Social Cognitive and Affective Neuroscience, 2015, 1577–1587

doi: 10.1093/scan/nsv050
Advance Access Publication Date: 28 April 2015
Original article

Emotion and goal-directed behavior: ERP evidence on cognitive and emotional conflict

Artyom Zinchenko,[1,2] Philipp Kanske,[3] Christian Obermeier,[2] Erich Schröger,[4] and Sonja A. Kotz[2,5]

[1]International Max Planck Research School on Neuroscience of Communication (IMPRS NeuroCom), Leipzig, Germany, [2]Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, [3]Department of Social Neuroscience, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, [4]Institute of Psychology, University of Leipzig, and [5]School of Psychological Sciences, University of Manchester, Manchester M13 9PL, UK

Correspondence should be addressed to Artyom Zinchenko, Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany. E-mail: zinchenko@cbs.mpg.de.

Abstract

Cognitive control supports goal-directed behavior by resolving conflict among opposing action tendencies. Emotion can trigger cognitive control processes, thus speeding up conflict processing when the target dimension of stimuli is emotional. However, it is unclear what role emotionality of the target dimension plays in the processing of emotional conflict (e.g. in irony). In two EEG experiments, we compared the influence of emotional valence of the target (emotional, neutral) in cognitive and emotional conflict processing. To maximally approximate real-life communication, we used audiovisual stimuli. Participants either categorized spoken vowels (cognitive conflict) or their emotional valence (emotional conflict), while visual information was congruent or incongruent. Emotional target dimension facilitated both cognitive and emotional conflict processing, as shown in a reduced reaction time conflict effect. In contrast, the N100 in the event-related potentials showed a conflict-specific reversal: the conflict effect was larger for emotional compared with neutral trials in cognitive conflict and smaller in emotional conflict. Additionally, domain-general conflict effects were observed in the P200 and N200 responses. The current findings confirm that emotions have a strong influence on cognitive and emotional conflict processing. They also highlight the complexity and heterogeneity of the interaction of emotion with different types of conflict.

Key words: emotion; cognitive control; conflict processing; cognitive conflict; emotional conflict

Introduction

Processing of emotional conflict is an inherent part of human social interactions. Common social phenomena such as irony and satire are good examples of emotional conflict, where emotion portrayed by the face and the voice do not coincide (e.g. in the remark ‘Great!’ in response to someone else’s failure). Emotional conflict is created between stimuli of different emotional valence (e.g. neutral face and angry voice). Processing conflict prolongs reaction times (RTs) and increases error rates (Stroop, 1935; Simon and Rudell, 1967; Egner and Hirsch, 2005; Ochsner *et al.*, 2009) reflecting increased competition for attentional resources by incongruent relative to congruent stimuli (West, 2003).

While recent studies have addressed the influence of emotion on cognitive conflict processing, the effect that emotionality of the target dimension can have on emotional conflict processing has not yet been investigated. Emotion facilitates the processing of cognitive conflict when a target stimulus is emotional (e.g. negative or positive; Kanske and Kotz, 2010, 2011a,c). For example, conflict processing is faster for emotional than for neutral stimuli (Kanske and Kotz, 2010), which also shows in a conflict-related N200 amplitude increase in

Received: 5 December 2014; **Revised:** 16 April 2015; **Accepted:** 24 April 2015

emotional trials. This suggests that conflict processing recruits additional resources via an emotionally salient target stimulus. The authors proposed that emotion triggers cognitive control thereby facilitating the processing of conflicting action tendencies (Kanske and Kotz, 2011b). In contrast, emotion impairs cognitive control when a non-target stimulus is emotional (Padmala *et al.*, 2011). For example, Hart *et al.* (2010) showed in a Stroop task that incongruent stimuli preceded by aversive pictures resulted in longer RTs relative to incongruent stimuli that were preceded by neutral pictures. To sum up, emotional stimuli influence cognitive control and facilitate cognitive conflict processing when they are part of the target dimension, while they impede conflict processing when they are non-targets, thus constituting particularly salient distractors (Kanske, 2012; Pessoa *et al.*, 2012).

The processing of emotional and cognitive conflict may differ at both the behavioral and neural levels (Egner *et al.*, 2008; Soutschek and Schubert, 2013; Torres-Quesada, *et al.*, 2014). The processing of emotional and cognitive conflict may differ at both the behavioral and neural levels (Egner *et al.*, 2008; Soutschek and Schubert, 2013; Torres-Quesada, *et al.*, 2014). For instance, Soutschek and Shubert (2013) tested whether cognitive and emotional conflict processing can be dissociated behaviorally. The authors examined the effects of a working memory and an emotional Go/Nogo task on executive control in an emotional and a cognitive Stroop task. The results support a behavioral dissociation: working memory load selectively suppressed conflict processing in the cognitive Stroop task, while an emotional Go/Nogo task selectively impaired conflict processing in the emotional Stroop task. Therefore, these findings argue in favor of domain-specific (cognitive, emotional) conflict processing mechanisms.

Furthermore, Egner *et al.* (2008) showed in a neutral and an emotional face Stroop task that both emotional and cognitive conflict processing resulted in comparable RTs, but domain-specific neural mechanisms were identified that were unique for either emotional or cognitive conflict (Egner *et al.*, 2008). In the emotional face Stroop task (Egner *et al.*, 2008), participants saw pictures of emotional faces with an overlay of either a congruent or incongruent emotional word. However, these studies could not address the question of how the emotionality of the target stimulus affects emotional conflict processing as stimuli were always emotional (Etkin *et al.*, 2006; Ochsner *et al.*, 2009). Therefore, this study aimed to test whether emotional conflict processing is influenced by the emotionality of a task-relevant target stimulus and to compare this effect with the influence of emotion on cognitive conflict processing. Furthermore, we were interested in the time course of cognitive and emotional conflict processing and how it varies as a function of the emotionality of the target dimension.

We devised an experimental protocol that varied the source of conflict between non-emotional (cognitive conflict) and emotional (emotional conflict) stimulus dimensions and that allowed the presentation of an identical stimulus for both types of conflict. Because of its precise temporal resolution, previous studies used event-related potentials (ERPs) to study the influence of emotion on conflict processing (Bruin and Wijers, 2002; Kanske and Kotz, 2011b). Here, we also measured ERPs to study early conflict-specific perceptual processes with regard to the two types of conflict. In two ERP studies, participants saw multisensory dynamic stimuli: short video clips of actors pronouncing different vocalizations in a negative or neutral tone of voice. Auditory and visual coding of the stimulus could be either congruent or incongruent with respect to vocalization (cognitive conflict task, Experiment 1) or emotion (emotional conflict task, Experiment 2). In Experiment 1 participants had to name the vowel pronounced by the voice (i.e. 'A' or 'O'), irrespective of the stimulus valence (emotional or neutral). In Experiment 2, participants were asked to report whether the voice was either emotional or neutral (i.e. the source of conflict was emotional). Importantly, while stimuli were identical in both tasks, emotionality was task-relevant (emotional conflict) or task-irrelevant (cognitive conflict) and the stimulus target dimension was either emotional or neutral.

This paradigm is unique among other tasks scrutinizing emotional conflict. First of all, it allows studying different types of conflict (cognitive, emotional), while keeping the stimulus dimension identical. Furthermore, it uses stimuli that have a higher ecological validity, relative to previously used face Stroop pictures (Egner *et al.*, 2008) as real-life conflict processing happens in a dynamic multisensory environment. The dynamic quality of information is important for emotion recognition (Fujimura and Suzuki, 2010), as static facial stimuli do not ideally represent real-life processes (Foley *et al.*, 2012). Therefore, we used dynamic multisensory videos to maximally approximate real-life stimuli, elicit robust neural responses (Schröger and Widmann, 1998; Klasen *et al.*, 2012; Donohue *et al.*, 2013) and answer a complex research question.

