind source localization algorithm found dipoles in the ACC, in all analyses. The coordinates of dipoles corresponded to the ACC activations identified by fMRI studies that employed Stroop tasks (for a review see Barch et al. 2001). This provides some convergent evidence for the view that the source of the congruency effect is the ACC (Liotti et al. 2000; West, 2003). Traditionally, ACC involvement in Stroop tasks has been interpreted as an expression of the response conflict detection/resolution. However, the lack of Congruency $\times$ Speed interactions in our data goes against the response conflict detection/resolution explanation. Response conflict was stronger in slow than in fast

trials. Hence, a Speed × Congruency interaction in ERP amplitude could have also been expected if the N450 had really been related to response conflict. While it is possible that interactions were missing because of lack of power, another possibility is that the N450 is not related to response conflict. For example, a recent fMRI study has shown that ACC activity can increase even when there is no response conflict, and concluded that the ACC may be involved in the selective activation of correct responses, rather than in response conflict detection/resolution per se (Roelofs, Turennout and Coles, 2006). In summary, it remained unclear whether the N450 was related to stimulus or response conflict. However, as a working hypothesis we assume that its source is indeed the ACC.

Difficulty main effects appeared in ERP amplitude between 360-420 ms, exactly at the peak of the P300. The P300 is known to be sensitive to task-difficulty effects (Donchin 1981), and therefore difficulty effects may reflect the modulation of the P300. The timing of the difficulty effect is in-line with the typical timing of the N400 component in digit-matching tasks (Szűcs, Soltész, Czigler and Csépe, 2007). Another possibility is that the difficulty effect may be a genuine N400. This explanation would require that the task context led subjects to interpret physical magnitude semantically, in a general sense of magnitude. However, the lack of a numerical distance effect in this time range goes against the N400 explanation. Hence, difficulty main effects were probably due to task-difficulty effects on the P300. Importantly, it is unlikely that increased difficulty in the incongruent, relative to the neutral, condition could also explain the congruency effects (N450) between 330-400 ms. Firstly, the difficulty main effect peaked at 400 ms, right at the offset of the congruency effect. Secondly, the difficulty effect was driven not only by the incongruent but also by the congruent

condition. Thirdly, there was no Congruency × Difficulty interaction during the congruency main effect. An interaction should have appeared if congruency effects would have been due to task difficulty alone (The incongruent condition was more difficult than the congruent condition, and this should have interacted with the main effect of task difficulty). The above suggest that the main effect of task difficulty was markedly different from the main effect of congruency.

Difficulty × Congruency interactions appeared right after the P300 peak (410-440 ms). Most probably, these effects were not directly related to response conflict as the ERP amplitude followed a different pattern than the Difficulty × Congruency interaction revealed by EMG. Rather, the interaction may be due to varying subjective difficulty in different congruency conditions. For example, this may be a consequence of the lack of parallel differential stimulus processing in the neutral condition relative to the congruent and incongruent conditions, as well as the contrasting, overall difficulty level of the congruent and incongruent conditions. Finally, it is noteworthy that several previous studies used only incongruent vs. congruent contrasts when examining the topography of congruency effects (Liotti et al. 2000; West 2003). Our data suggests that this may not be optimal; for example, the congruent vs. incongruent differences between 270-300 and 330-400 ms are qualitatively different from each other (P300 latency difference vs. genuine amplitude difference) while being relatively similar in topography. Consequently, omitting the use of a neutral condition may result in confusing qualitatively different ERP effects.

Methodologically, our data illustrates both the potential and the limitations of the LRP. The LRP does not provide hand-specific measures of response activation as it is

computed from volume-conducted ERPs, which are inevitably influenced by voltages emanating from both the left and right motor cortices. Consequently, when incorrect and correct response activation signals (voltages) from motor cortices ipsilateral/contralateral to the response hand are of similar size, signals cancel each other out. Then, if the correct response activation signal is of much larger amplitude than a temporally coincident incorrect response activation signal, the LRP will show correct response activation even if incorrect response activation genuinely occurred. The above can explain why the stimulus-locked LRP showed brief (20 ms duration), incorrect response activation only in slow difficult incongruent trials: the temporal difference between the activation of correct and incorrect response hands was the largest in these trials. Hence, the incorrect response activation in the LRP was not cancelled out by concurrent correct response activation. In contrast, in fast trials correct and incorrect response hand activation temporally coincided and the LRP effects cancelled out. It is well demonstrated in Fig. 6A., that unlike correct response hand activation, incorrect response hand activation did not tightly follow RT, but rather the incorrect response hand was activated during a relatively constant time-frame between 200-400 ms. This meant that response-locked LRPs had a smaller chance of detecting incorrect response activation in slow trials because the incorrect response activation was not time-locked to the RT.

In our data the weak LRP effects are in sharp contrast with the robust incorrect response hand activation detected by EMG. This exemplifies that the LRP does not have the sensitivity to detect incorrect response activation, nor to detect the true duration of incorrect response activation under certain circumstances because it is not a hand-specific measure of response activation (Gratton et al. 1988). Certain techniques, e.g. time-

locking LRP to the onset of EMG activity could enhance the sensitivity of the LRP (Masaki et al. 2000), however, analyzing single-trial EMG may provide a more robust method for identifying incorrect response activation than the LRP. Our data provides empirical support for our previous argument (Szűcs and Soltész, 2008), that it is invalid to draw conclusions from the *lack* (null finding) of LRP incorrect response activation about the non-existence of incorrect response activation. It also follows that incorrect response activation can happen even when the LRP shows a negative deflection (Szűcs and Soltész, 2008). However in such cases LRP data cannot provide clear evidence for incorrect response activation.

Conclusion

In summary, we have shown that incorrect response hand activation happens in a numerical Stroop task. The effect was related to performance, highly specific to the incongruent condition, and appeared consistently across all trials. These findings provide direct evidence for motor response conflict in a Stroop task. Incorrect and correct response hand activation was concurrent in 50-80% of the incongruent difficult trials. In 20-50% of trials correct hand activation followed incorrect hand activation by a lag of 60-230 ms (EMG peak latency to peak latency difference). This suggests that there may be two scenarios of incorrect response activation, probably depending on the balance of activation/inhibition between correct and incorrect response hands. P300 results suggest that not only response, but also stimulus conflict played a role in our Stroop task. Differences in ERP amplitude between the congruent and neutral conditions can be explained by the earlier peak of the P300 in the congruent than in the neutral condition. In contrast, ERPs were genuinely more negative in the incongruent than in the neutral

condition between 330-400 ms (N450 effect). The N450 was probably a correlate of the activity of the ACC (Liotti et al. 2000; West 2003). Overall, the parallel timing of correct and incorrect EMG activity supports the continuous flow model of information processing. In fact, this model can explain both stimulus and response conflict effects; the partially processed task-irrelevant stimulus information not only resulted in stimulus conflict but also prepared incorrect motor activity. Methodologically, we provide evidence that the LRP is an imperfect measure of incorrect response hand activation and that null findings in LRP cannot be taken to indicate that no incorrect response hand activation happened.

References

Aine CJ, Harter MR. 1984. Hemipsheric differences in event-related potentials to Stroop stimuli. Ann N Y Acad Sci. 425:154-6.

Atkinson CM, Drysdale KAD, Fulham WR. 2003. Event-related potentials to Stroop and reverse Stroop stimuli. *International Journal of Psychophysiology*, 47:1-21.

Barch DM, Braver TS, Akbudak E, Conturo T, Ollinger J, Snyder A. 2001. Anterior cingulate cortex and response conflict: Effects of response modality and processing domain. *Cerebral Cortex*, 11:837-848.

