Preserved sensory processing but hampered conflict detection when stimulus input is task-irrelevant

Stijn A. **Nuiten**[1,2*], Andrés **Canales-Johnson**[1,2,3.4], Lola **Beerendonk**[1,2,3], Nutsa **Nanuashvili**[1,2], Johannes J. **Fahrenfort**[1,2], Tristan **Bekinschtein**[3,4,^] and Simon **van Gaal**[1,2,*,^]

[1] Department of Psychology, University of Amsterdam, Nieuwe Achtergracht 129-B, 1018 WT, Amsterdam, The Netherlands

[2] Amsterdam Brain & Cognition, University of Amsterdam, Nieuwe Achtergracht 129-B, 1018 WT Amsterdam, The Netherlands

[3] Department of Psychology, University of Cambridge, Downing Street, CB2 3EB Cambridge, United Kingdom

[4] Behavioural and Clinical Neuroscience Institute, Downing Street, CB2 3EB Cambridge, United Kingdom

[*] Corresponding authors

[^] Shared senior authorship

Abstract (now 150, max 150)

Conflict detection in sensory input is central to adaptive human behavior. Perhaps unsurprisingly, past research has shown that conflict may even be detected in absence of conflict awareness, suggesting that conflict detection is an automatic process that does not require attention. To test the possibility of conflict processing in the absence of attention, we manipulated task relevance and response overlap of potentially conflicting stimulus features across six behavioral tasks. Multivariate analyses on human electroencephalographic data revealed neural signatures of conflict only when at least one feature of a conflicting stimulus was attended, regardless of whether that feature was part of the conflict, or overlaps with the response. In contrast, neural signatures of basic sensory processes were present even when a stimulus was completely unattended. These data reveal an attentional bottleneck at the level of objects, suggesting that object-based attention is a prerequisite for cognitive control operations involved in conflict detection.

Introduction

Every day we are bombarded with sensory information from the environment and we often face the challenge of selecting the relevant information and ignoring irrelevant –potentially conflicting– information to maximize performance. These selection processes require much effort and our full attention, sometimes rendering us deceptively oblivious to irrelevant sensory input (e.g. chest-banging apes), as illustrated by the famous inattentional blindness phenomenon (Simons & Chabris, 1999). However, unattended events that are not relevant for the current task might still capture our attention or interfere with ongoing task performance, for example when they are inherently relevant to us (e.g. our own name). This is illustrated by another famous psychological phenomenon: the cocktail party effect (Cherry, 1953; Moray, 1959). Thus, under specific circumstances, task-irrelevant information may capture attentional resources and be subsequently processed with different degrees of depth.

It is currently a matter of debate which processes require top-down attention (Dehaene et al., 2006; Koch & Tsuchiya, 2007; Koelewijn et al., 2010; Lamme, 2003; Lamme & Roelfsema, 2000; Rousselet et al., 2004; VanRullen, 2007). It was long thought that only basic physical stimulus features or very salient stimuli are processed in the absence of attention (Treisman & Gelade, 1980), due to an "attentional bottleneck" at higher levels of analysis (Broadbent, 1958; Deutsch & Deutsch, 1963; Lachter et al., 2004; Wolfe & Horowitz, 2004). However, there is now solid evidence that several tasks may in fact still unfold in the (near) absence of attention, including perceptual integration (Fahrenfort et al., 2017), the processing of emotional valence (Sand & Wiens, 2011; Stefanics et al., 2012), semantical processing of written words (Schnuerch et al., 2016) and visual scene categorization (Li et al., 2002; Peelen et al., 2009). Although one should be cautious in claiming complete absence of attention (Lachter et al., 2004), these and other studies have pushed the boundaries of input processing that is task-irrelevant (without attention), and may even question the existence of an attentional bottleneck at all, at least for relatively low-level information. Conceivably, the attentional bottleneck is only present at higher, more complex, levels of cognitive processing, like cognitive control functions.

Over the years, various theories have been proposed with regard to this attentional bottleneck among which the load theory of selective attention and cognitive control (Lavie et al., 2004), the multiple resources theory (Wickens, 2002), the hierarchical central executive bottleneck theory and formalizations thereof in a cortical network model for serial and parallel processing (Sigman & Dehaene, 2006; Zylberberg et al., 2010, 2011). These theories all hinge on the idea that resources for the processing of information are limited and that the brain therefore has to allocate resources to processes that are currently most relevant, via selective attention (Broadbent, 1958; Treisman, 1969). Resource (re-)allocation, and thus flexible behavior, is thought to be governed by an executive network, most prominently involving the prefrontal cortex (Goldman-Rakic, 1995, 1996). Information that is deemed task-irrelevant has fewer resources at its disposal and is therefore processed to a lesser extent. When more resources are necessary for processing the task-relevant information, e.g. under high perceptual load, processing of task-irrelevant information diminishes (Lavie et al., 2003, 2004). Yet even under high perceptual load, task-irrelevant features can be processed when they are part of an attended object (when object-based attention is present) (Chen, 2012; Chen & Cave, 2006; Cosman & Vecera, 2012; Kahneman et al., 1992; O'Craven et al., 1999; Schoenfeld et al., 2014; Wegener et al., 2014). There is currently no consensus which type of information can be processed in parallel by the brain and which attentional mechanisms determine what information passes the attentional bottleneck. One unresolved issue is that most empirical work has investigated the bottleneck with regard to sensory features, however it is unknown if the bottleneck and the distribution of processing resources also takes place for more complex, cognitive processes. Here, we test whether such a high-level attentional bottleneck indeed exists in the human brain.

Specifically, we aim to test whether cognitive control operations, necessary to identify and resolve conflicting sensory input, are operational when that input is irrelevant for the task at hand (and hence unattended) and what role object-based attention may have in conflict detection. Previous work has shown that the brain has

dedicated networks for the detection and resolution of conflict, in which the medial frontal cortex (MFC) plays a pivotal role (Ridderinkhof et al., 2004). Conflict detection and subsequent behavioral adaptation is central to human cognitive control and, hence, it may not be surprising that past research has shown that conflict detection can even occur unconsciously (Atas et al., 2016; D'Ostilio & Garraux, 2012a; Huber-Huber & Ansorge, 2018; van Gaal et al., 2008), suggesting that the brain may detect conflict fully automatically and that it may even occur without paying attention (e.g. Rahnev et al., 2012). Moreover, it has been shown that this automaticity can be enhanced by training, resulting in more efficient processing of conflict (Chen et al., 2013; MacLeod & Dunbar, 1988; van Gaal et al., 2008).

Conclusive evidence regarding the claim that conflict detection is fully automatic has, to our knowledge, not been provided and therefore the necessity of attention for cognitive control operations remains open for debate. Previous studies have shown that cognitive control processes are operational when to-be-ignored features from either a task-relevant or a task-irrelevant stimulus overlap with the behavioral response to be made to the primary task, causing interference in performance (Mao & Wang, 2008; Padrão et al., 2015; Zimmer et al., 2010). In these circumstances the interfering stimulus feature carries information related to the primary task and is therefore *de facto* not task-irrelevant. Consequently, it is currently unknown whether cognitive control operations are active for conflicting sensory input that is not related to the task at hand. Given the immense stream of sensory input we encounter in our daily lives, conflict between two (unattended) sources of perceptual information is inevitable.

Here, we investigated whether conflict between two features of an auditory stimulus (its content and its spatial location) would be detected by the brain under varying levels of task relevance of these features. The main aspect of the task was as follows. We presented auditory spoken words ("left" and "right" in Dutch) through speakers located on the left and right side of the body. By presenting these stimuli through either the left or the right speaker, content-location conflict arises on specific trials (e.g. the word "left" from the right speaker) but not on others (e.g. the word "right" from the right speaker) (Buzzell et al., 2013; Canales-Johnson et al., 2020). A wealth of previous studies has revealed that conflict arises between task-relevant and task-irrelevant features of the stimulus in these type of tasks (similar to the Simon task and Stroop task, Egner & Hirsch, 2005; Hommel, 2011). Here, these potentially conflicting auditory stimuli were presented during six different behavioral tasks, divided over two separate experiments, multiple experimental sessions and different participant groups (both experiments N=24). In all tasks we focus on the processing of content-location conflict of the auditory stimulus. There were several critical differences between the behavioral tasks: (1) task relevance of a conflicting feature of the stimulus, (2) task relevance of a non-conflicting feature that was part of a conflicting stimulus and (3) whether the response to be given mapped onto a conflicting feature of the stimulus. Note that in all tasks, only one feature could be task-relevant, and that all the other feature(s) had to be ignored. The systematic manipulation of task relevance and the response mapping allowed us to explore the full landscape of possibilities of how varying levels of attention affect sensory and conflict processing. Electroencephalography (EEG) was recorded and multivariate analyses on the EEG data were used to extract any neural signatures of conflict detection (i.e. theta-band neural oscillations; Cavanagh & Frank, 2014; Cohen & Cavanagh, 2011) and sensory processing for any of its features. Furthermore, in both experiments we measured behavioral and neural effects of task-irrelevant conflict before and after training on conflict-inducing tasks, aiming to investigate the role of automaticity in the detection of (task-irrelevant) conflict.

Results

Experiment 1: Can the brain detect fully task-irrelevant conflict?

In the first experiment, 24 human participants performed two behavioral tasks (**Figure 1A**). In the auditory conflict task (from hereon: content discrimination task I) the feature "sound content" was task-relevant. Participants were instructed to respond according to the content of the auditory stimulus ("left" vs "right"),

ignoring its spatial location that could conflict with the content response (presented from the left or right side of the participant). For the other behavioral task, participants performed a demanding visual random dot-motion (RDM) task in which they had to discriminate the direction of vertical motion (from hereon: vertical RDM task), while being presented with the same auditory stimuli—all features of which were thus fully irrelevant for task performance. Behavioral responses on this visual task were orthogonal to the response tendencies potentially triggered by the auditory features, excluding any task or response related interference (**Figure 1B**). Under this manipulation, all auditory features are task-irrelevant and are orthogonal to the response-mapping. To maximize the possibility of observing conflict detection when conflicting features are task-irrelevant, and to explore the effect of task automatization on conflict processing, participants performed the tasks both before and after extensive training, which may increase the efficiency of cognitive control (**Figure 1C**) (van Gaal et al., 2008).