On the basis of previous findings (Kanske and Kotz, 2012), we expected to observe enhanced RT conflict processing for emotional relative to neutral trials in both Experiment 1 and 2. Similarly, we expected to observe conflict facilitation in the N200 ERP response in emotional compared with neutral trials in both experiments (Kanske and Kotz, 2011b). Further, multisensory stimuli should modulate the N100 and P200 ERP responses when compared with unimodal stimuli. These effects are associated with early feature processing (Paulmann *et al.*, 2009; Jessen and Kotz, 2011; Kokinous *et al.*, 2014). Therefore, we expected to observe a conflict-driven enhancement of earlier ERP components (i.e. N100, P200) in comparison to previous studies that used unimodal stimuli (Kanske and Kotz, 2011b). The N100 is modulated by attention (Hillyard *et al.*, 1973), emotion (Scott *et al.*, 2009) and congruence (Atkinson *et al.*, 2003). Furthermore, the P200 has been shown to respond to task-relevance (Michalski, 2000) and emotion (Paulmann and Kotz, 2008). A common finding is that processing of incongruity leads to a decreased P200 amplitude (Kokinous *et al.*, 2014). Therefore, we expected to observe an enhanced N100 amplitude to incongruent relative to congruent trials and a reduced P200 amplitude to incongruent relative to congruent trials. Because previous studies showed emotional modulation of the conflict-specific N200 effect (Kanske and Kotz, 2011b), a somewhat open question was whether early conflict-specific N100 and P200 responses would also be modulated by the emotionality of the stimulus.

Experiment 1

Methods

Participants

Twenty-four healthy participants (female = 12, mean age = 24.5 years, s.d. = 4), all right-handed (Edinburgh Handedness Inventory score ME = 89.3, s.d. = 13.2), with normal or corrected-to-normal vision and normal hearing participated in the two experiments. All participants signed a consent form and received payment for their participation. The experiment was conducted in accordance with guidelines of the Declaration of Helsinki and approved by the Ethics Committee of the University of Leipzig.

Stimulus material

Stimuli consisted of short video clips of a male (27 years old) and a female (24 years old) semi-professional actor pronouncing the interjections 'A' and 'O' in a neutral and negative (i.e. angry) emotional tone of voice (see Figure 1A). The sound in all videos was normalized to 70 dB using root mean square in Final Cut Pro. We created 12 congruent and 12 incongruent videos by matching or mismatching vocalizations of the face and voice (e.g. voice pronouncing 'A' with facial lip movement corresponding to 'A' *vs* 'O', respectively). Videos were overlaid with the incongruent sound in Final Cut Pro 7 (Apple Inc) using the onset of the original video sound for alignment. The duration of congruent and incongruent videos ranged between 1 and 2 s (Table 1). There were no time differences between conditions before the audio onset and the total video durations (see supplementary material for details).

To check whether videos differed with regards to movement, we estimated the number of movements per video clip by quantifying the variation of light intensity (luminance) between pairs of frames for each pixel (Pichon *et al.*, 2008). The two emotions (neutral and emotional) and two vowels ('A' and 'O') were compared with a Kruskal-Wallis test. Emotional videos showed a higher number of movements relative to neutral videos ($X^2 = 5.33$, $P < 0.05$). As it is an inherent feature of angry expressions to be more dynamic and intense (e.g. Weyers *et al.*, 2006), the differences in mean motion were expected in naturalistic stimuli. Most importantly, difference in movement should not affect the outcome of the results, as we focused on the interaction of congruence and emotion. Motion difference was not significant for different vowels ($X^2 = 1.25$, $P > 0.2$).

Complete videos, video streams alone and audio streams alone were rated on a 7-point Likert scale using Self-Assessment Manikins in terms of expressiveness, arousal and emotion identification (Bradley and Lang, 1994) by 24 (12 female) participants (see Table 2). There was no difference in the level of perceived expressiveness and arousal. Participants rated emotional material as more emotional compared with neutral material (see supplementary material for details).

Experiment 1 had a 2 (emotional, neutral) × 2 (congruent, incongruent) factorial design and was split into four blocks. Each block comprised 52 videos (26 emotional and 26 neutral, half were congruent and the other half incongruent) presented in a pseudo-randomized order. Overall, there were 208 trials and testing took ~45 min per participant.

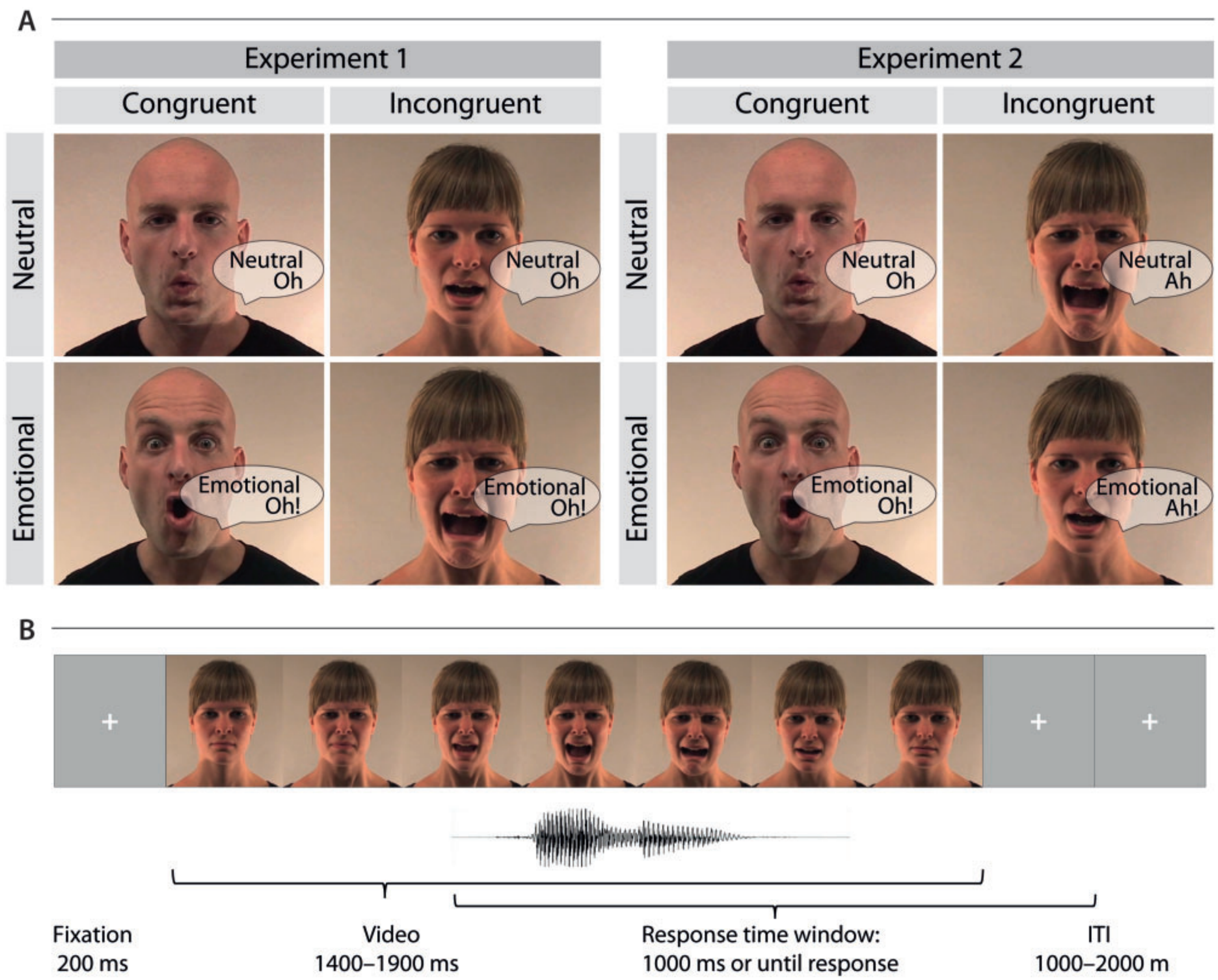

Fig. 1. A. Examples of video stimuli in Experiments 1 and 2: the female and the male actors vocalized the interjections 'A' and 'O' in either a negative or a neutral tone of voice. Incongruence was created by mismatches in the vocalizations of audio and video components in Experiment 1 and mismatches in emotion of audio and video components in Experiment 2. B. Example of a trial sequence.