Botvinick M, Nystrom LE, Fissel K, Carter SC, Cohen JD. 1999. Conflict monitoring versus selection for action in anterior cingulate cortex. *Nature*. 402:179-181.

Botvinick M, Cohen JD, Carter CS. 2004. Conflict monitoring and anterior congulate cortex: an update. *Trends in Cognitive Sciences*. 8:539-546.

Burle B, Possamai CA, Vidal F, Bonnet M, Hasbroucq T, 2002, *Psychological Research*. 66:324-336

Cohen-Kadosh R, Cohen-Kadosh K, Linden DEJ, Gevers W, Berger A, Henik A. 2007. The brain locus of interaction between number and size: A combined functional magnetic resonance imaging and event-related brain potential study. *Journal of Cognitive Neuroscience*. 19:957-970.

Coles MGH, Gratton G, Bashore TR, Eriksen CW, Donchin E. 1985. A psychophysiological investigation of the continuous flow model of human information processing. *Journal of Experimental Psychology: Human Perception and Performance.* 11:529-553.

Delorme A, Makeig S. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Meth.* 134:9-21.

De Jong R, Wierda M, Mulder G, Mulder LJM. 1988. Use of partial stimulus information in response processing. *Journal of Experimental Psychology: Human Perception and Performance.* 14:682-692.

DeSoto MC, Fabiani M, Geary DC, Gratton G. 2001. When in doubt, do it both ways: Brain evidence of the simultaneous activation of conflicting motor responses in a spatial Stroop task. *Journal of Cognitive Neuroscience.* 13:523-536.

Donchin, E. (1981), Surprise!...Surprise? *Psychophysiology*. 18:493-513.

Duncan-Johnson CC, Koppel BS. 1981. The Stroop effect: Brain potentials localize the source of interference. *Science*. 214:938-940.

Eriksen CW, Coles MGH, Morris LR, O'Hara WP (1985), An electromyographic investigation of response competition. *Bulletin of the psychonomic society*. 23:165-168.

Eriksen CW, Schultz DW. 1979. Information processing in visual search: A continuous flow conception of experimental results. *Perception and Psychophysics*, 25:249-263.

Fridlund AJ, Cacioppo JT. 1986. Guidelines for human electromyographic research. *Psychophysiology*, 23:567-589.

Gratton E, Limkeman M. 1983. A continuously variable frequency cross-correlation phase fluorometer with picoseconds resolution. *Biophysics Journal*. 44:315-324.

Gratton G, Coles MGH, Sirevaag EJ, Eriksen CW, Donchin E. 1988. Pre- and Poststimulus Activation of Response Channels: A Psychophysiological Analysis. *Journal of Experimental Psychology: Human Perception and Performance* 14:331-344.

Hasbroucq T, Possamai CA, Bonnet M, Vidal F. 1999. Effect of the irrelevant location of the response signal on choice reaction time: an electromyographic study in man. *Psychophysiology*. 36:522-526

Henik A, Tzelgov J. 1982. Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory and Cognition*, 10:389-395.

Hock H. S., and Egeth, H. 1970. Verbal interference with encoding in a perceptual classification task. *Journal of experimental Psychology,* 83:299-303.

Ilan AB, Polich J. 1999. P300 and response time from a manual Stroop task. *Clinical Neurophsyiology*. 110:367-373.

Kerns JG, Cohen JD, MacDonald AW, Cho RY, Stenger VA, Carter CS. 2004. Anterior Cingulate Confict Monitoring and Adjustments in Control, *Science*, 303:1023-1026.

Kutas M, Hillyard SA. 1980. Reading senseless sentences: Brain potentials reflect semantic incongruity. Science. 207:203-205.

Kutas M, McCarthy G, Donchin E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation in time. *Science.* 197:792-795.

Lehman GJ, McGill SM. 1999. The importance of normalization in the interpretation of surface electromyography: A proof of principle. *Journal of manipulative and physiological therapeutics*. 22:444-446.

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

Liotti M, Woldorff MG, Perez R, Mayberg HS, 2000. An ERP study of the temporal course of the Stroop color–word interference effect. Neuropsychologia. 38:701–711.

Liston C, Matalon S, Hare TA, Davidson MC, Casey BJ. (2006), Anterior cingulate and posterior parietal cortices are sensitive to dissociable forms of conflict in a task-switching paradigm. *Neuron*, 50:643-653.

MacLeod CM. 1991. Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109:163-203.

Masaki H, Takasawa N, Yamazaki K. 2000. An electrophysiological study of the locus of the interference effect in a stimulus-response paradigm. *Psychophysiology*. 37:464-472.

McCarthy G, Donchin E. 1981. A metric for thought: a comparison of P300 latency and reaction time. *Science*. 211:77-80.

Milham MP, Banich MT, Webb VB, Cohen NJ, Wszalek T, Kramer AF. 2001. the relative involvement of anterior cingulate processes in attentional control depends on the nature of conflict. *Cognitive Brain Research*, 12:467-473.

Morton J, Chambers SM. 1973. Selective attention to words and colours. *Quarterly Journal of Experimental Psychology,* 25:387-397.

Niedeggen M, Rösler F. 1999. N400 effects reflect activation spread during retrieval of arithmetic facts. *Psychological Science*, 10:271-276.

Posner M. 1978. *Chronometric explorations of mind*. Lawrence Erlbaum Associates.

Posner MI, Snyder CRR. 1975. Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: TheLoyola symposium* Hillsdale, NJ: Erlbaum. 55-85.

Rebai M, Bernard C, Lannou J. 1997. The Stroop's test evokes a negative brain potential, the N400. *Int. J. Neurosci*. 91:85–94.

Roelofs A, Turennout M, Coles MGH. 2006. Anterior cortex activity can be independent of response conflict in Stroop-like tasks. *Proc Natl Acad Sci*. 103:13884-13889.

Skrandies W. 1995. Visual information processing: topography of brain electrical activity. *Biological Psychology*. 40:1-15.

Smid HGOM, Mulder G, Mulder LJM. 1990. Selective response activation can begin before stimulus recognition is complete: a psychophysiological and error analysis of continuous flow. *Acta Psychologica*. 74:169-201.

Stroop JR. 1935. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*. 18:643-662.

Szűcs D, Csépe V. 2005. The effect of numerical distance and stimulus probability on ERP components elicited by numerical incongruencies in mental addition. *Cognitive Brain Research.* 22:282-300.

Szűcs D, Soltész F, Czigler I, Csépe V. 2007. Electroencephalography effects to semantic and non-semantic mismatch in properties of visually presented single-characters: the N2b and the N400. *Neuroscience Letters*. 412:18-23.

Szűcs, D. and Soltész, F. 2007. Event-related potentials dissociate facilitation and interference effects in the numerical Stroop paradigm**.** *Neuropsychologia.* 45:3190-3202.

Szűcs, D. and Soltész, F. 2008. The interaction of task-relevant and task-irrelevant stimulus features in the number/size congruency paradigm: an ERP study. *Brain Research,* 1190:143-158.

Szűcs D, Soltész F, Bryce D, Whitebread D. 2009. Real-time tracking of motor response activation and response competition in a Stroop task in young children: A lateralized readiness potential study. *Journal of Cognitive Neuroscience*, In Press.