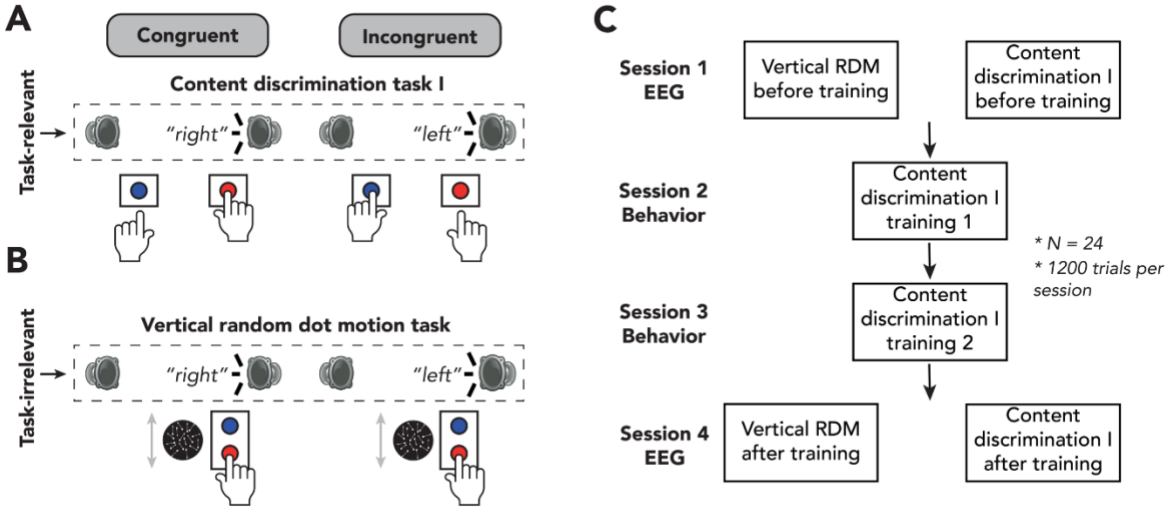

Figure 1. Experimental design of experiment 1. (A/B) Schematic representation of the experimental design for auditory content discrimination task I **(A)** and vertical RDM task **(B)**. In both tasks, the spoken words "left" and "right were presented through either a speaker located on the left or right side of the participant. Note that auditory stimuli are only task-relevant in auditory content discrimination task I and not in the vertical RDM task. In this figure, sounds are only depicted as originating from the right, whereas in the experiment the sounds could also originate from the left speaker. **(A)** In content discrimination task I, participants were instructed to report the content ("left" or "right") of an auditory stimulus via a button press with their left or right hand respectively, and to ignore the spatial location at which the auditory stimulus was presented. **(B)** During the vertical RDM task, participants were instructed to report the overall movement direction of dots (up or down) via a button press with their right hand, whilst still being presented with the auditory stimuli, which were therefore task-irrelevant. In both tasks, content of the auditory stimuli could be congruent or incongruent with its location of presentation (50% congruent/incongruent trials). **(C)** Overview of the sequence of the four experimental sessions of this study. Participants performed two EEG sessions during which they first performed the vertical RDM task followed by auditory content discrimination task I. Each session consisted of 1200 trials, divided over twelve blocks, allowing participants to rest in between blocks. In between experimental sessions, participants were trained on auditory content discrimination task I on two training sessions of 1 hour each.

Experiment 1: Conflicting information induces slower responses and decreased accuracy only for task-relevant sensory input

For content discrimination task I, mean error rates (ERs) were 2.6% (*SD* = 2.7%) and mean reaction times (RTs) 477.2ms (*SD* = 76.1ms), averaged over all four sessions. For the vertical RDM mean ERs were 19.2% (*SD* = 6.6%) and mean RTs were 711.4ms (*SD* = 151.3ms). The mean ER of vertical RDM indicates that our staircasing-procedure was effective. To investigate whether our experimental design was apt to induce conflict effects for task-relevant sensory input and to test whether conflict effects were still present when sensory input was task-irrelevant, we performed repeated measures (rm-)ANOVAs (2x2x2 factorial) on mean RTs and ERs gathered

during the EEG-recording sessions (session 1, "before training"; session 4, "after training"). This allowed us to include 1) task relevance (yes/no), 2) training (before/after) and 3) congruency of auditory content with location of auditory source (congruent/incongruent). Note that congruency is always defined based on the relationship between two features of the auditorily presented stimuli, also when participants performed the visual task (and therefore the auditory features were task-irrelevant).

Detection of conflict is typically associated with behavioral slowing and increased error rates. Indeed, we observed that, across both tasks, participants were slower and made more errors on incongruent trials as compared to congruent trials (the conflict effect, RT: $F(1,23) = 52.83$, $p < .001$, $\eta_p^2 = 0.70$; ER: $F(1,23) = 9.13$, $p = .01$, $\eta_p^2 = 0.28$). This conflict effect was modulated by task relevance of the auditory features (RT: $F(1,23) = 152.76$, $p < .001$, $\eta_p^2 = 0.87$; ER: $F(1,23) = 11.15$, $p = .01$, $\eta_p^2 = 0.33$) and post-hoc ANOVAs (see **Methods**) showed that the conflict effect was present when the auditory feature content was task-relevant ($RT_{cont(I)}$: $F(1,23) = 285.00$, $p < .001$, $\eta_p^2 = 0.93$; $ER_{cont(I)}$: $F(1,23) = 23.85$, $p < .001$, $\eta_p^2 = 0.51$**; Figure 2A, left panel**), but not when all auditory features were task-irrelevant (RT_{VRDM}: $F(1,23) = 1.96$, $p = .18$, $\eta_p^2 = 0.08$, $BF_{01} = 5.41$; ER_{VRDM}: $F(1,23) = 0.26$, $p = .62$, $\eta_p^2 = 0.01$, $BF_{01} = 4.55$; **Figure 2A, right panel**). Because responses in the vertical RDM were made with the right hand only, we subsequently tested whether the auditory features *in isolation* affected the speed and accuracy of right-hand responses. For example, the spoken word 'left' may slow down responses made with the right hand more so than the spoken word 'right' (the same logic holds for stimulus location). However this was not the case. A 2x2x2 factorial rm-ANOVA on mean RTs with session (before/after training), stimulus content ("left"/"right"), and stimulus location (left/right) showed that RTs were unaffected by sound content ($F(1,23) = 0.01$, $p = .92$, $\eta_p^2 = 0.00$, $BF_{01} = 6.16$) and sound location ($F(1,23) = 0.49$, $p = .49$, $\eta_p^2 = 0.02$, $BF_{01} = 6.36$).

Participants performed both behavioral tasks before and after extensive training of the content discrimination task, to be able to investigate the role of training on conflict processing (**Figure 1C**). RTs and ERs in the vertical RDM task were not modulated by behavioral training (RT_{VRDM}: $F(1,23) = 2.07$, $p = .16$, $\eta_p^2 = 0.08$, $BF_{01} = 0.32$; ER_{VRDM}: $F(1,23) = 0.24$, $p = .63$, $\eta_p^2 = 0.01$, $BF_{01} = 3.79$). Training did result in a decrease of overall RT on content discrimination task I, although ERs were not affected ($RT_{cont(I)}$: $F(1,23) = 45.05$, $p < .001$, $\eta_p^2 = 0.66$; $ER_{cont(I)}$: $F(1,23) = 1.77$, $p = .20$, $\eta_p^2 = 0.07$, $BF_{01} = 0.89$). Moreover, the effect of conflict on RTs and ERs in this task decreased after behavioral training ($RT_{cont(I)}$: $F(1,23) = 29.86$, $p < .001$, $\eta_p^2 = 0.57$; $ER_{cont(I)}$: $F(1,23) = 9.76$, $p = .005$, $\eta_p^2 = 0.30$; **Figure 2 – Supplement 1A-B**) suggesting increased efficiency of within-trial conflict resolution mechanisms. All other effects were not reliable ($p > .05$).

Experiment 1: Neural signatures of conflict detection only for task-relevant sensory features

The observation that task-irrelevant conflicting features had no effect on RTs and ERs, even after substantial training, whereas task-relevant conflicting features had, may not come as a surprise, because manual responses on the visual task (motion up/down with index and middle finger of right hand) were fully orthogonal to the potential conflicting nature of the auditory features (i.e. left/right). Further, content discrimination task I and the vertical RDM were independent tasks, requiring different cognitive processes. For example, mean RTs on the vertical RDM were on average 267ms longer than mean RTs for content discrimination task I. However, caution is required in concluding that conflict detection is absent for task-irrelevant features based on these behavioral results alone, as neural and/or behavioral effects can sometimes be observed in isolation (one is observed but not the other, e.g. Canales-Johnson et al., 2020; van Gaal et al., 2014). Therefore, in order to test whether unattended conflict is detected by the brain we turn to the multivariate pattern analysis (MVPA) of our neural data.

Plausibly, the neural dynamics of conflict processing for task-irrelevant sensory input are different -in physical (across electrodes) and frequency space- from those related to the processing of conflict when sensory input is task-relevant. Therefore, we applied multivariate decoding techniques in the frequency domain to inspect whether and—if so—to what extent certain stimulus features were processed. These multivariate approaches

have some advantages over traditional univariate approaches, for example, they are less sensitive to individual differences in spatial topography, because decoding accuracies are derived at a single participant level (Fahrenfort et al., 2018; Grootswagers et al., 2017; Haxby et al., 2001). Therefore, group statistics do not critically depend on the presence of effects in specific electrodes or clusters of electrodes. Further, although a wealth of studies have shown that conflict processing is related to an increase in power of theta-band neural oscillations (~4Hz-8Hz) after stimulus presentation (Cohen & Cavanagh, 2011; Jiang, Zhang, & van Gaal, 2015; Nigbur et al., 2012), it is unknown whether this is also the case for task-irrelevant conflict. By performing our MVPA analysis in frequency space, we could potentially find neural signatures in non-theta frequency-bands related to the processing of task-irrelevant conflict. However, due to the temporal and frequency space that has to be covered, strict multiple comparison corrections have to be performed (across time and frequency, see **Methods**). Therefore, we adopted an additional hypothesis-driven analysis, which also allowed us to obtain evidence for the absence of effects. Throughout this paper, we will discuss our neural data in the following order. First, the MVPA analyses in the frequency domain are presented for all critical features of the task (congruency, content, location, corrected for multiple comparisons). Then, we report results from the additional hypothesis-driven analysis, where we extracted classifier accuracies from a predefined time-frequency ROI (100ms-700ms, 2Hz-8Hz) on which we performed (Bayesian) tests (see **Methods**). This ROI was selected based on previous observations of conflict related theta-band activity (Cohen & Cavanagh, 2011; Cohen & van Gaal, 2014; Jiang, Zhang, & Van Gaal, 2015; Nigbur et al., 2012). Specifically, for every task and every stimulus feature (i.e. congruency, content, location), we extracted average decoding accuracies from the ROI per participant and performed analyses on these values.