Table 1. Timing of video stimuli of Experiment 1

Video condition (the 'vowel' specifies the interjection)	Time before start of the movement (ms)	Time before start of the audio sound (ms)	Total video duration (ms)
Female			
Neutral congruent 'A'	240	561	1400
Neutral congruent 'O'	240	740	1480
Negative congruent 'A'	240	665	1880
Negative congruent 'O'	240	846	1840
Face Neutral 'A'—voice neutral 'O'	240	540	1400
Face Neutral 'O'—voice neutral 'A'	240	562	1480
Face Negative 'A'—voice negative 'O'	240	680	1880
Face Negative 'O'—voice negative 'A'	240	630	1840
Male			
Neutral congruent 'A'	240	475	1400
Neutral congruent 'O'	240	560	1400
Negative congruent 'A'	240	450	1400
Negative congruent 'O'	240	540	1400
Face neutral 'A'—voice neutral 'O'	240	520	1400
Face neutral 'O'—voice neutral 'A'	240	635	1400
Face negative 'A'—voice negative 'O'	240	580	1400
Face negative 'O'—voice negative 'A'	240	490	1400

Table 2. Results of the video rating

Stimuli		Arousal	Expressiveness	Valence
Complete video	Neutral	4.97 (2.92)	4.80 (2.08)	5.38 (0.45)
	Negative	4.89 (2.51)	4.94 (2.90)	7.35 (1.19)
Audio stream	Neutral	4.23 (1.18)	4.29 (1.08)	4.96 (1.25)
	Negative	4.74 (0.95)	4.42 (1.05)	6.10 (0.79)
Videos stream	Neutral	4.37 (1.45)	4.63 (1.46)	4.91 (0.93)
	Negative	4.60 (1.05)	4.68 (0.88)	6.59 (0.66)

Note. Arousal and expressiveness were rated on a scale from 1 (low) to 9 (high). Valence of videos was rated on a scale from 1 (positive) to 9 (negative). Values represent mean and s.d.

Procedure

Participants were seated in a dimly lit sound-attenuated booth about 1 m from a computer screen. The sound was delivered via headphones. Each trial started with a fixation cross on a blank computer screen for 200 ms (see Figure 1B). Subsequently, a video was presented and played in full duration (i.e. video was not cut off after the response). Participants were instructed to report whether the vocalization of the voice was 'A' or 'O' irrespective of the emotional quality. The emotion of the face and voice was always matched. To ensure that faces were not ignored, participants were sometimes (10% of all trials) asked a second question about the vocalization (lip movement, i.e. 'A' or 'O') of the face. These additional questions were presented randomly and the response time to these questions was not limited. The response time window to the main question was 1000 ms starting from voice onset. If participants did not respond within 1000 ms, the words 'try to respond faster' appeared on the screen for 200 ms. If participants answered incorrectly, a word 'wrong' appeared on the screen. Button presses were counterbalanced across participants. An ITI of 1000, 1250, 1500, 1750 and 2000 ms was used randomly before the onset of the next trial.

EEG recording and preprocessing

EEG was recorded from 59 Ag/AgCl scalp electrodes (10-10 system). We used brain vision recorder (Brain Products GmbH, Munich, Germany) and a BRAINAMP amplifier (DC to 250 Hz). The sampling rate was 500 Hz. The left mastoid served as a reference and the sternum as ground. The vertical and horizontal electro-oculogram was measured for artifact rejection purposes. The impedance was kept below 5 kΩ.

EEG data were analyzed using the FieldTrip (v0.20120501) toolbox (Oostenveld *et al.*, 2011) running on Matlab 8.1 R2013a (The Mathworks, Natrick, USA). Electrodes were rereferenced offline to linked mastoids. Only correct trials were chosen for data processing. Extended epochs of 2000 ms before and after video onset were extracted. Epochs containing excessive muscle activity or jump artifacts were rejected. Data were band-pass filtered using a two-pass Butterworth IIR filter with a frequency passband of 0.1–100 Hz (order of four). Independent component analysis (ICA) was conducted using the runica algorithm. Subsequently, components were visually inspected and those showing ocular, muscle, heart and electrode artifacts were manually rejected (mean number of components removed = 17.41, s.d. = 6). In a subsequent step, individual epochs were visually scanned and epochs containing artifacts were manually discarded. Approximately 22% of trials (incorrect responses, artifacts) were excluded from further analysis.

Data analysis

For the statistical analysis, smaller epochs were selected (−200 to 1000 ms time-locked to the voice onset). Single-subject averages were calculated for each session and condition. Based on previous studies (Kanske and Kotz, 2010, 2011a,d), four regions of interests were defined: left anterior (FP1, AF3, AF7, F3, F5, F7, FC3, FC5, FT7), right anterior (P2, AF4, AF8, F4, F6, F8, FC4, FC6, FT8), left posterior: (CP3, CP5, TP7, P3, P5, P7, PO3, PO7, O1) and right posterior (CP4, CP6, TP8, P4, P6, P8, PO4, PO8, O2)[1]. Peak latencies were detected separately for each participant and

1 Conflict processing has repeatedly been observed for the N200 component over anterior midline electrodes. Please see the Supplementary material for an additional analysis of midline electrode sites for the N200 component in both experiments. These data were identical to the results of analyses of lateral electrode sites.