Szűcs D, Soltész F, Jármi É, Csépe V. 2007. The speed of magnitude processing and executive functions in controlled and automatic number comparison in children: an electro-encephalography study. *Behavioral and Brain Functions,* 23:1-20.

van Veen V, Carter SC. 2005. Separating semantic conflict and response conflict in the Stroop task: A functional fMRI study. *Neuroimage*. 27:497-504.

van Veen V, Cohen JD, Botvinick MM, Stenger VA, Carter S. 2001. Anterior cingulate cortex, conflict monitoring and levels of processing. *Neuroimage*, 14:1302-1308.

West R. 2003. Neural correlates of cognitive control and conflict detection in the Stroop and digit location tasks. *Neuropsychologia.* 41:1122-1135.

West R, Alain C. 1999. Event-related neural activity associated with the Stroop task. *Cogn. Brain Res*. 8:157-164.

TABLES AND CAPTIONS

Table 1

Means and standard errors for accuracy (Acc.: %) and reaction time (RT: milliseconds) by Congruency $\times$ Difficulty in all trials and in fast and slow trials.

Congruency:	Neutral		Congruent		Incongruent	
Task difficulty:	Hard	Easy	Hard	Easy	Hard	Easy
A. All trials: Acc.	98 ± 1	99 ± 1	98 ± 1	99 ± 0	92 ± 3	98 ± 1
B. All trials: RT	460 ± 28	435 ± 21	457 ± 27	430 ± 24	492 ± 27	455 ± 24
C. Fast trials: RT	404 ± 13	402 ± 12	398 ± 13	396 ± 13	401 ± 13	401 ± 13
D. Slow trials: RT	506 ± 24	492 ± 20	513 ± 24	493 ± 20	524 ± 22	501 ± 21

Table 2

Proportions of fast and slow trials (%).

Congruency:	Neutral		Congruent		Incongruent	
Task difficulty:	Hard	Easy	Hard	Easy	Hard	Easy
A. Fast trials	43.4	61.0	54.1	67.2	30.3	49.0
B. Slow trials	32.3	18.3	29.4	18.9	52.2	32.2

Table 3

Time periods with significant EMG deviations in the correct response hand in milliseconds.

Congruency:	Neutral		Congruent		Incongruent	
Task difficulty:	Hard	Easy	Hard	Easy	Hard	Easy
A. All trials	216:888	214:828	208:876	200:894	206:910	208:870
B. Fast trials	216:646	212:734	204:653	200:668	218:662	204:650
C. Slow trials	274:880	270:772	260:826	278:712	254:932	268:872

FIGURES AND CAPTIONS

Congruent		Incongruent		Neutral	
Diff.	Easy	Diff.	Easy	Diff.	Easy
2 9	2 9	2 9	2 9	2 2	2 2
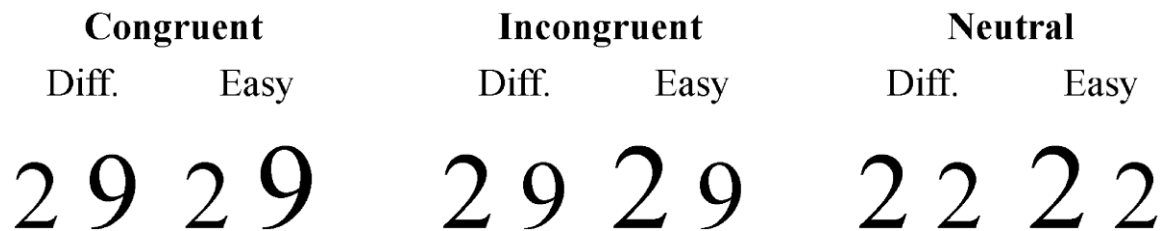
Figure 1

Example stimuli from the numerical Stroop task (Diff. = Difficult condition. Easy = Easy condition.). Participants have to press a button on the side where they see the physically larger number.
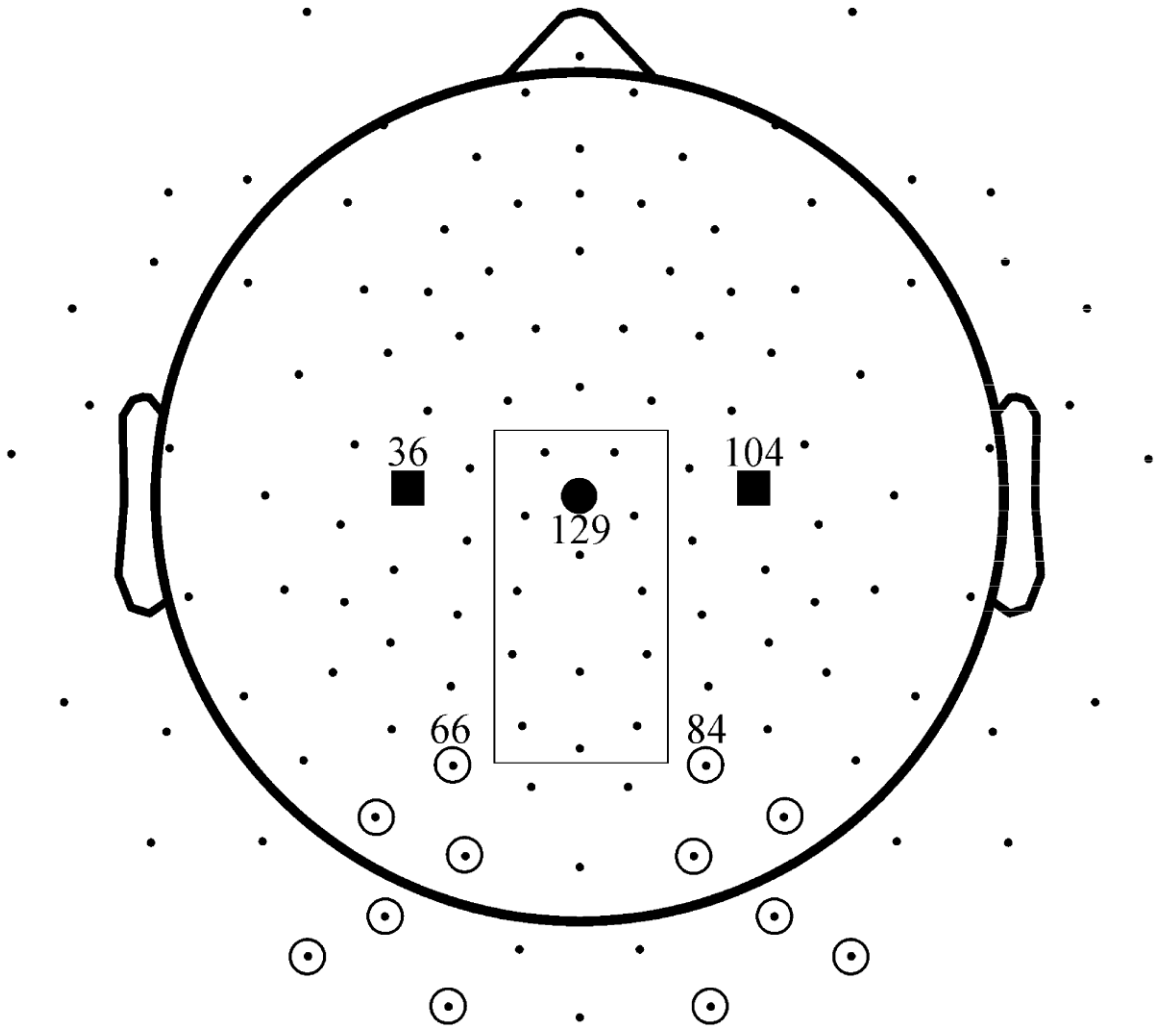
Figure 2