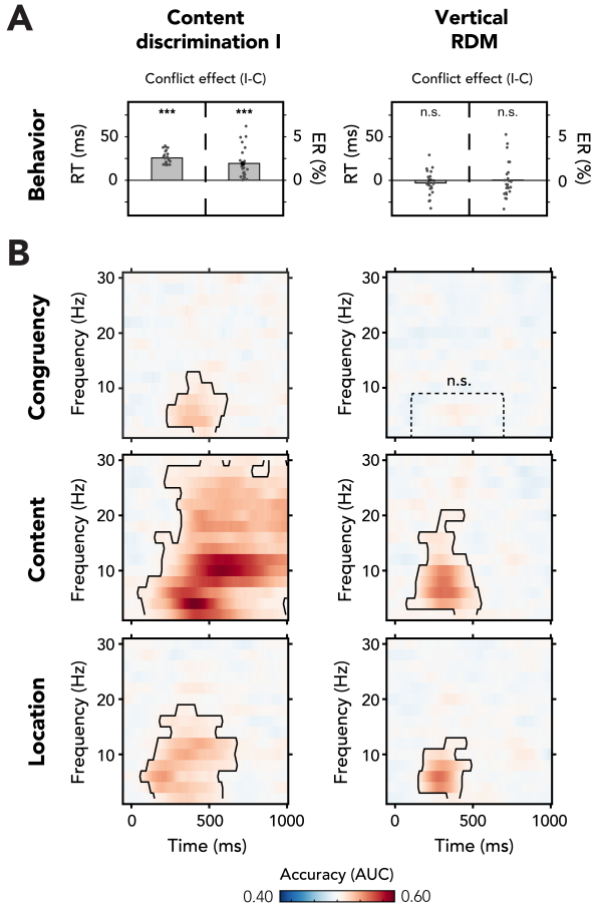

Figure 2. Behavioral and multivariate decoding results of experiment 1. (A/B) All results depicted here are from the merged data of both experimental sessions. The left column of plots shows the results for content discrimination task I, where auditory stimuli and conflicting features were task-relevant. The right column of plots shows the results for the vertical RDM, where neither the auditory stimulus nor its conflicting features were task-relevant. **(A)** The behavioral results are plotted as conflict effects (incongruent – congruent). Effects of conflict were present in content discrimination task I, with longer RTs (left bar) and increased error rates (right bar) for incongruent compared to congruent trials. For the vertical RDM task, no

significant effects of conflict were found in behavior. Dots represent individual participants. The behavioral data that is shown here can be found in **Figure 2 source data**. **(B)** Multivariate classifier accuracies for different stimulus features. We trained classifiers on three stimulus features: auditory congruency, auditory content and auditory location. Classifier accuracies (AUC) are plotted across a time-frequency window of -100ms to 1000ms and 2Hz to 30Hz. Classifier accuracies are thresholded (cluster-based corrected, one-sided: $\bar{x} > 0.5$, $p < .05$) and significant clusters are outlined with a solid black line. The dotted box shows the predefined ROI on which we performed a hypothesis-driven analysis. The classifier accuracies within this ROI were not significantly greater than chance for the vertical RDM task. Note that conflicting features of the auditory stimulus, content and location, could be decoded from neural data regardless of attention to the auditory stimulus. Information related to auditory congruency was present in a theta-band cluster, but only when the auditory stimulus was attended. ***: $p < .001$, **n.s.:** $p > .05$.

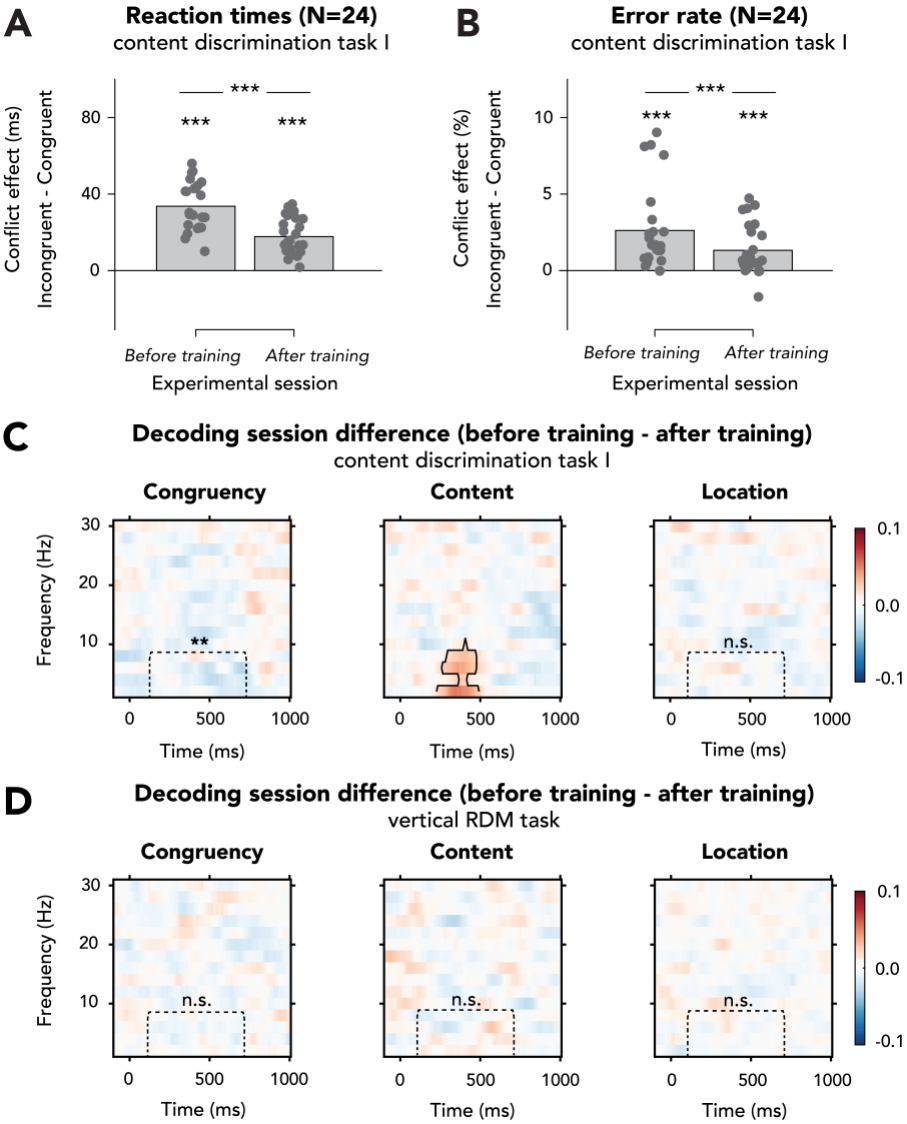

Figure 2 – Supplement 1. Effects of behavioral training on behavioral effects of conflict and decoding performance in Experiment 1. (A/B) We performed 2x2 rm-ANOVAs on **(A)** RTs, **(B)** ERs in content discrimination task I, with the factors being session and congruency of the auditory stimulus. I**n A-B** data are plotted as conflict effects (incongruent – congruent) and for separate sessions. The top horizontal line shows significance of the interaction between session and congruency and markers above the bars indicate significance of paired sample t-tests comparing incongruent and congruent for each run (shown data and results of t-tests can be found in **Figure 2 – Supplement 1 source data**). Effects of conflict on RTs **(A)** and ERs **(B)** significantly decreased after behavioral training on this task, suggesting more efficient processing of conflict. Effects of conflict on RTs and ERs were nonetheless present during both sessions. **(C)** There were no clusters for which the difference in congruency decoding between the two sessions in content discrimination task I was significant (left panel), although decoding accuracies within the preselected ROI did decrease with training for content discrimination task I, suggesting more efficient conflict resolution, in line with the behavioral results plotted in **A-B**. Classifier accuracies for sound content (middle panel) were higher in a delta-theta band cluster after behavioral training, showing that the task-relevant feature was processed better. Location decoding accuracy was not affected by behavioral training, as we observed no clusters where the

difference between sessions and classifier accuracies within the ROI were also not different between sessions. **(D)** Behavioral training on the content discrimination task did not affect neural processing of auditory features in the vertical RDM (no significant clusters and none of the results were significant when tested for the predefined ROI). Thresholded (cluster-based corrected, $p < .05$) accuracies are depicted across the frequency-range (2-30Hz) and significant clusters are outlined with a solid black line. ***: $p < .001$, **: $p < .01$, **n.s.**: $p > .05$.

First, we trained a classifier on data from all EEG electrodes to distinguish between congruent vs. incongruent trials, for both content discrimination task I and the vertical RDM task. Above-chance classification accuracies imply that relevant information about the decoded stimulus feature is present in the neural data, meaning that some processing of that feature occurred (Hebart & Baker, 2017). We performed our main analysis on the combined data from both EEG sessions, thereby maximizing power to establish effects in our crucial comparisons. We also performed similar analyses on the session specific data to investigate the role of behavioral training on processing of conflict. These results are discussed more in depth below and are shown in **Figure 2 – Supplement 1C-D.**

Congruency decoding reveals that stimulus congruency was represented in neural data only when conflict was task-relevant ($p < .001$, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 2Hz-12Hz, peak frequency: 4Hz, time-range: 234ms-609ms, peak time: 438ms; **Figure 2B**, left panel). The conflict effect roughly falls in the theta-band (4Hz-8Hz), which confirms a firm body of literature linking conflict detection to post-conflict modulations in theta-band oscillatory dynamics (Cavanagh & Frank, 2014; Cohen & Cavanagh, 2011; Cohen & van Gaal, 2014; Nigbur et al., 2012). Activation patterns that were calculated from classifier weights within the predefined time-frequency theta-band ROI (2Hz-8Hz, 100ms-700ms) revealed a clear midfrontal distribution of conflict related activity (**Figure 5 – Supplement 1A**). No significant time-frequency cluster was found for the vertical RDM task (**Figure 2B**, right panel). To quantify the absence of this effect, we followed up this hypothesis-free (with respect to frequency and time) MVPA-analysis with a hypothesis-driven analysis focused on the post-stimulus theta-band. This more restricted analysis showed no significant effect ($t(23) = -0.50$, $p = .69$, $d = -0.10$) and an additional Bayesian analysis revealed moderate evidence in favor of the null hypothesis (i.e. no effect of conflict on theta-band power) than the alternative hypothesis ($BF_{01} = 6.53$).

Similar to our observation of decreased behavioral effects of conflict after behavioral training (**Figure 2 – Supplement 1A-B**), decoding accuracies in the content discrimination task I were also lower after training ($t(23) = -3.01$, $p = .01$, $d = -0.63$; **Figure 2 – Supplement 1C**), suggesting more efficient conflict resolution, as reflected in neural theta oscillations as well. In the vertical RDM, behavioral training did not affect decoding accuracies of sound congruency ($t(23) = -1.24$, $p = .23$, $d = -0.25$, $BF_{01} = 2.36$; **Figure 2 – Supplement 1D**).

Experiment 1: Stimulus features are processed in parallel, independent of task relevance

Thus, cognitive control networks -or possible substitute networks- are seemingly not capable of detecting conflict when sensory features are task-irrelevant. However, the question remains whether this observation is specific to the conflicting nature of the auditory stimuli, or whether the auditory stimuli are not processed whatsoever when attention is reallocated to the visually demanding task. To address this question, we trained classifiers on two other features of the auditory stimuli, i.e. location and content, to test whether these features were processed by the brain regardless of task relevance. Indeed the content of auditory stimuli was processed both when the stimuli were task-relevant ($p < .001$, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 2Hz-30Hz, peak frequency: 4Hz, time-range: 47ms-1000ms, peak time: 422ms; **Figure 2B**, left panel) and task-irrelevant ($p < .001$, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 2Hz-20Hz, peak frequency: 6Hz, time-range: 78ms-547ms, peak time: 297ms; **Figure 2B**, right panel).