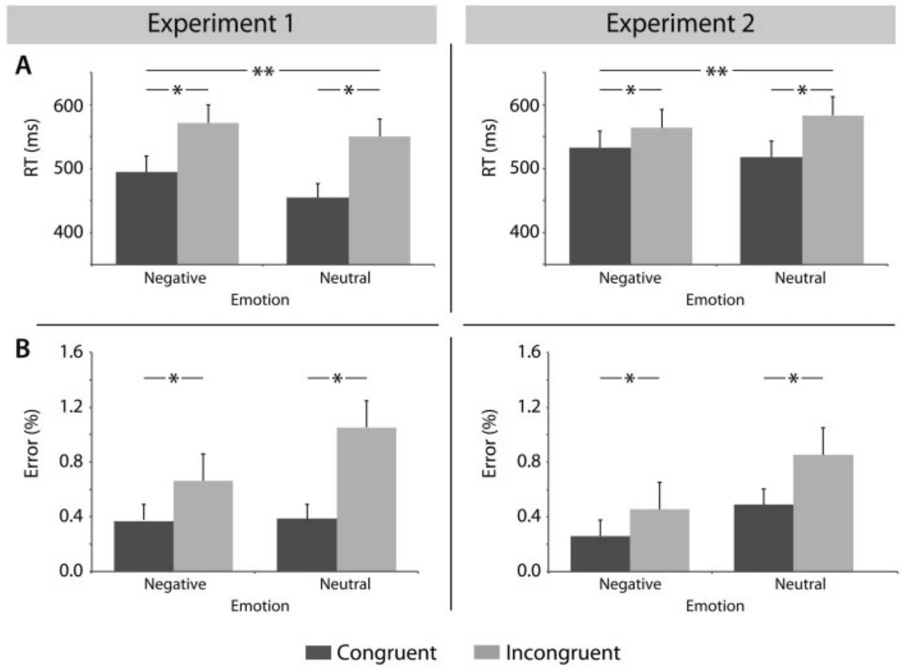

Fig. 2. Reaction time (A) and Error rate (B) data (mean + SEM) for congruent and incongruent/emotional and neutral conditions of Experiments 1 and 2. Incongruent videos elicited slower RTs and higher error rates (asterisk represents the main effect of congruence). Importantly, emotional stimuli resulted in smaller conflict effect (double asterisk represents a significant interaction of emotion by congruence).

each condition within the following time windows: 70–110 ms (N100), 180–225 ms (P200) and 225–285 ms (N200) as suggested by Luck and Kappenman (2011). A 20 ms time window was applied before and after each of the individual peaks for a mean amplitude analysis. For each time window, a repeated-measures ANOVA was calculated, using emotion (emotional, neutral), congruence (congruent, incongruent), region (anterior, posterior) and side (left, right) as within-subject factors. As videos prior to voice onset in Experiment 1 were identical (as emotion of face and voice was kept constant and only vocalization varied), no baseline correction was applied before the voice onset.

Results

Behavioral data

RT data

The main effects of emotion [$F(1, 23) = 28.154$, $P < 0.001$, $\eta_p^2 = 0.537$] and congruence were significant [$F(1, 23) = 88.093$, $P < 0.0001$, $\eta_p^2 = 0.769$], as was the interaction of congruence and emotion [$F(1, 23) = 9.193$, $P < 0.006$, $\eta_p^2 = 0.211$]. The conflict effect (incongruent–congruent) was smaller for emotional [77 ms; $F(1, 23) = 60.65$, $P < 0.001$, $\eta_p^2 = 0.671$] when compared with neutral [97 ms; $F(1, 23) = 98.35$, $P < 0.001$, $\eta_p^2 = 0.806$] videos (Figure 2).

Error rate

The main effect of congruence was significant [$F(1, 23) = 18.26$, $P < 0.001$, $\eta_p^2 = 0.477$]. Participants made more errors in response to incongruent (0.85%) than to congruent (0.38%) trials (Figure 2). No other effects reached significance (all $P > 0.1$).

ERP data

N100 range. We observed a significant main effect of congruence [$F(1, 23) = 9.571$, $P < 0.005$, $\eta_p^2 = 0.294$] and a significant interaction of emotion by region [$F(1, 23) = 30.045$, $P < 0.001$, $\eta_p^2 = 0.566$]. We resolved this interaction by region. Emotional trials elicited an enhanced amplitude compared with neutral trials in the anterior region [$F(1, 23) = 11.80$, $P < 0.01$, $\eta_p^2 = 0.339$], but not in the posterior region, where this effect was reversed [$F(1, 23) = 8.78$, $P < 0.01$ $\eta_p^2 = 0.276$].

Importantly, the interaction of emotion by congruence was also significant [$F(1, 23) = 4.767$, $P < 0.039$, $\eta_p^2 = 0.172$] (Figure 3). We resolved this interaction by emotion. For emotional stimuli, the N100 range was significantly larger for incongruent compared with congruent trials [$F(1, 23) = 15.743$, $P < 0.001$, $\eta_p^2 = 0.406$], whereas there was no such difference for neutral stimuli [$F(1, 23) = 0.523$, $P < 0.477$, $\eta_p^2 = 0.022$].

P200 range. The main effect of congruence was significant [$F(1, 23) = 8.739$, $P < 0.007$, $\eta_p^2 = 0.275$]. Overall, congruent stimuli elicited more positive amplitude relative to incongruent stimuli (Figure 3). The interaction of emotion by region was also significant [$F(1, 23) = 5.770$, $P < 0.025$, $\eta_p^2 = 0.201$]. We resolved this interaction by region. Emotional compared with neutral trials elicited an enhanced amplitude in the posterior region [$F(1, 23) = 4.266$, $P < 0.05$, $\eta_p^2 = 0.156$], but not in the anterior region [$F(1, 23) = 2.327$, $P > 0.141$, $\eta_p^2 = 0.092$].

N200 range. The main effect of congruence was significant: incongruent videos elicited enhanced amplitudes relative to congruent videos [$F(1, 23) = 12.393$, $P < 0.002$, $\eta_p^2 = 0.350$]. The congruence by region interaction was significant as well [$F(1, 23) = 10.788$, $P < 0.003$, $\eta_p^2 = 0.319$]. This interaction was resolved by region. The main effect of congruence was significant in the posterior brain region: incongruent videos elicited enhanced amplitudes relative to congruent videos [$F(1, 23) = 25.695$, $P < 0.001$, $\eta_p^2 = 0.528$]. The main effect of congruence was not significant over anterior brain regions [$F(1, 23) = 0.108$, $P > 0.745$, $\eta_p^2 = 0.005$].