Schematic representation of the 129-channel Hydro-cell Net used in the experiment. Electrode 129 is at the vertex. Electrodes 36 and 104 were used for LRP computation. Electrodes in the bordered area were used for determining P300 peak latency. Electrodes marked by open circles were used for determining P100 and N200 peak latency.
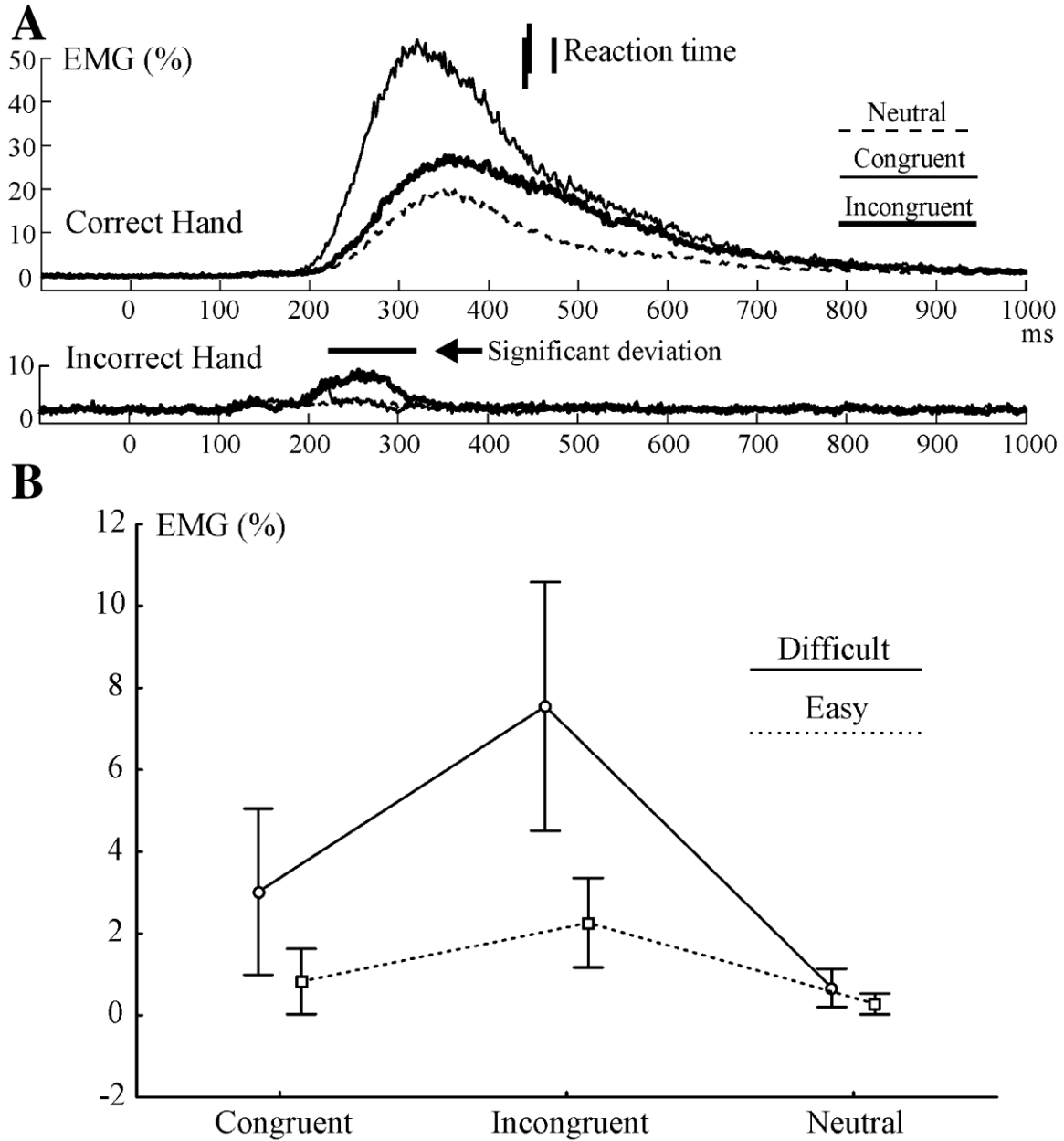
Figure 3

Group-average electro-myography (EMG) signal in correct and incorrect response hands.

(A) Time-course of EMG. Reaction time is marked for the congruent, neutral and

incongruent conditions. The horizontal line marks the interval where the EMG

significantly deviated from the baseline in the incorrect hand in the incongruent

condition. (B) Mean EMG amplitude between 222-322 ms. A 95% confidence interval is

shown.
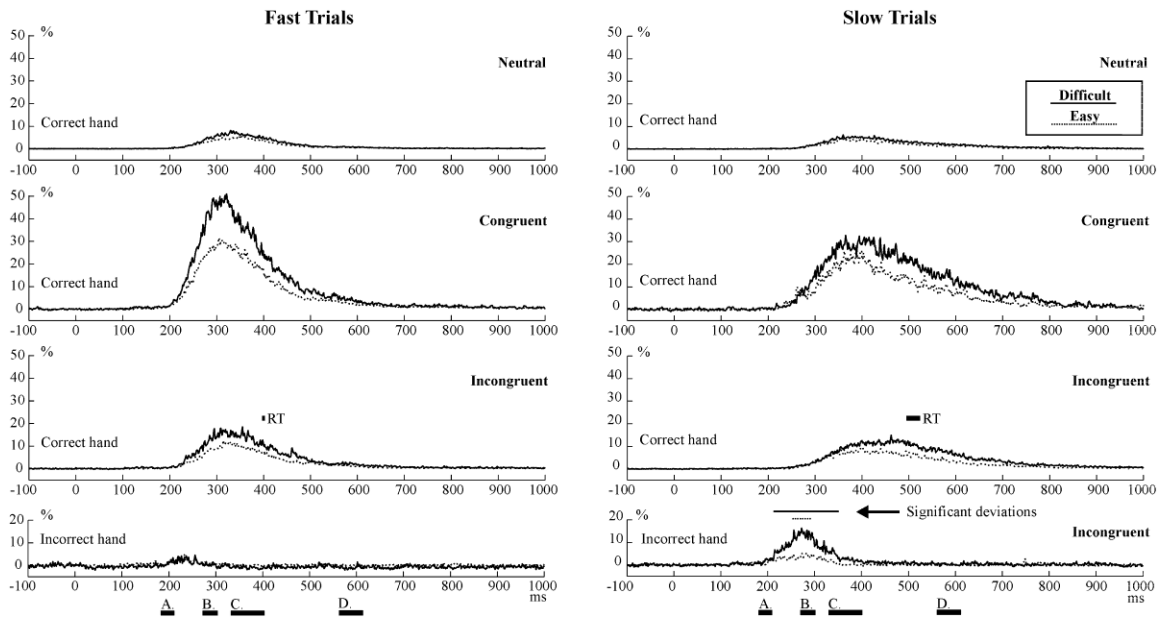
Figure 4

Group-average electro-myography (EMG) signal with significant deviations from baseline; in correct and incorrect response hands, for fast and slow trials. The range of mean reaction times (RT) in the 3 congruency conditions is marked in the correct hand/incongruent condition panels. In the bottom two panels horizontal lines, above the graphs, mark intervals where the EMG significantly deviated from the baseline in the incorrect hand in the easy and difficult incongruent conditions. Lines below the graphs, marked A-D, refer to time intervals with significant congruency effects as shown in Figure 8.