Similarly, the location of auditory stimuli could also be decoded from neural data for both content discrimination task I ($p < .001$, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 2Hz-18Hz, peak frequency: 6Hz, time-range: 63ms-672ms, peak time: 203ms; **Figure 2B**, left panel) and the vertical RDM task ($p < .001$, one-sided:

$\bar{x}$>0.5, cluster-corrected; frequency-range: 2Hz-12Hz, peak frequency: 6Hz, time-range: 156ms-484ms, peak time: 281ms; **Figure 2B**, right panel). The above chance performance of the classifiers for the auditory stimulus features demonstrates that location and content information were processed, even when these features were task-irrelevant. Processing of task-irrelevant stimulus features was, however, more transient in time and more narrow-band in frequency as compared to processing of the same features in a task-relevant setting. Further, content decoding revealed a much broader frequency spectrum than any of the other comparisons in content discrimination task I. In the next experiment we show that this is related to the fact that this feature was response relevant and that this effect therefore partially reflects response preparation and response execution processes. Summarizing, we show that when (conflicting) features of an auditory stimulus are truly and consistently task-irrelevant, the conflict between them is no longer detected by –nor relevant to– the conflict monitoring system, but the features (content and location) are still processed in isolation.

To investigate if, and how, behavioral training affects processing of sound content and location we tested whether decoding accuracies for these features were different between the two experimental sessions. Decoding accuracies for the task-relevant feature of content discrimination task I (i.e. sound content) were significantly increased after behavioral training in a delta- to theta-band cluster (p < .001, one-sided: $\bar{x}$ > 0.5, cluster-corrected; frequency-range: 2Hz-10Hz, peak frequency: 2Hz, time-range: 234ms-484ms, peak time: 344ms; **Figure 2 – Supplement 1C**). This suggests that processing of task-relevant information (i.e. sound content) is improved as a results of training. Decoding accuracies for sound location in the content discrimination task were not different before and after behavioral training (no significant clusters; predefined ROI: $t(23) = 0.12$, $p = .91$, $d = 0.02$, $BF_{01} = 4.63$). In the vertical RDM task, behavioral training did not affect the decoding accuracies within the predefined ROI for sound content ($t(23) = 0.75$, $p = .46$, $d = 0.15$, $BF_{01} = 3.61$) and location ($t(23) = 0.04$, $p = .97$, $d = 0.01$, $BF_{01} = 4.66$, also no other significant clusters; **Figure 2 – Supplement 1D**). This suggests that processing of sound content and location (**Figure 2B**), both task-irrelevant auditory features in the vertical RDM task, is automatic and not dependent on training.

In conclusion, we observed neural signatures of the processing of sensory stimulus features (i.e. location and content of an auditory stimulus) regardless of task relevance of these features, but a lack of integration of these features to form conflict when the auditory stimulus was fully task-irrelevant. Considerable training in content discrimination task I resulted in more efficient conflict processing (i.e. decreased behavioral conflict effects and theta-band activity after training; **Figure 2 – Supplement 1**) when the auditory stimulus was task-relevant, but this increased automaticity did not lead to detection of conflict when the auditory stimulus was fully task irrelevant.

Experiment 2: Does detection of conflict depend on task relevance of the stimulus or its individual features?

The experimental design of the first experiment rendered the auditory features to be located at the extreme ends of the scale of task relevance, i.e. either the conflicting features were task-relevant and the conflicting features were consistently mapped to specific responses, or the conflicting features were task-irrelevant and the conflicting features were not mapped to responses. However, to further understand the relationship between the relevance of the conflicting features and the overlap with responses, we performed a second experiment containing four behavioral tasks. For this second experiment we recruited 24 new participants. We included two auditory conflicting tasks, similar to content discrimination task I. In one of the auditory tasks (from hereon: content discrimination task II, **Figure 3A**) participants again had to respond according to the content of the auditory stimulus, whereas in the other auditory task (from hereon: location discrimination task, **Figure 3B**) they were instructed to report from which side the auditory stimulus was presented (i.e. left or right speaker).

Furthermore, we included two new tasks in which the conflicting features (location and content) were not task-relevant and participants responded to a non-conflicting feature that was part of the conflicting stimulus (from

hereon: volume oddball detection task, **Figure 3C**) or the auditory stimulus was task-irrelevant but its features location and content overlapped with the responses to be given (from hereon: horizontal RDM task, **Figure 3D**). The horizontal RDM task was similar to vertical RDM task of Experiment 1, however the dots were now moving on a horizontal plane. In other words, participants were instructed to classify the overall movement of moving dots to either the left or the right. As this is a visual paradigm, the simultaneously presented auditory stimuli are fully task-irrelevant. However, both features of conflict, the content (i.e. 'left' and 'right') and the location (i.e. left and right speaker), of the auditory stimuli could potentially interfere with participants' responses on the visual task, thereby inducing a crossmodal Stroop-like type of conflict (Stroop, 1935).

In the volume oddball detection task, participants were presented with the same auditory stimuli as before, however one out of eight stimuli (12.5%) was presented at a lower volume. Participants were instructed to detect these volume oddballs by pressing the spacebar with their right hand as fast as possible. If they did not hear an oddball, they were instructed to withhold from responding. In this task, theoretically, the selection of an object's feature (e.g. volume) could lead to the selection of all of its features (e.g. sound content, location), as suggested by theories of object-based attention (Chen, 2012). This in turn may lead to conflict detection, even if the conflicting features are task-irrelevant. Similar to Experiment 1, we included behavioral training in conflict inducing tasks to inspect if enhanced automaticity of conflict processing would affect conflict detection under task-irrelevant sensory input. Participants performed 500 trials of the volume oddball detection task twice, at the very beginning of a session and at the end of a session (**Figure 3E**). During the first run of the task neither sound content nor sound location was related to any behavioral responses, whereas during the second run these features might have acquired some intrinsic relevance through training on the other tasks. Furthermore, repeated exposure to conflict may prime the conflict monitoring system to exert more control over sensory inputs necessary for more efficient conflict detection, even when these sensory inputs are not task-relevant within the context of the task the participant is performing at that time.

In order to keep sensory input similar, moving dots (coherence: 0) were presented on the monitor during content discrimination task II, the location discrimination task and the volume oddball detection task, but these could be ignored. Again, EEG was recorded while participants performed these tasks, in order to see if auditory conflict was detected when the auditory stimulus, or its conflicting features (i.e. location and content) were task-irrelevant. We performed the same artefact rejection procedure as in Experiment 1. For one participant, on average 64.5% (*SD* = 9.9%) of all epochs within each task were removed in this procedure, which is 3.9 standard deviations from the average ratio of removed epochs in this experiment (*M* = 10.3%, *SD* = 13.9%). Therefore, this participant was excluded from the EEG analysis of Experiment 2, resulting in N=23 for the analysis of EEG data.

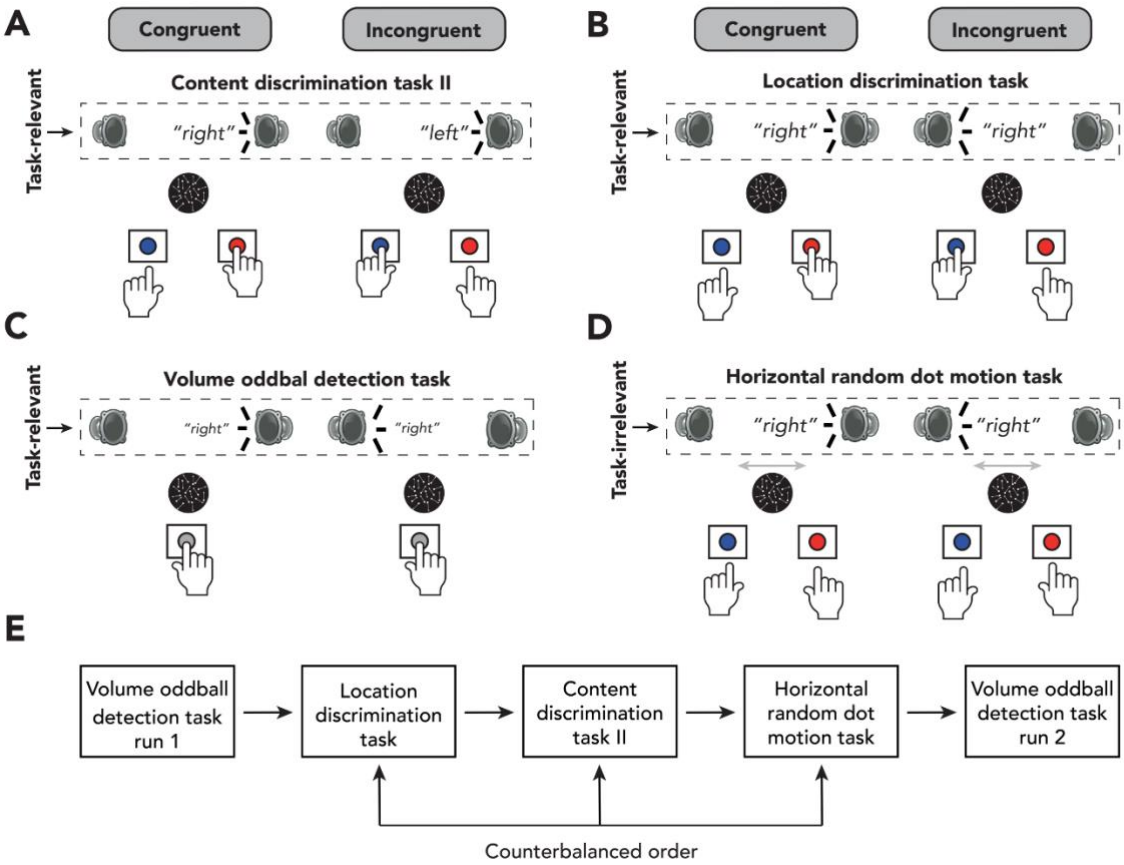

Figure 3. Experimental design of experiment 2. (A/B/C/D) Schematic representation of the experimental design for auditory content discrimination task II **(A)**, location discrimination task **(B)**, volume oddball task **(C)** and horizontal RDM task **(D)**. In all tasks, the spoken words "left" and "right were presented through either a speaker located on the left or right side of the participant. **(A)** In auditory content discrimination task II, participants were instructed to report the content ("left" or "right") of an auditory stimulus via a button press with their left or right hand, respectively, and to ignore the location of the auditory stimulus that was presented. **(B)** In the auditory location discrimination task, participants were instructed to report the location (left or right speaker) of an auditory stimulus via a button press with their left or right hand respectively, and to ignore the content of the auditory stimulus that was presented. **(C)** During the volume oddball task, participants were instructed to detect auditory stimuli that were presented at a lower volume than the majority of the stimuli (i.e. oddballs) by pressing the spacebar with their right hand. **(D)**. In the horizontal RDM, participants were instructed to report the overall movement of dots (left or right) via a button press with their left and right hands respectively, whilst still being presented with the auditory stimuli. In all four tasks, content of the auditory stimuli could be congruent or incongruent with its location of presentation (50% congruent/incongruent trials). **(E)** Order of behavioral tasks in experiment 2. Participants always started with the volume oddball task, followed by the location discrimination task, content discrimination task and horizontal RDM, in randomized order. Participants ended with another run of the volume oddball task.