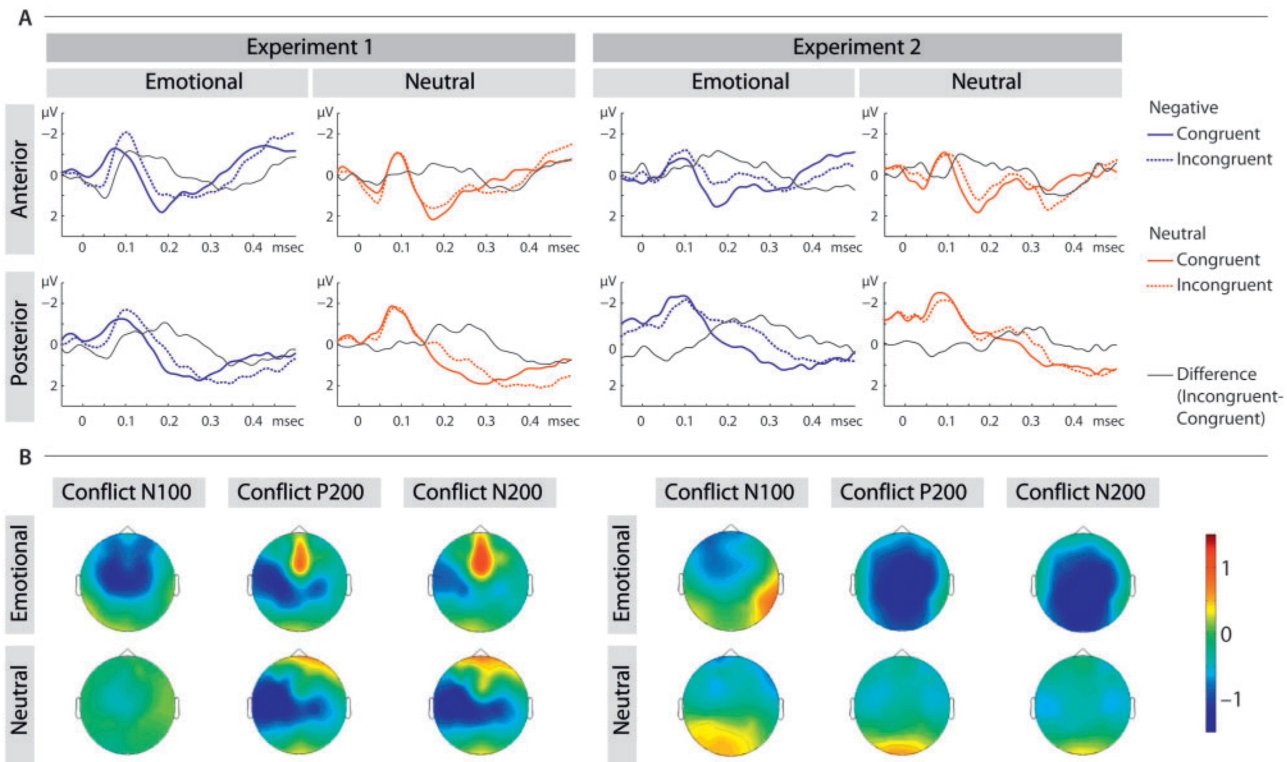

Fig. 3. (A) ERP waves at selected electrodes depicting the conflict effect for emotional and neutral videos of Experiments 1 and 2. (B) Conflict represents topographic distribution of amplitude difference (incongruent–congruent) for each of the ERP components (i.e. N100, P200 and N200 range).

In summary, Experiment 1 tested the influence of emotionality of the target stimulus on cognitive conflict processing using dynamic multisensory videos. We show that emotional stimuli facilitate RT conflict processing by reducing the conflict effect. Additionally, we found an interaction of emotion and congruence in the N100 time range, showing much earlier modulation of conflict processing by emotion compared with what was reported previously (i.e. N200, Kanske and Kotz, 2011b). To test the influence of emotionality of the target dimension on the emotional conflict processing, we conducted Experiment 2.

Experiment 2

Methods

Participants

Participants who took part in Experiment 1 also participated in Experiment 2.

Stimulus material and procedure

We modified original videos of Experiment 1 and created 12 congruent and 12 incongruent videos by matching or mismatching the emotional valence of the face and voice (e.g. voice pronouncing a neutral ‘A’ and the corresponding audio ‘A’ was pronounced emotionally, Figure 1). To create incongruent trials with a neutral target voice (i.e. audio trace), the original emotional facial video streams were combined with neutral auditory streams. Similarly, the originally neutral facial video streams were combined with emotional auditory streams to create incongruent trials with an emotional target. The original voice onset was used to align the incongruent voice with the lip movement in both incongruent conditions. The vocalization of the face and voice was always matched. Participants were instructed to report whether the voice was either emotional or neutral.

The duration of congruent and incongruent videos ranged between 1 and 2 s (Table 3). There were no differences between conditions in time before the audio onset and total video durations (see supplementary material for details).

Experiment 2 had a 2 (emotional, neutral) × 2 (congruent, incongruent) factorial design and consisted of four blocks. Each block comprised 52 videos (26 emotional and 26 neutral, half were congruent and the other half incongruent) presented in random order. To ensure that faces were not ignored, second questions were introduced (10% of all trials), for which participants were asked to report the emotion of the face. Overall, 208 trials were presented and testing took ∼45 min per participant.

EEG recording and preprocessing

The EEG recording and preprocessing were identical to Experiment 1. Artifact components were removed after ICA (mean number of components removed = 18.3, s.d. = 4.6). Approximately 24% of the trials (incorrect responses, artifacts) were excluded from further analysis.

Data analysis

Data analysis was identical to that of Experiment 1. In Experiment 2 a voice target (emotional or neutral) was preceded by either a congruent or an incongruent video. As we found motion differences for emotional compared with neutral videos, a 200 ms baseline correction was applied before the voice onset to account for these differences.

Table 3. Timing of video stimuli of Experiment 2

Video condition (the 'vowel' specifies the interjection)	Time before start of the movement (ms)	Time before start of the audio stream (ms)	Total video duration (ms)
Female			
Neutral congruent 'A'	240	561	1400
Neutral congruent 'O'	240	740	1480
Negative congruent 'A'	240	665	1880
Negative congruent 'O'	240	846	1840
Face neutral—voice negative 'A'	240	590	1400
Face neutral—voice negative 'O'	240	860	1480
Face negative—voice neutral 'A'	240	683	1880
Face negative—voice neutral 'O'	240	659	1840
Male			
Neutral congruent 'A'	240	475	1400
Neutral congruent 'O'	240	560	1400
Negative congruent 'A'	240	450	1400
Negative congruent 'O'	240	540	1400
Face neutral—voice negative 'A'	240	328	1400
Face neutral—voice negative 'O'	240	500	1400
Face negative—voice neutral 'A'	240	590	1400
Face negative—voice neutral 'O'	240	600	1400

Results

Behavioral data

RT data

We obtained a main effect of congruence [$F(1, 23) = 69.31$, $P < 0.0001$, $\eta_p^2 = 0.786$, see Figure 2] and an interaction of emotion by congruence [$F(1, 23) = 8.69$, $P < 0.008$, $\eta_p^2 = 0.316$]. The congruence effect (incongruent–congruent) was smaller in emotional target videos [32 ms; $F(1, 23) = 16.40$, $P < 0.001$, $\eta_p^2 = 0.467$] relative to neutral target videos [66 ms; $F(1, 23) = 60.17$, $P < 0.001$, $\eta_p^2 = 0.761$].

Error rate

There was a significant main effect of congruence: incongruent videos (0.66%) produced more errors compared with congruent videos (0.37%) [$F(1, 23) = 12.07$, $P < 0.01$, $\eta_p^2 = 0.365$], confirming the general effect of congruence. The main effect of emotion was significant [$F(1, 23) = 8.76$, $P < 0.007$, $\eta_p^2 = 0.294$, see Figure 2]. Emotional target stimuli elicited fewer errors (0.35%) relative to neutral target stimuli (0.65%).

ERP data

N100 range. We observed an interaction of congruence by side [$F(1, 23) = 5.520$, $P < 0.029$, $\eta_p^2 = 0.208$], congruence by region [$F(1, 23) = 10.145$, $P < 0.004$, $\eta_p^2 = 0.326$] and emotion by side [$F(1, 23) = 5.392$, $P < 0.03$, $\eta_p^2 = 0.204$]. Follow-up analyses did not confirm any significant main effects for any of the three interactions (all $P > 0.1$). Most importantly, a two-way interaction of congruence by emotion was significant [$F(1, 23) = 6.585$, $P < 0.018$, $\eta_p^2 = 0.239$, see Figure 3]. We resolved this interaction by emotion. The congruence effect was significant for neutral target videos, that is, the N100 range component showed a larger amplitude for congruent relative to incongruent trials [$F(1, 23) = 5.216$, $P < 0.033$, $\eta_p^2 = 0.199$]. The congruence effect for emotional target trials was not significant [$F(1, 23) = 2.291$, $P > 0.145$, $\eta_p^2 = 0.098$].

P200 range. We found a significant main effect of emotion [$F(1, 23) = 5.835$, $P < 0.025$, $\eta_p^2 = 0.217$] and a main effect of congruence [$F(1, 23) = 9.299$, $P < 0.006$, $\eta_p^2 = 0.307$], an interaction of congruence by emotion [$F(1, 23) = 5.945$, $P < 0.024$, $\eta_p^2 = 0.221$] and an interaction of congruence by region [$F(1, 23) = 8.492$, $P < 0.008$, $\eta_p^2 = 0.288$].