Fast Trials

Slow Trials
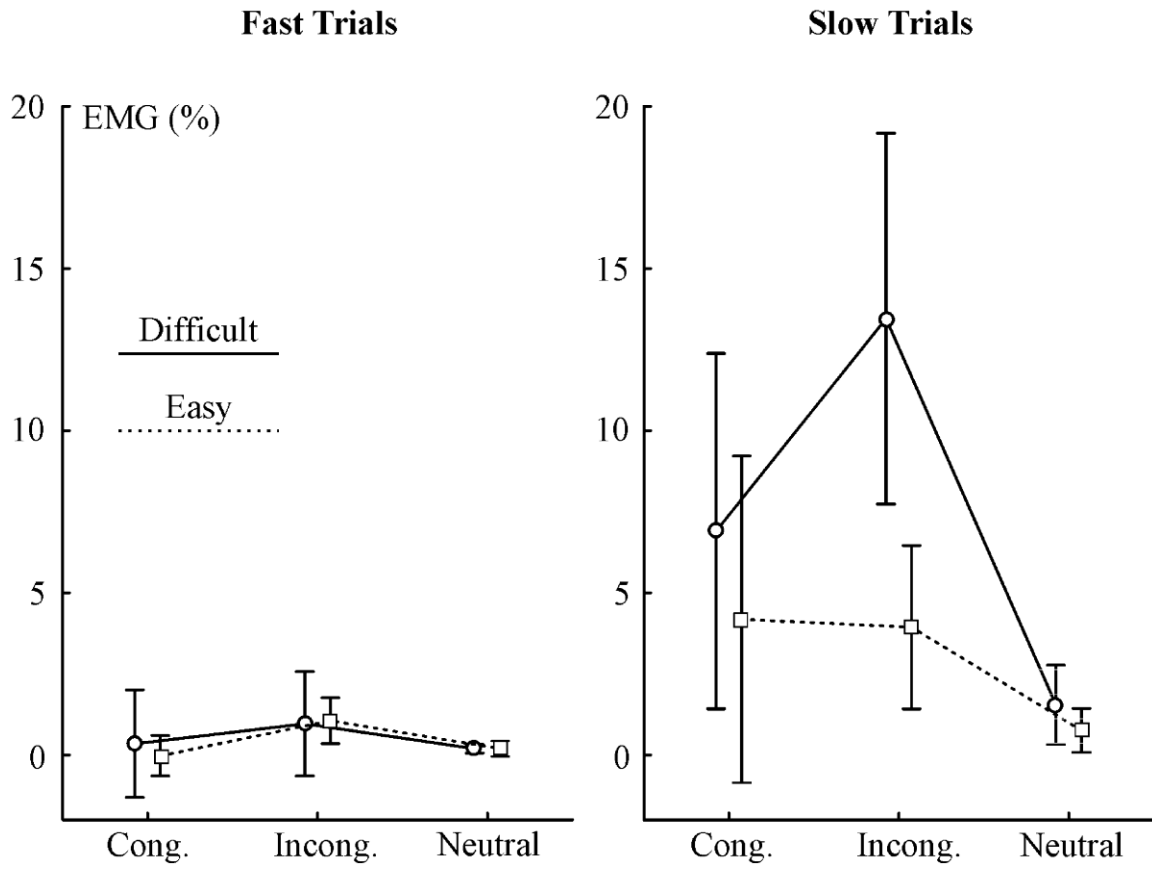
Figure 5

Mean group-average electro-myography (EMG) signal between 254-294 ms. A 95% confidence interval is shown.
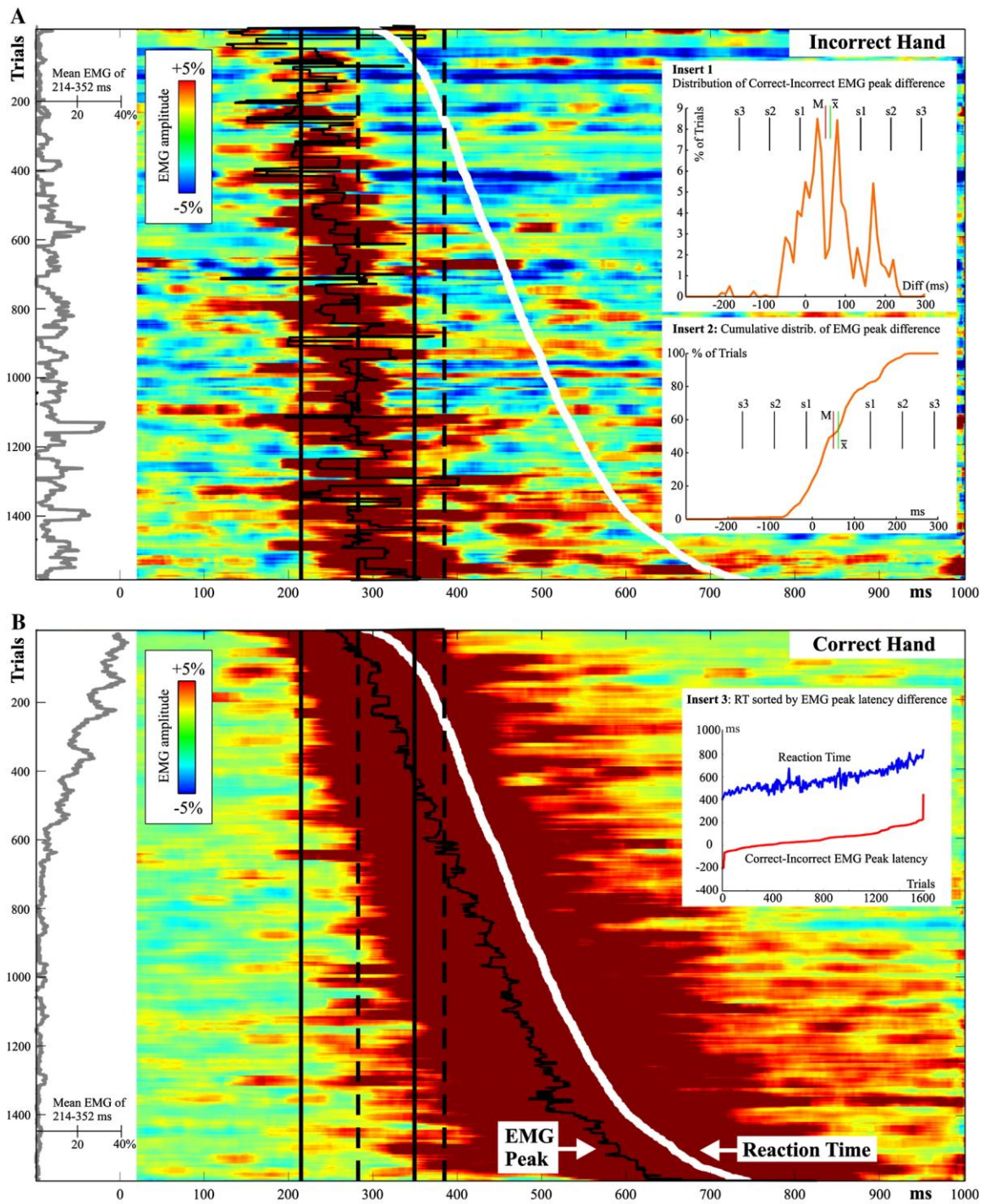
Figure 6

Single trial EMG data in the incorrect (A) and correct (B) response hands for the difficult incongruent condition. Trials are pooled for all subjects and sorted by reaction time.