Experiment 2: Behavioral effects of conflict only for task-relevant auditory sensory input

Mean RT in the location discrimination task was 338.2ms (*SD* = 112.7ms) and mean ER was 4.5% (*SD* = 3.2%). For content discrimination task II, RTs were on average 364.0ms (*SD* = 127.4ms) and ERs were 5.5% (*SD* = 5.8%). For the horizontal RDM, RTs were on average 362.5ms (*SD* = 116.5ms) and ERs were 27.7% (*SD* = 5.2%). Mean RTs of the volume oddball task were calculated for correct trials in which a response was made (i.e. hit trials) and was 504.5s (*SD* = 178.7ms, on hit trials). On average participants had 40.5 hits (*SD* = 12.8) out of 61.8 oddball trials (*SD* = 8.6), per run of 500 trials. We will first discuss the behavioral results of the content discrimination, location discrimination and horizontal RDM tasks, as behavioral performance for the volume oddball task is represented in d', rather than ER.

Repeated measures ANOVAs (3x2 factorial) were performed on mean RTs and ERs from these three tasks, with the factors 1) task and 2) congruency of the auditory features. Again, congruency always relates to the

combination of the auditory stimulus features sound content ("left" vs. "right") and sound location (left speaker vs. right speaker). We observed that participants were slower and made more errors on incongruent trials as compared to congruent trials (RT: $F(1,23) = 75.41$, $p < .001$, $\eta_p^2 = 0.77$; ER: $F(1,23) = 68.00$, $p < .001$, $\eta_p^2 = 0.75$; **Figure 4A**). This conflict effect was modulated by task (RT: $F(1.55,35.71) = 22.80$, $p < .001$, $\eta_p^2 = 0.50$; ER: $F(1.58, 36.36) = 10.18$, $p < .001$, $\eta_p^2=0.31$) and post-hoc paired samples t-tests (incongruent – congruent) showed that conflict effects were only present in tasks where one of the conflicting features was task-relevant (location discrimination task: RT_{loc}: $t(23) = 5.03$, $p < .001$, $d = 1.03$; ER_{loc}: $t(23) = 6.25$, $p < .001$, $d = 1.28$; content discrimination task II: $RT_{cont(II)}$: $t(23) = 8.95$, $p < .001$, $d = 1.83$; $ER_{cont(II)}$: $t(23) = 5.93$, $p < .001$, $d = 1.21$; horizontal RDM task: RT_{HRDM}: $t(23) = 1.44$, $p = .16$, $d = 0.29$, $BF_{01} = 1.88$; ER_{HRDM}: $t(23) = 1.65$, $p = .11$, $d = 0.34$, $BF_{01} = 1.44$). Although in the horizontal RDM task, conflict between sound content and location did not affect the speed of responses, stimulus content and location *in isolation* could have potentially interfered with behavioral performance given the overlap of these features with both the plane of dot direction (left/right) and the response scheme (left/right hand). Indeed, trials containing conflict between sound location and dot direction resulted in slower RTs and increased ERs (RT: $t(23) = 2.12$, $p = .045$, $d = 0.44$; ER: $t(23) = 5.94$, $p < .001$, $d = 1.21$). Similar effects were observed for trials where sound content conflicted with the dot-direction, most prominently observed in RTs (RT: $t(23) = 1.72$, $p = .10$, $d = 0.35$, $BF_{01} = 2.85$; ER: $t(23) = 5.55$, $p < .001$, $d = 1.13$). This shows that sound content and location *in isolation*, even though both features were task-irrelevant, interfered with task performance.

For the volume oddball task, we tested the effect of auditory congruency on RTs and d' of trials which, by virtue of task-instruction, only covers oddball trials in which a correct response was made (i.e. hits). Repeated measures ANOVAs (2x2 factorial) with the factors run number and feature congruency revealed no effects of auditory conflict on RT ($F(1,23) = 2.78$, $p = .11$, $\eta_p^2 = 0.10$, $BF_{01} = 3.34$; **Figure 4A,** no effects of training, see **Figure 4 – Supplement 1A**). To test whether *individual* features of the auditory stimulus interfered with right-hand responses, we performed an additional 2x2 factorial rm-ANOVA with sound content and location as factors. Auditory content significantly affected RTs, whereas sound location did not (content: $F(1,23) = 25.41$, $p < .001$, $\eta_p^2 = 0.53$; location: $F(1,23) = 0.99$, $p = .33$, $\eta_p^2 = 0.04$, $BF_{01} = 3.90$). Specifically, RTs were slower for the spoken word "left" (incongruent with the responding hand, $M = 528.6$ms, $SD = 193.1$ms) as compared to the spoken word "right" (congruent with the responding hand, $M = 493.3$ms, $SD = 174.9$ms), revealing interference of sound content *in isolation* on right hand responses, similar to the horizontal RDM task. Although conflict between the auditory features was not present in RTs, we did observe that sensitivity (d') increased for incongruent ($M = 2.66$, $SD = 0.77$) compared to congruent trials ($M = 2.11$, $SD = 0.81$, $F(1,23) = 45.62$, $p < .001$, $\eta_p^2 = 0.67$; **Figure 4A**). These results show that volume oddball detection performance increases on trials that contain conflict between sound content and location. This effect of conflict on behavioral performance can already be found in the first run, when sound content and location had not yet been related to any responses/task and were thus fully task-irrelevant ($t(23) = 5.71$, $p < .001$, $d = 1.17$). There was no significant interaction between run number and auditory stimulus congruency ($F(1,23) = 1.40$, $p = .25$, $\eta_p^2 = 0.06$, $BF_{01} = 2.13$; **Figure 4 – Supplement 1B,** hit rates and false alarms are plotted in this supplementary figure as well).

Experiment 2: Detection of conflict occurs when any feature of a stimulus is related to the response on the primary task or task-relevant (even a non-conflicting feature)

We again trained multivariate classifiers on single-trial time-frequency data to test whether the auditory stimulus features (i.e. content, location and congruency) were processed when (I) the auditory conflicting features were task-relevant and overlapped with the response-mapping (content and location discrimination tasks), (II) the auditory conflicting features were task-irrelevant and another stimulus of the conflicting stimulus was task-relevant (volume oddball task), or when (III) the auditory conflicting features was task-irrelevant, but its conflicting features overlapped with the response-mapping in the task (horizontal RDM task).

Cluster-based analyses across the entire T-F range, revealed neural signatures of conflict processing in the theta-

band when the content of the auditory stimulus was task-relevant (content discrimination task II: $p < .001$, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 4Hz-10Hz, peak frequency: 4Hz, time-range: 328ms-953ms, peak time: 438ms; **Figure 4B**) and when the volume of the auditory stimulus was task-relevant (volume oddball task: $p = .03$, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 2Hz-6Hz, peak frequency: 2Hz, time-range: 234ms-516ms, peak time: 438ms; **Figure 4B,** see **Figure 4 – Supplement 1E:** no training effects for the volume oddball task). Both observations of congruency decoding are in line with the presence of conflict-related behavioral effects in these tasks (**Figure 4A**). No significant clusters of above-chance classifier accuracy were found after correcting for multiple comparisons in the location discrimination task and the horizontal RDM task (**Figure 4B**). However, a hypothesis driven analysis focused on the post-stimulus theta-band (2Hz-8Hz, 100ms-700ms) revealed that congruency decoding accuracies within this ROI were significantly above chance for both tasks as well (location discrimination: $t(22) = 2.00$, $p = .03$, $d = 0.42$; horizontal RDM: $t(22) = 2.89$, $p = .004$, d=0.60). Thus, we observed that conflict of the auditory stimulus is detected when one of the auditory conflicting features is task-relevant (content and location discrimination tasks), when one of its non-conflicting features is task relevant (volume oddball task), and when none of the auditory features is task-relevant but these features overlap with the response mapping of the task (horizontal RDM task). To qualify the differences between tasks, we combine the data from all experiments and compare effect sizes across tasks at the end of this **Results** section (**Figure 5**).

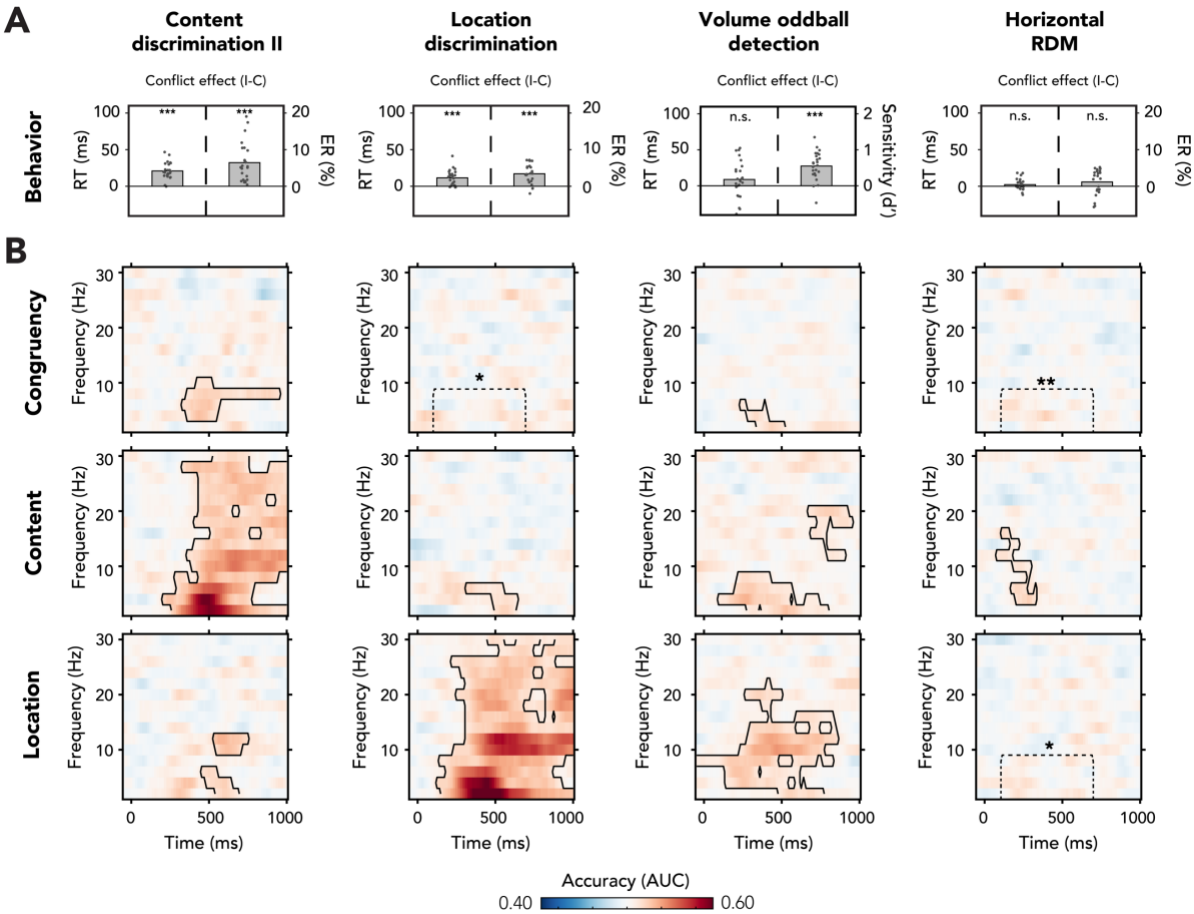

Figure 4. Behavioral and multivariate decoding results for experiment 2. (A/B) The four columns show data belonging to, from left to right, content discrimination task II, the location discrimination task, the volume oddball detection task and the horizontal RDM task. **(A)** Behavioral results are plotted as conflict effects (incongruent – congruent). Effects of conflict were present in all tasks where the auditory stimulus was task-relevant (content discrimination task II, location discrimination task and volume oddball). In both auditory discrimination tasks we observed longer RTs (left bar) and increased error rates (right bar) for incongruent compared to congruent trials. For the volume oddball we did not observe an effect in RT, but increased sensitivity (d') on incongruent compared to congruent trials. Dots represent individual participants. The data that is shown here can be found in **Figure 4 source data**. **(B)** Multivariate classifier accuracies for different stimulus features (auditory

congruency, auditory content and auditory location). Classifier accuracies (AUC) are plotted across a time-frequency window of -100ms to 1000ms and 2Hz to 30Hz. Classifier accuracies are thresholded (cluster-based corrected, one-sided: $\bar{x} > 0.5$, $p < .05$) and significant clusters are outlined with a solid black line. The dotted box shows the predefined ROI on which we performed a hypothesis-driven analysis. Note that the data shown for the volume oddball task was merged over both runs. ***:** $p < .05$**, **:** $p < .01$**, ***:** $p < .001$**; n.s.:** $p > .05$**.**