Finally, there was also a three-way interaction of emotion by congruence by region [$F(1, 23) = 4.026$, $P < 0.05$, $\eta_p^2 = 0.161$]. We resolved this three-way interaction by region. In the anterior brain region, we observed a main effect of congruence [$F(1, 23) = 22.467$, $P < 0.001$, $\eta_p^2 = 0.517$]. The P200 range amplitude response to congruent videos was enhanced relative to the one in incongruent videos in the anterior region.

Over posterior electrodes, there was a main effect of emotion [$F(1, 23) = 6.402$, $P < 0.019$, $\eta_p^2 = 0.234$]. Additionally, we observed an interaction of emotion by congruence [$F(1, 23) = 10.919$, $P < 0.003$, $\eta_p^2 = 0.342$]. This interaction was resolved by emotion. The congruence effect was significant for videos with emotional target voices [$F(1, 23) = 8.906$, $P < 0.007$, $\eta_p^2 = 0.298$]: incongruent videos produced reduced amplitudes relative to congruent videos. The congruence effect was not significant for videos with neutral target voices [$F(1, 23) = 1.534$, $P > 0.229$, $\eta_p^2 = 0.068$].

N200 range. As can be seen in the Figure 3, incongruent videos elicited enhanced amplitudes relative to congruent videos [$F(1, 23) = 19.110$, $P < 0.001$, $\eta_p^2 = 0.476$].

In summary, Experiment 2 set out to test the role of emotionality of the target dimension in emotional conflict processing. Consistent with our prediction, we show that an emotional compared with a neutral target facilitated emotional conflict processing by reducing the RT conflict effect. Additionally, we found an interaction of emotion and congruence in the N100 and P200 range components, marking early ERP correlates of emotional conflict processing.

Discussion

This study directly compared the effects of emotion on cognitive and emotional conflict processing by using behavioral measures and ERPs. Four main findings emerged: First, emotional target stimuli facilitated both cognitive and emotional conflict processing in RTs. The conflict effect was reduced for emotional compared with neutral trials in both experiments. Second, emotional but not neutral videos led to a conflict effect in the N100 range in Experiment 1. This result was reversed in Experiment 2; neutral but not emotional videos led to a conflict effect in the N100 range. Third, in Experiment 1 we found a main effect of congruence in the P200 range component. Incongruent videos elicited less positive (i.e. more negative) amplitudes relative to congruent videos, irrespective of emotional valence. This effect was also observed in the P200 range component of Experiment 2, except for posterior ERP responses to neutral videos. Finally, we observed a general conflict effect in the N200 range in both experiments.

The behavioral results of Experiment 1 replicated previous findings: cognitive conflict processing is facilitated if the target stimulus dimension is emotional rather than neutral (Kanske and Kotz, 2010, 2011a). Importantly, we showed that emotion facilitated emotional conflict processing when the target dimension (voice) of the stimulus was emotional. This is in line with the hypothesis that emotion can trigger control processes (Norman and Shallice, 1986). Norman and Shallice (1986) suggested that the selection of one of the conflicting action schemas (e.g. competing responses) is influenced by the 'activation value' of the stimulus. Emotional stimuli have a high activation value, due to their motivational relevance for survival (Ohman *et al.*, 2001; Vuilleumier *et al.*, 2001) and, thus, seem to facilitate conflict processing. Additionally, our findings extend previous reports on cognitive conflict to emotional conflict (Kanske and Kotz, 2012). In other words, the present findings show that the behavioral pattern that was previously observed for cognitive conflict also holds for emotional conflict.

Furthermore, we observed a conflict-related ERP component that has been repeatedly described earlier: the N200 (Bruin and Wijers, 2002). Previous studies found that the N200 amplitude is enhanced in incongruent relative to congruent trials (Kopp *et al.*, 1996; Van Veen and Carter, 2002). The N200 is also enhanced when an incorrect response is suppressed (Luck, 2005; Azizian *et al.*, 2006). Consistent with previous findings, the current N200 response was enhanced for incongruent relative to congruent trials. This result reflects a domain-general conflict processing mechanism that is common in both cognitive and emotional conflict processing (Egner *et al.*, 2008; Ochsner *et al.*, 2009; Torres-Quesada *et al.*, 2014).

The conflict-related ERPs have been described as arising from generators in the ACC (Liotti *et al.*, 2000; Hanslmayr *et al.*, 2008). However, the distribution of the ERP conflict effect varies considerably across studies and experimental manipulations. There are studies that report conflict related ERP activations that are anteriorly distributed (e.g. Markela-Lerenc *et al.*, 2004; Hanslmayr *et al.*, 2008), centrally distributed (e.g. Donohue *et al.*, 2013) and posteriorly distributed (Liotti *et al.*, 2000; Appelbaum *et al.*, 2009). These observed distribution differences of conflict related ERP activations may reflect differences in the anterior and posterior attentional control systems that have been described earlier (Eimer *et al.*, 2003; Petersen and Posner, 2012). Therefore, further controlled studies are required to clarify the distribution of conflict-specific ERP components over the anterior, central and posterior brain regions.

We also found earlier ERP responses (i.e. N100, P200 range) that were affected by congruence and emotion. These results imply that conflict information and information regarding the emotional quality of a stimulus are already integrated 100 ms after stimulus onset. Because multisensory and dynamic stimuli receive prioritized perceptual processing relative to unimodal stimuli (Franconeri and Simons, 2003; Brockelmann *et al.*, 2011), we conclude that earlier modulation of conflict-sensitive ERPs may be due to the multisensory nature of stimuli (Ho *et al.*, 2014; Kokinous *et al.*, 2014).

In contrast to the comparable behavioral results, cognitive and emotional conflict processing showed reversed conflict-related N100 effects. We interpret the observed N100 reversal as reflecting conflict-specific mechanisms that are used to process the task (Egner and Hirsch, 2005; Etkin *et al.*, 2006). Specifically, cognitive conflict processing amplifies the processing of the target dimension of a stimulus (Egner and Hirsch, 2005). Therefore, emotional stimuli in Experiment 1 triggered executive attention and facilitated conflict processing (Kanske and Kotz, 2010) by increasing the attentional focus to the target (Derryberry and Reed, 1994; Fredrickson, 2001). This corresponds to the enhanced incongruent N100 range response (Hillyard *et al.*, 1973). Previous studies also showed enhanced early negative ERP responses to incongruent, compared with congruent stimuli in cognitive conflict tasks, even though these effects occurred in the N200 time window for unisensory stimuli (Kanske and Kotz, 2010, 2011b).

Emotional conflict processing (Experiment 2), on the other hand, has been found to involve facilitated inhibition of the non-target affective distractor dimension of a stimulus (Etkin *et al.*, 2006). As the inhibition of neutral relative to emotional non-target faces is less demanding (Ohman *et al.*, 2001; Vuilleumier *et al.*, 2001), we observed a reduced N100 conflict effect for emotional relative to neutral videos in the emotional conflict task. Therefore, the reduced N100 conflict effect is likely the result of lower interference from neutral distractors. A reduced N100 response to incongruent compared with congruent emotional audiovisual face–voice combinations has been reported previously (Pourtois *et al.*, 2000; Ho *et al.*, 2014). Usually, the N100 modulation in multisensory perception is explained as a cross-sensory anticipatory process (Stekelenburg and Vroomen, 2007; Vroomen and Stekelenburg, 2010; Jessen and Kotz, 2011). Visual information precedes acoustic information (P.M.T. Smeele, Unpublished doctoral dissertation), which is also the case in the naturalistic stimuli used here (see Tables 1 and 3). Therefore, visual information can cue the upcoming auditory signal and help to form predictions, which, if correct, lead to faster and more efficient processing of the auditory modality (Klucharev *et al.*, 2003; Besle *et al.*, 2004; van Wassenhove *et al.*, 2005). However, when an audiovisual combination is incompatible, an incongruent visual cue leads to processing costs, that is, conflict (Jakobs *et al.*, 2009). This prediction violation drives the observed N100 suppression in response to incongruent videos (Pourtois *et al.*, 2000; Ho *et al.*, 2014). However, contrary to previous studies, our findings show that this suppression only happens when the distractor modality is emotional rather than neutral. Given the present findings, we propose that the emotional-conflict N100 suppression observed previously and in this study is primarily driven by the to-be-suppressed emotional distractor of the audiovisual input (Etkin *et al.*, 2006).