Trials were smoothed by a 30-trial-wide sliding window and data points were smoothed with a 50-ms-wide sliding window. The white curve depicts reaction time. The black curve traces the peak amplitude of the EMG signal. Continuous black vertical lines surround the interval where significant group-level EMG activity was detected in the difficult incongruent condition (214-334 ms). Dashed vertical lines mark the interval where significant correlations appeared between single-trial EMG amplitude and reaction time. Insert 1 shows the distribution of the temporal difference (correct minus incorrect) of EMG peak latency in correct and incorrect hands (M: median; x: mean; s1-3: standard deviations). Insert 2 shows the cumulative distribution of the above temporal difference. Insert 3 shows reaction times sorted according to response hand EMG peak latency difference (correct minus incorrect). The side panels to the left show the mean EMG amplitude between 214-352 ms, in each trial.
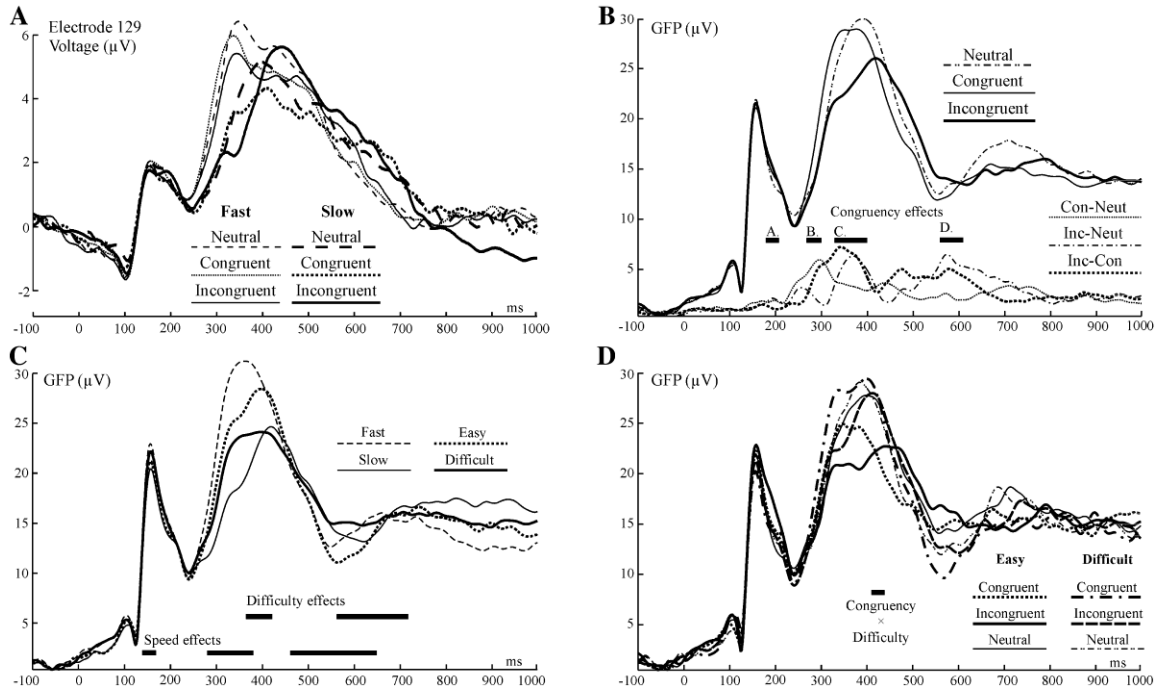
Figure 7

The Global Field Power (GFP) summarizes the time-course of the ERPs. (A) ERPs on electrode 129. (B) GFP in congruency conditions and GFP computed from; Congruent – Neutral (Con-Neut), Incongruent – Neutral (Inc-Neut) and Incongruent – Congruent (Inc-Con) difference potentials. Horizontal markers denote the time course of Congruency effects (see Fig. 8. for intervals A-D.). (C) GFP in fast vs. slow trials and in easy vs. difficult conditions. Horizontal markers denote Speed and Difficulty effects. (D) Congruency × Difficulty interactions. The horizontal marker denotes the interval with interactions.

56
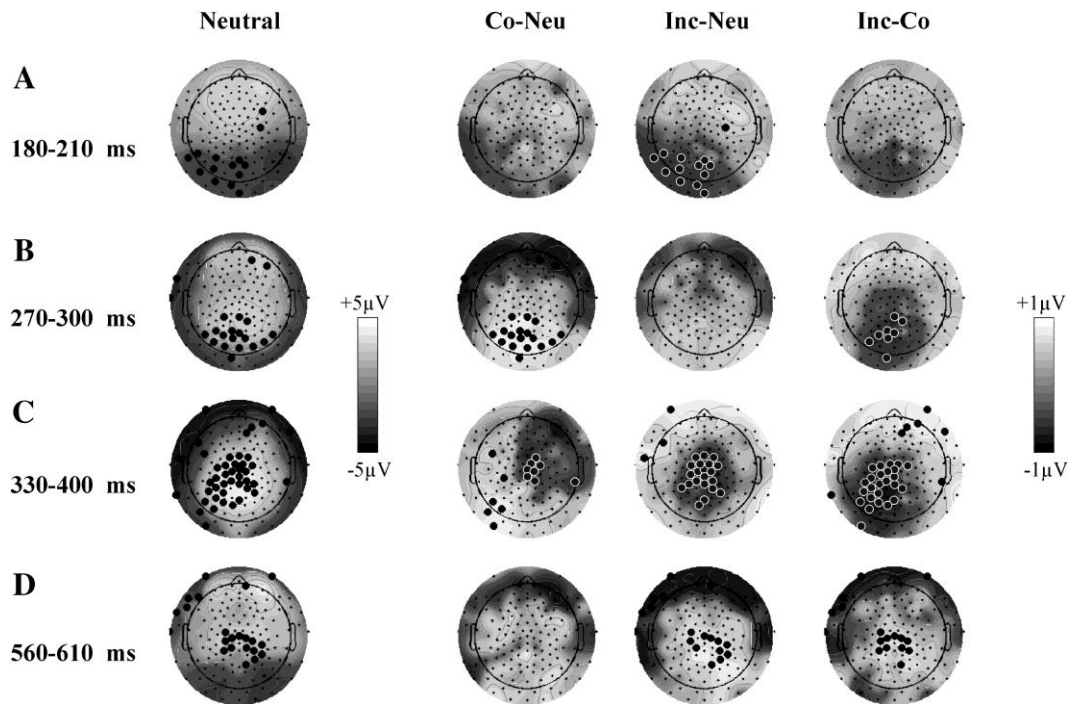
Figure 8

The main effect of congruency on the amplitude of ERPs. (A-D) Time intervals with significant congruency effects. Electrodes with significant effects (p<0.005) are represented by bold dots (N.B. Some dots have white contour lines merely for better visibility). The 'Neutral' column represents congruency effects from the main ANOVA with 3 levels of congruency. This column shows raw voltage in the Neutral condition (±5 µV). Other columns to the right represent pair-wise comparisons from ANOVAs with 2 levels of congruency (Co-Neu: Congruent vs. Neutral; Inc-Neu: Incongruent vs. Neutral; Inc-Co: Incongruent vs. Congruent). These columns show voltage from appropriate difference potentials (±1 µV).