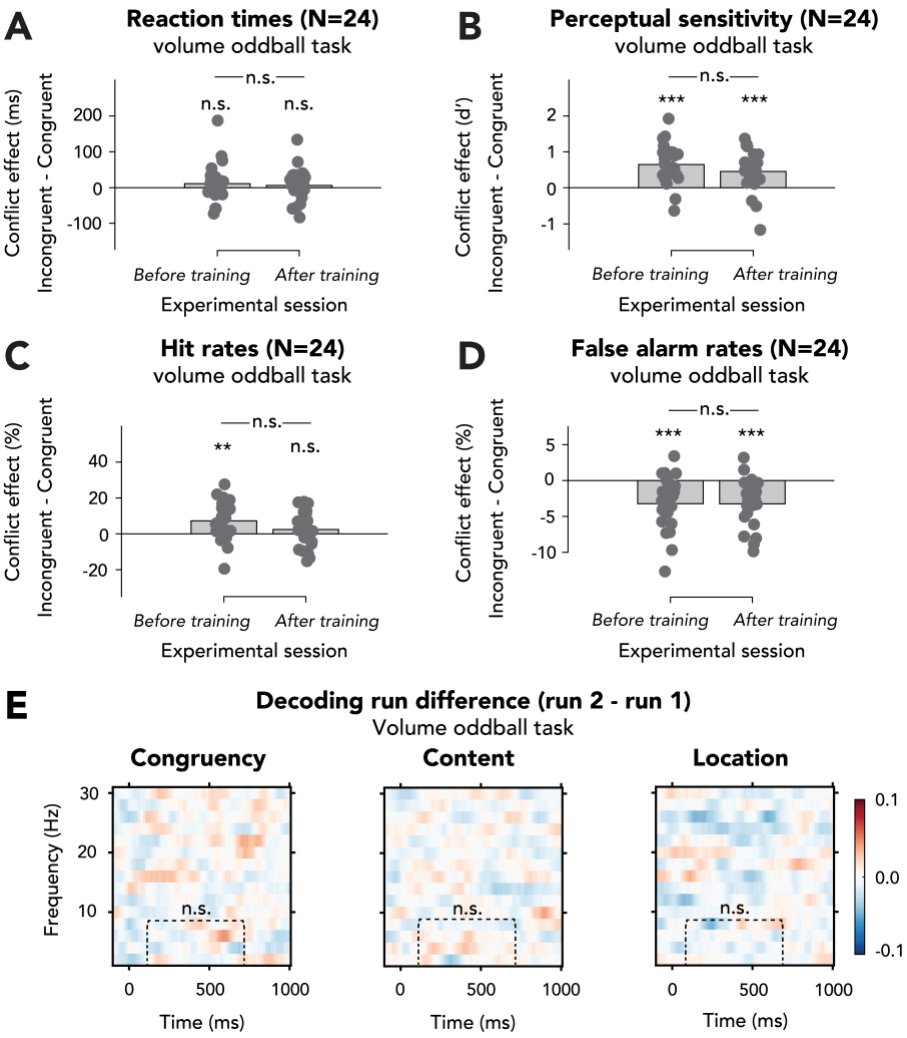

Figure 4 – Supplement 1. Effects of exposure to conflict inducing task on behavioral effects of conflict and decoding performance in the volume oddball task of Experiment 2. (A-D) We performed 2x2 rm-ANOVAs on **(A)** RTs, **(B)** perceptual sensitivity (d'), **(C)** hit rates and **(D)** false alarm rates with the factors run number and congruency of the auditory stimulus. I**n A-D** data are plotted as conflict effects (incongruent – congruent) and for separate runs. The top horizontal line shows significance of the interaction between session and congruency and markers above the bars indicate significance of paired sample t-tests comparing incongruent and congruent for each run (shown data and results of t-tests can be found in **Figure 4 – Supplement 1 source data**).**(A)** RTs were unaffected by auditory congruency and run number (statistics in **Results**). **(B)** There was no interaction effect between congruency and run number on perceptual sensitivity (d'; statistics in **Results**) and post-hoc paired sample t-tests (incongruent – congruent) revealed that the effect of congruency on d' was present during both runs. **(C)** The interaction between congruency and run number was not significant ($F(1,23) = 2.99$, $p = .10$, $\eta_p^2 = 0.12$, $BF_{01} = 1.64$), showing that the effect of conflict on hit rate was not different for both runs, although the effect of conflict was present during the first, but not second run. **(D)** False alarm rates were not modulated by the interaction between congruency and run number ($F(1,23) = 0.00$, $p = .99$, $\eta_p^2 = 0.00$, $BF_{01} = 3.45$), showing that the effects of conflict were not different between runs. This conflict effect was present during both runs. **(E)** There were no clusters for which the difference in decoding of all features between the two runs of the volume oddball task was significant and there were also no differences within the pre-selected ROI (congruency: $t(22) = 0.07$, $p = .95$, $d = 0.01$, $BF_{01} = 4.56$; content: $t(22) = 0.64$, $p = .53$, $d = 0.13$, $BF_{01} = 3.81$; location: $t(22) = -1.25$, $p = .22$, $d = -0.26$, $BF_{01} = 2.29$), suggesting that processing of these features was not affected by training. Thresholded (cluster-based corrected, $p < .05$) accuracies are depicted across the frequency-range (2-30Hz). Plots show the difference in classifier accuracy between the two runs (run 2 - run 1) of stimulus congruency (left panel), stimulus content (middle panel) and stimulus location (right panel) in the volume oddball task. **n.s.**: $p > .05$.

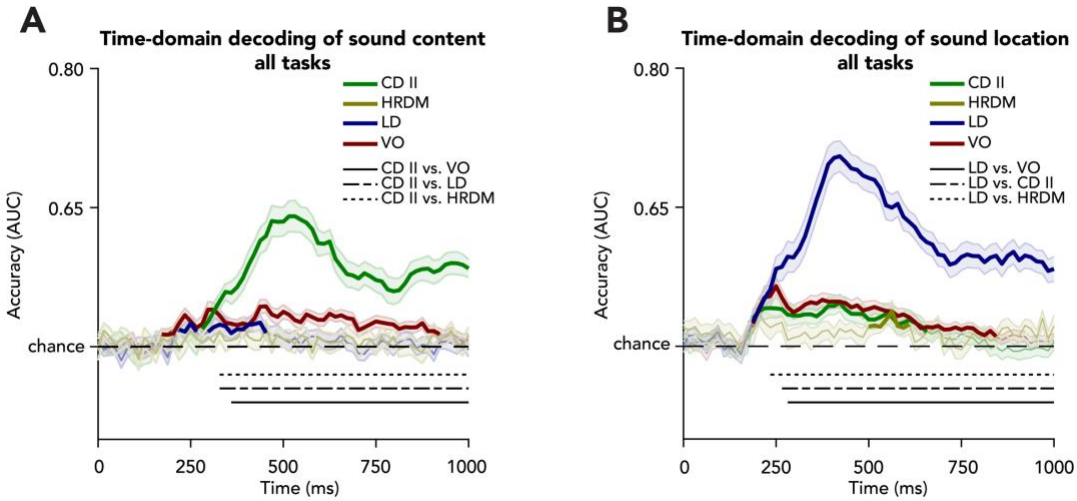

Figure 4 - Supplement 2. Sensory feature decoding in the time-domain. We trained classifiers on either sound content (**A**) or sound location (**B**) in order to see how neural representations of sensory processing were affected by our manipulation of task relevance of these features. (**A**) Sound content could be decoded from most tasks (except the horizontal RDM) and decoding accuracies for sound content were highest for the task in which sound content was the task-relevant feature, i.e. content discrimination task II. Decoding accuracies were higher for content discrimination task II as compared to the other three tasks (difference start location discrimination: 328ms, horizontal RDM: 313ms, volume oddball: 344ms). (**B**) Sound location could be decoded from all tasks. Again, the task in which the decoded feature was task-relevant, i.e. the location discrimination task, showed the highest decoding accuracies. Location decoding performance was improved for the task in which this feature was task-relevant (i.e. location discrimination task) as compared to the other tasks. These differences started from 250ms (vs. content discrimination task II), 234ms (vs. horizontal RDM task), and 266ms (vs. volume oddball task). Shaded areas represent the SEM. Bold traces indicate that feature decoding was significantly (cluster-corrected, one-sided t-test, $\bar{x} > 0.5$, $p < .05$) above chance. Horizontal black lines at the bottom depict where feature decoding is significantly different (cluster-corrected, two-sided t-test, p<0.05) between the task in which the feature was task-relevant versus where it was task-irrelevant. *Abbreviations*: **CD II**: content discrimination task II, **HRDM**: horizontal RDM task, **LD**: location discrimination task, **VO**: volume oddball detection task.

Experiment 2: Sensory features are processed in all tasks

Next, we trained classifiers to distinguish trials based on sound location and content in order to inspect sensory processing. We found neural signatures of the processing of sound content for all four tasks: content discrimination II ($p < .001$, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 2Hz-30Hz, peak frequency: 4Hz, time-range: 203ms-1000ms, peak time: 469; **Figure 4B**), location discrimination ($p = .01$, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 2Hz-6Hz, peak frequency: 2Hz, time-range: 313-641ms, peak time: 563ms; **Figure 4B**), horizontal RDM task (p < .001, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 4Hz-16Hz, peak frequency: 4Hz, time-range: 78ms-328ms, peak-time: 281ms; **Figure 4B**) and the . For the volume oddball task we observed a delta/theta-band cluster and a late beta-band cluster (delta/theta: $p < .001$, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 2Hz-8Hz, peak frequency: 4Hz, time-range: 94ms-797ms, peak-time: 281ms; late beta: $p = .01$, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 12Hz-20Hz, peak frequency: 20Hz, time-range: 672ms-953ms, peak-time: 828ms; **Figure 4B**).