Our findings also showed a conflict effect in the P200 range, which was modulated by the emotional target dimension in emotional but not in cognitive conflict processing. These results go hand in hand with a recent study on multisensory emotional audio visual integration that reported an incongruent P200 response to be less positive than a congruent P200 response (Kokinous *et al.*, 2014), possibly reflecting a reduced information gain from incongruent stimuli. Furthermore, Knowland *et al.* (2014) have suggested that the audiovisual P200 suppression reflects competition between different multisensory inputs with greater competition for incompatible stimulation. Importantly, in both cases (i.e. reduced information gain, increased competition) we would expect to observe a larger P200 reduction for emotional distractors relative to neutral distractors in emotional conflict processing. However, we observed the opposite. Therefore, reduced information gain and greater competition from a distractor cannot account for the observed effects. Finally, a decreased P200 amplitude has also been linked to increased allocation of attention (Crowley and Colrain, 2004). Therefore, a reduced P200 response in incongruent trials could imply attentional capture by mismatching vocalizations. As an emotional face requires more resources to be inhibited than a neutral face (Algom *et al.*, 2004; Dolcos and McCarthy, 2006), we observed less attentional allocation to the vocalization when the distractor was emotional rather than neutral.

Note that the observed ERP responses could also be a result of a sustained negativity caused by having to process incongruent stimuli (Giard *et al.*, 1987). However, we reason that this possibility is less likely since we did not observe a sustained negativity over all the ERP components and some components produced the opposite effect (i.e. more positive amplitudes for incongruent stimuli). Nevertheless, a sustained negativity cannot be completely ruled out.

Conclusion

Successful communication often requires processing of emotional conflicts, produced by different sensory modalities (e.g. audio, visual). For instance, in irony or satire emotions depicted in the face and the voice of the interlocutor do not match. In two ERP experiments, we have shown that emotional target stimuli facilitate both cognitive and emotional conflict processing. Although the behavioral outcomes of the two types of conflict were comparable, we identified a domain-general (N200 range) and two domain-specific (N100, P200 range) ERP correlates of emotional and cognitive conflict processing. We suggest that the emotional target dimension has a facilitatory effect on conflict processing, irrespective of conflict type, although neuropsychological mechanisms underlying these processes may differ. Lastly, as this study concentrated on negative emotions, one question that remains is whether we would observe the same result with positive emotional conflict (e.g. schadenfreude, courtesy laugh) that also represents an important social function (e.g. trying not to laugh in a serious situation so as to not offend others).

Supplementary data

Supplementary data are available at *SCAN* online.

Conflict of interest. None declared.

References

Algom, D., Chajut, E., Lev, S. (2004). A rational look at the emotional Stroop phenomenon: a generic slowdown, not a Stroop effect. *Journal of Experimental Psychology: General*, **133**, 323–38.

Appelbaum, L.G., Meyerhoff, K.L., Woldorff, M.G. (2009). Priming and backward influences in the human brain: processing interactions during the Stroop interference effect. *Cerebral Cortex*, **19**, 2508–21.

Atkinson, C.M., Drysdale, K.A., Fulham, W.R. (2003). Event-related potentials to Stroop and reverse Stroop stimuli. *International Journal of Psychophysiology*, **47**(1), 1–21.

Azizian, A., Freitas, A.L., Parvaz, M.A., Squires, N.K. (2006). Beware misleading cues: perceptual similarity modulates the N2/P3 complex. *Psychophysiology*, **43**(3), 253–60.

Besle, J., Fort, A., Delpuech, C., Giard, M.H. (2004). Bimodal speech: early suppressive visual effects in human auditory cortex. *The European Journal of Neuroscience*, **20**(8), 2225–34.

Bradley, M.M., Lang, P.J. (1994). Measuring emotion: the Self-Assessment Manikin and the Semantic Differential. *J Behav Ther Exp Psychiatry*, **25**:49–59.

Brockelmann, A.K., Steinberg, C., Elling, L., Zwanzger, P., Pantev, C., Junghofer, M. (2011). Emotion-associated tones attract enhanced attention at early auditory processing: magnetoencephalographic correlates. *The Journal of Neuroscience*, **31**(21), 7801–10.

Bruin, K.J., Wijers, A.A. (2002). Inhibition, response mode, and stimulus probability: a comparative event-related potential study. *Clinical Neurophysiology*, **113**(7), 1172–82.

Crowley, K.E., Colrain, I.M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clinical Neurophysiology*, **115**(4), 732–44.

Derryberry, D., Reed, M.A. (1994). Temperament and attention: orienting toward and away from positive and negative signals. *Journal of Personality and Social Psychology*, **66**(6), 1128–39.

Dolcos, F., McCarthy, G. (2006). Brain systems mediating cognitive interference by emotional distraction. *J. Neurosci*, **26**, 2072–9.

Donohue, S.E., Todisco, A.E., Woldorff, M.G. (2013). The rapid distraction of attentional resources. *Journal of Cognitive Neuroscience*, **12**, 1–15.

Egner, T., Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, **8**(12), 1784–90.

Egner, T., Etkin, A., Gale, S., Hirsch, J. (2008). Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. *Cerebral Cortex*, **18**(6), 1475–84.

Eimer, M., Forster, B., van Velzen, J. (2003). Anterior and posterior attentional control systems use different spatial reference frames: ERP evidence from covert tactile-spatial orienting. *Psychophysiology*, **40**, 924–33.

Etkin, A., Egner, T., Peraza, D.M., Kandel, E.R., Hirsch, J. (2006). Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron*, **51**(6), 871–82.

Foley, E., Rippon, G., Thai, N.J., Longe, O., Senior, C. (2012). Dynamic facial expressions evoke distinct activation in the face perception network: a connectivity analysis study. *J Cogn Neurosci*, **24**:507–20.

Franconeri, S.L., Simons, D.J. (2003). Moving and looming stimuli capture attention. *Perception and Psychophysics*, **65**(7), 999–1010.

Fredrickson, B.L. (2001). The role of positive emotions in positive psychology. The broaden-and-build theory of positive emotions. *The American Psychologist*, **56**(3), 218–26.

Fujimura, T. & Suzuki, N. (2010). Recognition of dynamic facial expressions in peripheral and central vision. *The Japanese Journal of Psychology*, **81**(4), 348–55.

Giard, M.H., Perrin, F., Pernier, J., Peronnet, F. (1988). Several attention-related wave forms in auditory areas: a topographic study. *Electroencephalography and Clinical Neurophysiology*, **69**, 371–84.

Hart, S.J., Green, S.R., Casp, M., Belger, A. (2010). Emotional priming effects during Stroop task performance. *Neuroimage,* **49**(3), 2662–70.

Hillyard, S.A., Hink, R.F., Schwent, V.L., Picton, T.W. (1973). Electrical signs of selective attention in human brain. *Science,* **182**(4108), 177–80.