All trials Fast trials Slow trials

12.34% 18.77% 23.18%

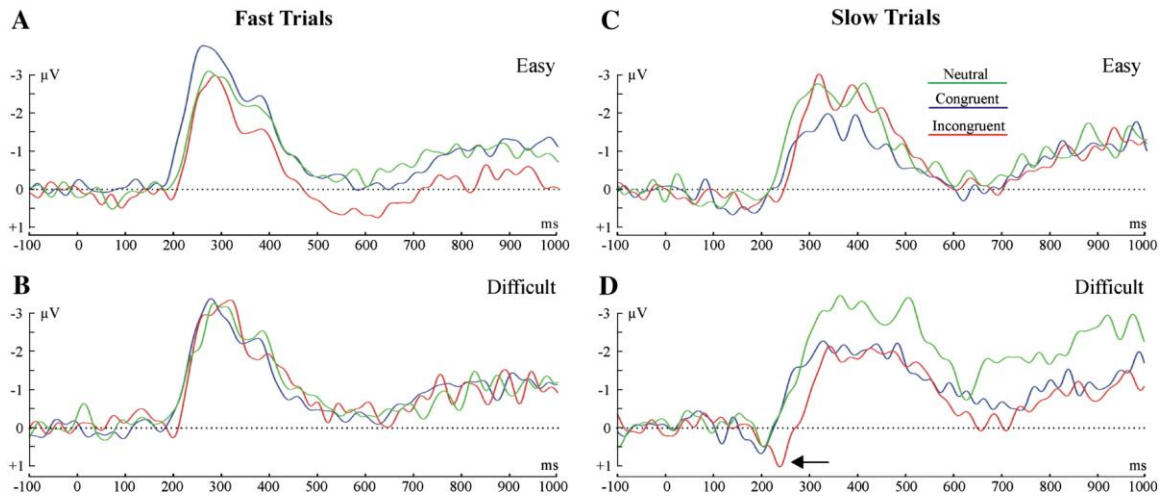x=-2 y=-3 z=33   x=-7 y=14 z=25   x=2 y=-11 z=36

-1μV ▮▮▮▮▮▮▮ +1μV
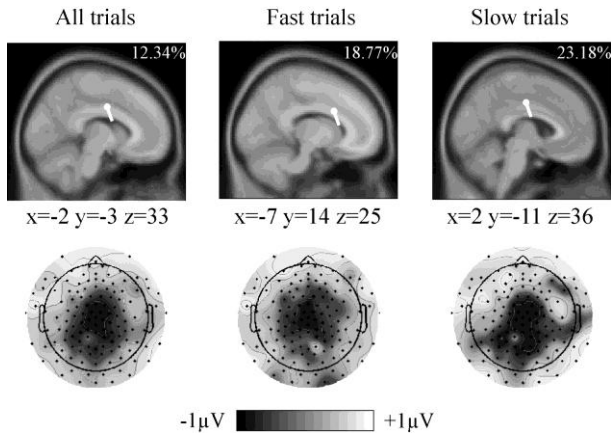
Figure 9

Source localization results for the incongruent condition. All trials: results for data containing all trials in the incongruent condition. Fast trials: results for data from fast trials only. Slow trials: results for data from slow trials only. The percentage at the right of brain images shows the residual variance. Below the images, the x, y and z values show MNI coordinates and scalp maps show the topography of data used for source localization.
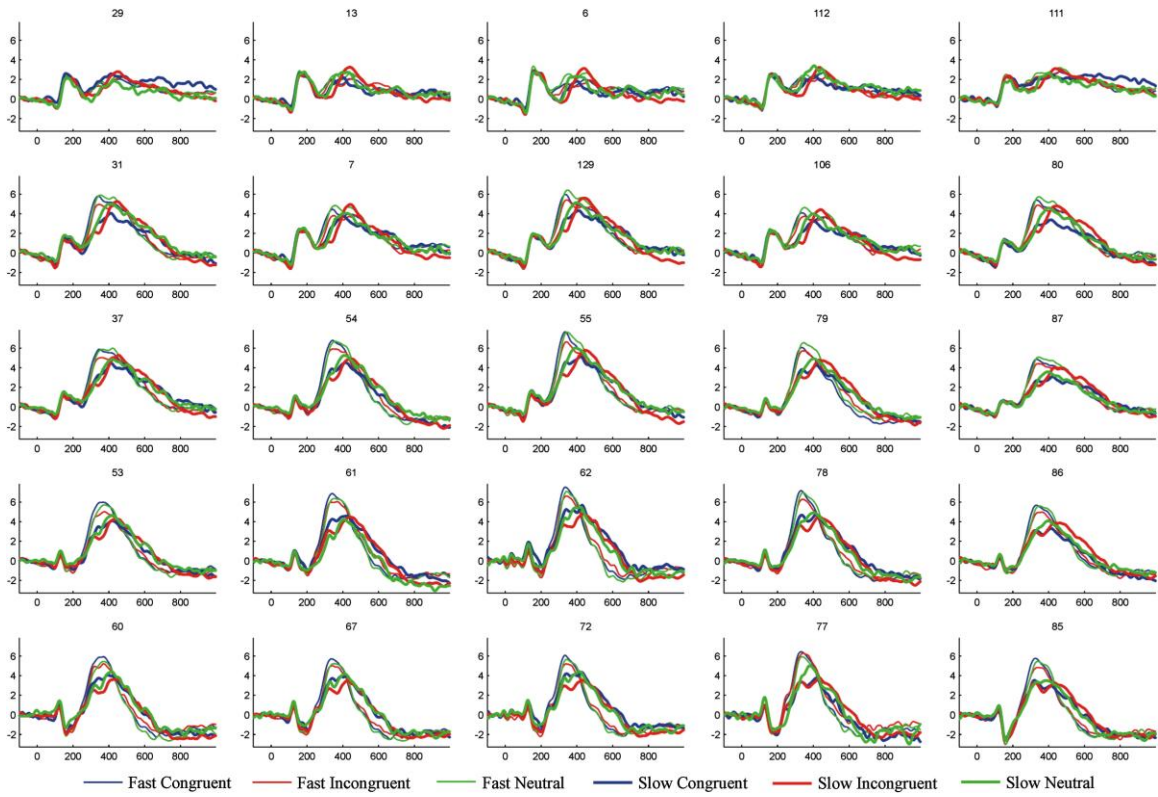
Supplementary Figure 1

The locations of all electrodes. ERPs at electrode locations marked by bold dots are shown in Supplementary Figure 4.