Furthermore, sound location could be decoded from the content discrimination task II (delta/theta: $p = .02$, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 2Hz-6Hz, peak frequency: 2Hz, time-range: 453ms-688ms, peak time: 609ms; alpha: $p = .03$, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 10Hz-12Hz, peak frequency: 12Hz, time-range: 531ms-750ms, peak time: 578ms; **Figure 4B**), the location discrimination task ($p < .001$, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 2Hz-30Hz, peak frequency: 2Hz, time-range: 109ms-1000ms, peak time: 453ms; **Figure 4B**) and the volume oddball task ($p < .001$, one-sided: $\bar{x} > 0.5$, cluster-corrected; frequency-range: 2Hz-22Hz, peak frequency: 10Hz, time-range: -47ms-891ms, peak time: 469ms; **Figure 4B**). Initially, we did not observe a significant cluster of location decoding in the horizontal RDM, although the hypothesis-driven analysis revealed that accuracies within the predefined theta-band ROI were significantly

above chance-level as well ($t(23) = 2.47$, $p = .01$, $d = 0.51$).

One aspect of these results is worth highlighting. When participants responded to the location of the auditory stimulus, location decoding revealed a broadband power spectrum, similar to sound content decoding when sound content was task-relevant (content discrimination tasks). This broad frequency decoding may be due to the fact that these features were task-relevant, but these results may also partially reflect response preparation and response execution processes, as these auditory features were directly associated with a specific response. In order to test whether the earliest sensory responses were already modulated by task relevance, and to link this to previous ERP studies (Alilović et al., 2019; Molloy et al., 2015; Woldorff et al., 1993), we performed an additional time-domain multivariate analysis on these sensory features (T-F analysis are not well suited to address questions about the timing of processes). Because we were interested in the earliest sensory responses we performed this analysis on data from Experiment 2 only, as all task parameters were best matched (e.g. in all tasks a visual stimulus was presented, no training etc.). We observed increased decoding for task-relevant sensory features, compared to task-irrelevant features, starting ~250ms (sound location RT: M = 338ms) and ~330ms (sound content RT: M = 364ms) after stimulus presentation (**Figure 4 – Supplement 2**). The onset of these differences starts before a response is made, which may suggest that sensory processing of these features is indeed affected by task relevance, however, processes building up towards motor execution, such as decision-making and response preparation processes, cannot be excluded as potential factors driving the higher decoding accuracies in tasks where specific features are task-relevant and hence correlated with decision and motor processes. These results are elaborated upon in the **Discussion**.

In conclusion, in line with the behavioral results, we observed that the processing of conflict between two stimulus features (i.e. location and content of an auditory stimulus) was present in all tasks of Experiment 2. This indicates that conflict can be detected when one of the auditory conflicting features is task-relevant (content and location discrimination tasks), when one of its non-conflicting features is task relevant (volume oddball task), and when there is overlap in the response mapping with any of its task-irrelevant conflicting features (horizontal RDM task). Overall, this reveals that when the conflicting stimulus itself is attended, or when its conflicting features overlap with the response scheme, all of its features seem to be processed and integrated to elicit conflict.

All experiments: Decreasing task relevance hampers cognitive control operations, but not sensory processing

The neural data from all six different tasks over two experiments suggests that if sensory input is task-irrelevant, processing of that information is preserved, while cognitive control operations are strongly hampered (**Figures 2B** and **4B**, **5B**). To quantify this observation, we calculated Cohen's d_z for all tasks and features (based on the preselected ROI), sorted tasks according to the effect sizes of congruency decoding and plotted Cohen's d_z across tasks for all features (conflict, content and location, **Figure 5B**). To be able to compare results from the different tasks we used the same time-frequency ROI ranging from 2Hz-8Hz and 100ms-700ms post-stimulus (**Figure 5A**). For each feature and task, we extracted individual classifier AUC values and performed an analysis of covariance (ANCOVA) on these accuracies, with fixed effects being task and stimulus feature. We found main effects for behavioral task and stimulus feature (task: $F(5,402) = 17.25$, $p < .001$, $\eta_p^2 = 0.18$; stimulus feature: $F(2,402) = 44.61$, $p < .001$, $\eta_p^2 = 0.18$). Crucially, the interaction between task and stimulus feature was also significant ($F(10,402) = 18.80$, $p < .001$, $\eta_p^2 = 0.32$), showing that the accuracy of conflict decoding decreased more across tasks as compared to content and location decoding. We next performed (Bayesian) t-tests (one-sided, $\bar{x} > 0.5$) on ROI accuracies from every task/feature-combination to assess decoding performance for all stimulus features for the different task relevance manipulations. We observed that congruency decoding accuracies were strongly influenced by task, whereas this was not the case for decoding accuracies of stimulus content and location (**Figure 5B**). Note that these results were robust and not dependent on the specific ROI that was selected,

because using other ROI windows led to similar patterns of results, i.e. decreased congruency decoding under task-irrelevant input, but relatively stable sensory feature decoding (**Figure 5 – Supplement 1C-D**). Classifier weights were extracted from the ROI for all tasks and features, transformed to activation patterns and plotted in topomaps, to show the patterns of activity underlying the decoding results (**Figure 5 – Supplement 1A**).

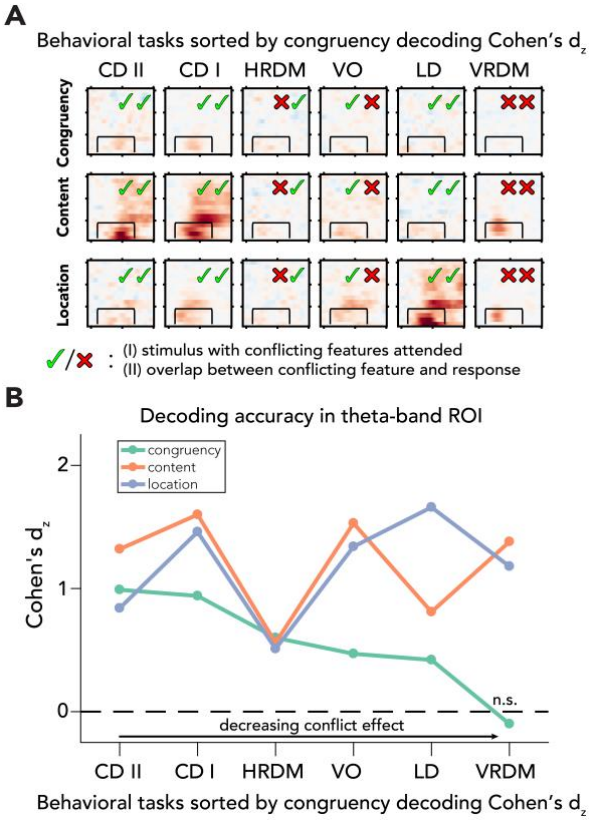

Figure 5. Processing of sensory and conflict features for different levels of task-relevance. (**A)** Summary of the decoding results of all behavioral tasks, sorted by congruency decoding effect size (Cohen's d_z) in a preselected time-frequency ROI. The data in these plots are identical to the ones shown in **figures 2** and **4**. (**B**) Effect sizes are shown for all task/feature combinations derived from a predefined ROI (2Hz-8Hz and 100ms-700ms) and sorted according to effect size of congruency decoding. Effect sizes for congruency decoding were dependent on behavioral task (downward slope of the green line), whereas this was not the case, or less so, for the decoding of content and location. The data can be found in **Figure 5 source data**. *Abbreviations:* **CD II**: content discrimination task II, **CD I**: content discrimination task I, **HRDM**: horizontal RDM task, **VO**: volume oddball detection task, **LD**: location discrimination task, **VRDM**: vertical RDM task, **n.s.**: $p > .05$.

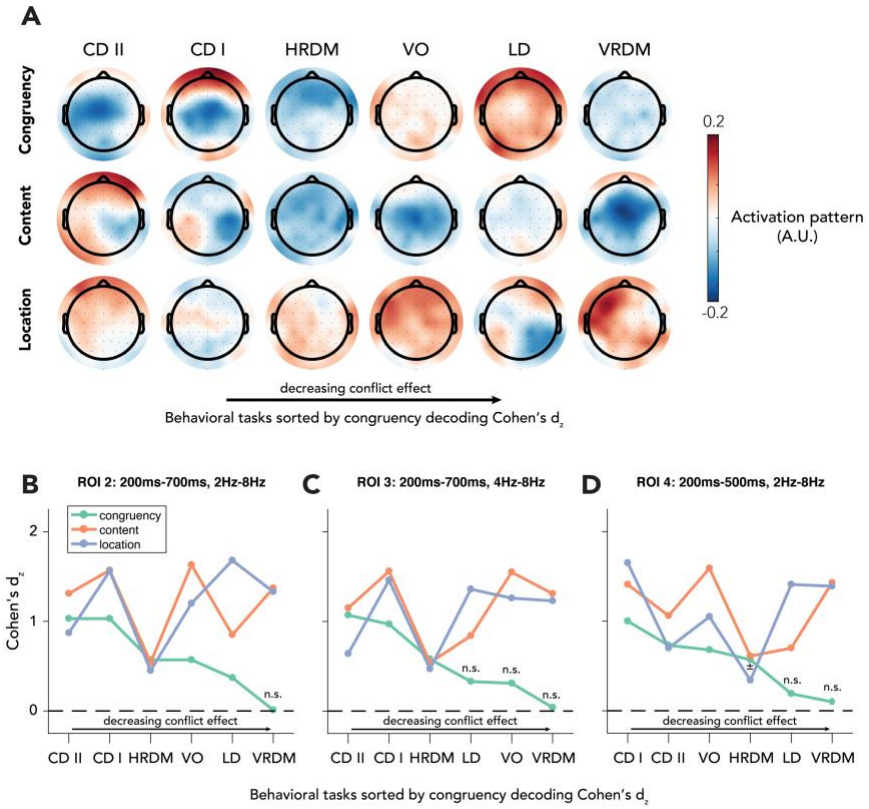

Figure 5 – Supplement 1. Topographic maps of reconstructed activation patterns and effects sizes for alternative ROIs. **(A)** . Decoding weights were extracted from the ROI that was used in the final analysis (100ms-700ms, 2Hz-8Hz). Then these weights were transformed into activation patterns by multiplying them with the covariance in the EEG data. The topographical maps of the two behavioral tasks in which congruency decoding was most accurate (content discrimination task I and II) reveal a clear midfrontal distribution, which is commonly found in the literature (Cohen & Cavanagh, 2011; Cohen & van Gaal, 2014; Jiang, Zhang, & van Gaal, 2015; Nigbur et al., 2012). **(B/C/D)** The reported results from our ROI-analyses (**Figure 5B**) could have been accidental, due to the fact that the ROI was chosen on the basis of previous results. In order to exclude this possibility, we performed the exact same analyses (rm-ANCOVA with factors being task and feature) on data that were extracted from three different ROIs. Crucially, we found interaction effects between behavioral task and stimulus feature for all ROIs (ROI 2: $F(10,402) = 20.31$, $p < .001$, $\eta_p^2 = 0.34$; ROI 3: $F(10,402) = 12.25$, $p < .001$, $\eta_p^2 = 0.23$; ROI 4: $F(10,402) = 14.47$, $p < .001$, $\eta_p^2 = 0.27$). Behavioral tasks on the x-axis are sorted by magnitude of the effect size of congruency decoding, similar to **Figure 5B**. Note that in all figures, the VRDM task is the one with the lowest effect size in congruency. Interestingly, for all ROIs the same pattern as in **figure 5B** is visible, namely that congruency decoding deteriorates to a point where accuracies are no longer significant, whereas for decoding of sensory stimulus features this is not the case. *Abbreviations*: **CD II**: content discrimination task II, **CD I**: content discrimination task I, **HRDM**: horizontal RDM task, **VO**: volume oddball detection task, **LD**: location discrimination task, **VRDM**: vertical RDM task. **±**: $p < .06$, **n.s.**: $p > .06$.