Hanslmayr, S., Pastotter, B., Bauml, K.H., Gruber, S., Wimber, M., Klimesch, W. (2008). The electrophysiological dynamics of interference during the Stroop task. *Journal of Cognitive Neuroscience*, **20**, 215–25.

Ho, H.T., Schröger, E., Kotz, S.A. (2014). Selective attention modulates early human evoked potentials during emotional face1-voice perception. *Journal of Cognitive Neuroscience*, **30**, 1–21.

Jakobs, O., Wang, L.E., Dafotakis, M., Grefkes, C., Zilles, K., Eickhoff, S.B. (2009). Effects of timing and movement uncertainty implicate the temporo-parietal junction in the prediction of forthcoming motor actions. *Neuroimage*, **47**, 667–77.

Jessen, S., Kotz, S.A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *Neuroimage*, **58**, 665–74.

Kanske, P. (2012). On the influence of emotion on conflict processing. *Frontiers in Integrative Neuroscience,* **6**, 42.

Kanske, P., Kotz, S.A. (2010). Modulation of early conflict processing: N200 responses to emotional words in a flanker task. *Neuropsychologia,* **48**(12), 3661–4.

Kanske, P., Kotz, S.A. (2011a). Emotion speeds up conflict resolution: a new role for the ventral anterior cingulate cortex? *Cerebral Cortex,* **21**(4), 911–9.

Kanske, P., Kotz, S.A. (2011b). Emotion triggers executive attention: anterior cingulate cortex and amygdala responses to emotional words in a conflict task. *Human Brain Mapping,* **32**(2), 198–208.

Kanske, P., Kotz, S.A. (2011c). Conflict processing is modulated by positive emotion: ERP data from a flanker task. *Behavioural Brain Research*, **219**, 382–6.

Kanske, P., Kotz, S.A. (2011d). Positive emotion speeds up conflict processing: ERP responses in an auditory Simon task. *Biological Psychology*, **87**, 122–7.

Klasen, M., Chen, Y.H., Mathiak, K. (2012). Multisensory emotions: perception, combination and underlying neural processes. *Reviews in the Neurosciences*, **23**, 381–92.

Klucharev, V., Möttönen, R., Sams, M. (2003). Electrophysiological indicators of phonetic and non1phonetic multisensory interactions during audiovisual speech perception. *Brain Research. Cognitive Brain Research,* **18**(1), 65–75.

Knowland, V.C.P., Mercure, E., Karmiloff-Smith, A., Dick, F., Thomas, M.S.C. (2014). Audio1visual speech perception: a developmental ERP investigation. *Developmental Science*, **17**(1), 110–24.

Kokinous, J., Kotz, S.A., Tavano, A., Schroger, E. (2014). The role of emotion in dynamic audiovisual integration of faces and voices. *Social Cognitive and Affective Neuroscience*, **10**(5), 713–20.

Kopp, B., Mattler, U., Goertz, R., Rist, F. (1996). N2, P3 and the lateralized readiness potential in a nogo task involving selective response priming. *Electroencephalography and Clinical Neurophysiology,* **99**(1), 19–27.

Liotti, M., Woldorff, M.G., Perez, R., Mayberg, H.S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia*, **38**, 701–11.

Luck, S. (2005). *An Introduction to the Event-Related Potential Technique*, **Vol. 1**, MA: The MIT Press, 1–49.

Luck, S.J. Kappenman, E.S. (2011). *The Oxford Handbook of Event-Related Potential Components*. New York: Oxford University Press.

Markela-Lerenc, J., Ille, N., Kaiser, S., Fiedler, P., Mundt, C., Weisbrod, M. (2004). Prefrontal-cingulate activation during executive control: which comes first? *Brain Research on Cognitive Brain Research*, **18**, 278–87.

Michalski, A. (2000). Expectation of an important event affects responses to irrelevant stimuli of different modalities *Acta Neurobiologiae Experimentalis*, **60**, 467–78.

Norman, D.A., Shallice, T. (1986). Attention to action: willed and automatic control of behavior. In: Davidson, R.J., Schwartz, G.E., Shapiro D., editors. *Consciousness and Self-Regulation*. New York: Plenum, 1–18.

Ochsner, K.N., Hughes, B., Robertson, E.R., Cooper, J.C., Gabrieli, J.D. (2009). Neural systems supporting the control of affective and cognitive conflicts. *Journal of Cognitive Neuroscience,* **21**(9), 1842–55.

Ohman, A., Flykt, A., Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology: General,* **130**(3), 466–78.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J. M. (2001). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience,* **2011**(156869), 1–9.

Padmala, S., Bauer, A., Pessoa, L. (2011). Negative emotion impairs conflict-driven executive control. *Frontiers in Psychology,* **2**, 192.

Paulmann, S., Kotz, S.A. (2008). An ERP investigation on the temporal dynamics of emotional prosody and emotional semantics in pseudo- and lexical-sentence context. *Brain and Language*, **105**(1), 59–69.

Paulmann, S., Jessen, S., Kotz, S.A. (2009). Investigating the multimodal nature of human communication. *Journal of Psychophysiology*, **23**(2), 63–76.

Pessoa, L., Padmala, S., Kenzer, A. Bauer, A. (2012). Interactions between cognition and emotion during response inhibition *Emotion*, **12**(1), 192–7.

Petersen, S.E. Posner, M.I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, **35**, 73–89.

Pichon, S., de Gelder, B., Grezes, J. (2008). Emotional modulation of visual and motor areas by dynamic body expressions of anger. *Social Neuroscience,* **3**(3–4), 199–212.

Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport,* **11**, 1329–33.

Scott, G.G., O'Donnell, P.J., Leuthold, H., Sereno, S.C. (2009). Early emotion word processing: evidence from event-related potentials. *Biological Psychology,* **80**(1), 95–104.

Schröger, E., Widmann, A. (1998). Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology*, **35**(6), 755–9.

Simon, J.R., Rudel, A.P. (1967). Auditory S-R compatibility: the effect of an irrelevant cue on information processing. *J Appl Psychol*, **51**:300–4.

Soutschek, A., Schubert, T. (2013). Domain-specific control mechanisms for emotional and nonemotional conflict processing. *Cognition,* **126**(2), 234–45.

Stekelenburg, J.J., Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *Journal of Cognitive Neuroscience*, **19**(12), 1964–73.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**(6), 643–62.

Torres-Quesada, M., Korb, F.M., Funes, M.J., Lupianez, J., Egner, T. (2014). Comparing neural substrates of emotional vs.

non-emotional conflict modulation by global control context. *Frontiers in Human Neuroscience,* **8**, 66.
Van Veen, V., Carter, C.S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience,* **14**(4), 593–602.
Van Wassenhove, V., Grant, K.W., Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(4), 1181–6.
Vroomen, J., Stekelenburg, J.J. (2010). Visual anticipatory information modulates multisensory interactions of artificial audiovisual stimuli. *Journal of Cognitive Neuroscience*, **22**(7), 1583–96.
Vuilleumier, P., Armony, J.L., Driver, J., Dolan, R.J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron,* **30**(3), 829–41.
West, R. (2003). Neural correlates of cognitive control and conflict detection in the Stroop and digit-location tasks. *Neuropsychologia,* **41**, 1122–35. doi: 10.1016/s0028-3932(02)00297-x.
Weyers, P., Muhlberger, A., Hefele, C., Pauli, P. (2006). Electromyographic responses to static and dynamic avatar emotional facial expressions. *Psychophysiology,* **43**(5), 450–3.