Supplementary Figure 2

The stimulus-locked Lateralized Readiness Potential in Easy (A) and Difficult (B) trials. The horizontal arrow points to the significant positive deflection of LRP, signalling incorrect response activation in the difficult incongruent condition (significant deviation during 228-248 ms).
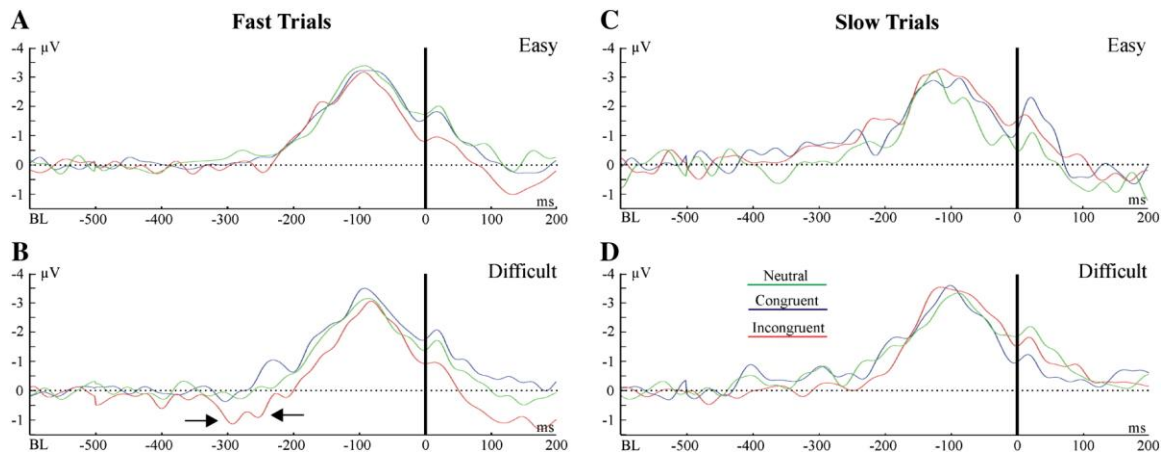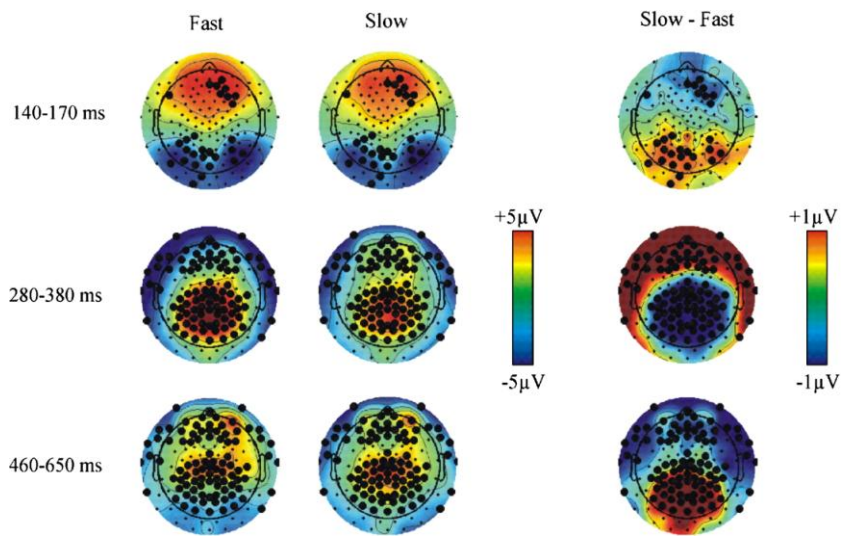
Supplementary Figure 3

The response-locked Lateralized Readiness Potential in Easy (A) and Difficult (B) trials. The horizontal arrow points to the significant positive deflection of LRP, signalling incorrect response activation in the easy incongruent condition (significant deviation during -298 to -284 ms and between -262 to -248 ms, relative to response).
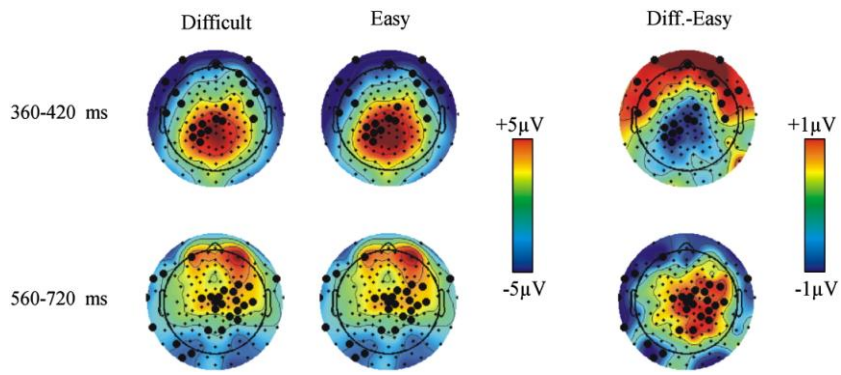
Supplementary Figure 4

ERPs at several electrode sites according to Speed × Congruency. For electrode positions see Supplementary Figure 1.
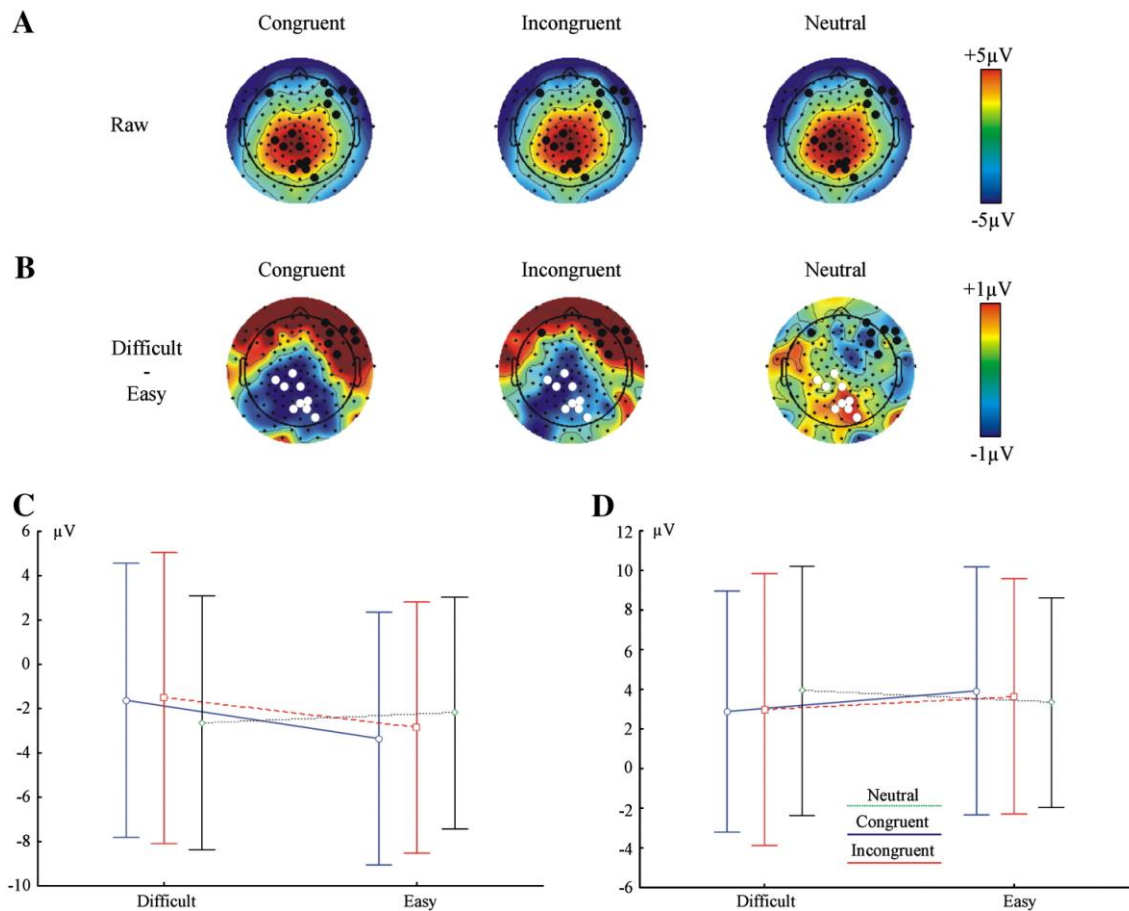
Supplementary Figure 5

Speed effects on the amplitude of ERPs. Bold dots represent electrodes with significant effects (p<0.005).

Supplementary Figure 6

Difficulty effects on the amplitude of ERPs. Bold dots represent electrodes with significant effects ($p<0.005$).

Supplementary Figure 7

Congruency × Difficulty interactions between 410-440 ms. (A) Topography of mean voltage in congruency conditions. (B) Topography of difficulty effects (difficult – easy) in congruency conditions. (C) Mean voltage at electrodes with negative overall means (marked by black bold dots in topographic maps). (D) Mean voltage at electrodes with positive overall means (marked by white bold dots in topographic maps).