Discussion

Although it has been hypothesized for a long time that only basic physical properties of task-irrelevant sensory input are processed (Treisman & Gelade, 1980), over the past few years an overabundance of processes have been found to be preserved in the absence of attention (Fahrenfort et al., 2017; Li et al., 2002; Peelen et al., 2009). Here we aimed to push the limits of the brains capacity to process unattended information and addressed whether cognitive control networks can be recruited when conflicting features of sensory input are task-irrelevant. Interestingly, similar cognitive control functions have been shown to occur when stimuli are masked and hence conscious awareness is strongly reduced (Atas et al., 2016; D'Ostilio & Garraux, 2012b; Huber-Huber & Ansorge, 2018; Jiang, Zhang, & van Gaal, 2015; Jiang et al., 2018; van Gaal et al., 2008, 2010).

In two omnibus experiments with 6 different tasks, we presented stimuli with potentially auditory-spatial conflicting stimulus features (e.g. the word "left" presented on the right side) to participants, whilst they were performing several behavioral tasks. These tasks manipulated whether the features sound content and location of the auditory stimulus were task-relevant and whether these features were mapped to specific overlapping responses of the primary task. We observed clear signals of conflict processing in behavior (i.e. longer reactions times, increased error rates, increased sensitivity) and brain activity (i.e. above chance decoding accuracy in the theta-band) when the conflicting features of the auditory stimulus were task-relevant, i.e. in the content and location discrimination tasks, when another non-conflicting feature of the auditory stimulus was task-relevant, but the conflicting features content and location were not (volume oddball task) and when the conflicting features were not task-relevant but when they overlapped with the response scheme of the task (horizontal RDM task). When the features of the auditory stimulus were task-irrelevant and orthogonal to the response scheme of the primary task, i.e. in the vertical RDM task, we did not observe any effects of conflict in behavior or neural measures. The absence of conflict effects was supported by Bayesian analyses, showing reliable evidence in favor of the null hypothesis. Strikingly, the individual stimulus features, i.e. stimulus location/content, were always processed, regardless of their task relevance and response relevance. Note that this dissociation -hampered conflict processing yet preserved perceptual processing- cannot be explained by a lack of statistical power, because decoding accuracy of stimulus location/content was comparable between behavioral tasks. These results highlight that relatively basic stimulus properties escape the attentional bottleneck, lending support to previous studies (e.g. Fahrenfort et al., 2017; Li et al., 2002; Peelen et al., 2009; Sand & Wiens, 2011; Treisman & Gelade, 1980), but furthermore showcase that an attentional bottleneck for detecting conflict (integration of stimulus features) exists upstream in the hierarchy of cognitive processing. Below we link the observed results to the existing literature.

Object-based attention as a prerequisite for feature integration (leading to conflict)

So why are auditory content and location not integrated to form conflict when all of the auditory features are task-irrelevant? It has been suggested that the medial frontal cortex is crucial for monitoring the presence of conflict, through the detection of coinciding inputs (Cohen, 2014). In our paradigm, it thus seems crucial that information related to the auditory features content and location reaches the medial frontal cortex to be able to detect conflict, although control networks can undergo reconfiguration under certain circumstances (Canales-Johnson et al., 2020). Previous studies have shown that task-irrelevant stimuli can still undergo elaborate processing (Li et al., 2002; Peelen et al., 2009). Our decoding results show that task-irrelevant features are indeed processed by the brain (**Figures 2B**, **4B** and **5**). Interestingly, conflict between two task-irrelevant features was detected when another feature of the conflicting stimulus was task-relevant (volume oddball task) or when the conflicting features had overlap with the overall response scheme (horizontal RDM task), but remained undetected when none of the auditory features was task-relevant and there was no overlap with the response scheme (vertical RDM task). We argue that this difference is due to the fact that in the volume oddball and horizontal RDM tasks, the task-irrelevant conflicting features were selected through object-based attention. Theories of object-based attention have suggested that when one stimulus feature of an object is task-relevant and selected, attention "spreads" to all other features of the attended stimulus, even when these features are task-irrelevant or part of a different stimulus or modality (Chen, 2012; Chen & Cave, 2006; O'Craven et al., 1999; Turatto et al., 2005; Wegener et al., 2014; Xu, 2010). In the volume oddball task, a non-conflicting feature of the auditory stimulus (volume) was task-relevant, but this allowed for the selection of the other task-irrelevant features through object-based attention. In the horizontal RDM task on the other hand, the conflicting features of the task-irrelevant auditory stimulus overlapped with the overall response scheme or task-set of the participant, namely discriminating rightwards versus leftwards moving dots. This may have led to the automatic classification of all sensory input according to this task-set (as either coding for "left" or "right"), even when that input was not relevant for the task at hand. Possibly, through this classification, attentional resources could be exploited for the processing of these task-irrelevant features. This is especially interesting, because in all conflict analyses, incongruency was defined as the mismatch between the two features of the auditory stimulus

(location and sound) and not between a visual feature (leftward moving dots) and one feature of the auditory stimulus (e.g. the word "right"). Note that we report additional behavioral results that show clear indications of conflict when the task-relevant feature of the visual stimulus interferes directly with a *single* task-irrelevant feature of the auditory task (e.g., auditory content-dot motion conflict).

When inspecting the T-F maps for the vertical RDM task, the relatively fleeting temporal characteristics of the processing of the task-irrelevant stimulus features (sound content and location) might suggest that the integration of these features may not be possible due to a lack of time, as proposed in the incremental grouping model of attention (Roelfsema, 2006; Roelfsema & Houtkamp, 2011). However, the time-window in which conflict was decodable when the auditory conflicting features were task-relevant coincides with the time-range in which these features could be decoded when the auditory conflicting features were task-irrelevant (**Figure 5A**). Therefore, it seems unlikely that the more temporally constrained processing of task-irrelevant stimulus features is the cause of hampered conflict detection. Besides time being a factor, the processing of task-irrelevant features in the vertical RDM task may have also been too constrained to (early) sensory cortices and therefore could not progress to integration networks, including the MFC, necessary for the detection of conflict. Speculatively, the processing of task-irrelevant auditory features was relatively superficial, due to the relatively few remaining resources (Lavie et al., 2004; Sigman & Dehaene, 2006; Zylberberg et al., 2010, 2011), and combined with a lack of object-based attention, this may have prevented the propagation of the information to the MFC. It has been hypothesized that unattended (sometimes referred to as "preconscious"; Dehaene et al., 2006; Dehaene & Changeux, 2011) stimuli are not propagated deeply in the brain, but still allow for shallow recurrent interactions in sensory cortices. The poor spatial resolution of EEG-measurements and the specifics of our experimental setup however, do not allow to test these ideas regarding the involvement of spatially distinct cortices. Yet, previous work of our group suggests that task-irrelevant nonconscious information does not propagate to the frontal cortices, whereas task-relevant nonconscious information does. We demonstrated that masked task-irrelevant conflicting cues induced similar early processing in sensory cortices as compared to masked task-relevant cues, but prohibit activation of frontal cortices (van Gaal et al., 2008). These findings are not conclusive and so we believe that uncovering the role of task relevance in processing of (nonconscious) information deserves more attention in future work (see also van Gaal et al., 2012 for a discussion on this issue).

Sensory processing is weakened, but conflict processing hampered in absence of task relevance

We show that conflict processing is absent when conflicting features are fully task-irrelevant, while evidence of sensory processing is still present in neural data (**Figures 2B**, **4B** and **5**). Even though sensory processing of auditory features seems relatively preserved under various levels of task relevance of these features, it appears that sensory processing may in fact also be affected when the feature is task-irrelevant (**Figure 4 – Supplement 2**), in line with previous studies (e.g. Alilović et al., 2019; Jehee et al., 2011; Kok et al., 2012; Kouider et al., 2016), although to a lesser extent than conflict processing (**Figure 5B**). For example, when sound location is the task-relevant feature (i.e. in the location discrimination task), decoding accuracies for that feature are more broadband in the frequency domain (**Figure 4B**) and higher in the time-domain (**Figure 4 – Supplement 2**), compared to location decoding performance in other tasks. This increased decoding accuracy is present even before a response has been made, suggesting decreased early stage sensory processing in tasks where the decoded feature is not task-relevant. However, despite sensory processing being weakened under decreasing levels of task relevance, it is not diminished, in line with previous findings of ongoing processing in the (near) absence of attention (Fahrenfort et al., 2017; Li et al., 2002; Peelen et al., 2009). Processing of conflict between the two interfering auditory features, on the contrary, is hampered when the features are fully task-irrelevant. This is further supported by the significant interaction between task and feature in terms of decoding performance within the predefined ROI (**Figure 5B**). Summarizing, although processing of sensory features is degraded under decreasing levels of task relevance it is present regardless of attention, whereas detection of

conflict between these features is no longer possible when the features are fully task-irrelevant.

Besides object-based attention, the process through which attentional resources are allocated to the processing of task-irrelevant features of a task-relevant object, other mechanisms might also play a role in the extent to which sensory information is processed, such as the active suppression of task-irrelevant information. It has been shown that task-irrelevant information that is response-relevant, and can thus potentially interfere with performance on the primary task, can be suppressed to minimize interference (Appelbaum et al., 2011; Janssens et al., 2018; Polk et al., 2008; but see Egner & Hirsch, 2005). This would result in more reduced sensory processing, indexed by lower decoding performance, for task-irrelevant features that are response-relevant than for task-irrelevant features that are not. Disentangling the effects of such mechanisms, object-based attention and their possible interactions on the processing of sensory and cognitive information, however, falls outside the scope of this work.

Disentangling effects of conflict and task-difficulty

For our main analysis we trained a multivariate classifier on congruent versus incongruent trials and observed effects of task relevance of the performance of the classifier, i.e. decoding performance was hampered when conflicting features were fully task-irrelevant (**Figures 2B**, **4B** and **5B**). Moreover, we report behavioral effects of conflict in all *auditory* tasks as well (**Figures 2A** and **4A**). Given that behavioral performance on the auditory tasks is worse for incongruent trials as compared to congruent trials, one may wonder whether our multivariate decoder is in fact picking up information related to conflict detection or processes related to task-difficulty. Whether medial frontal theta-band oscillations are a reflection of conflict detection or task difficulty, and whether these factors can be dissociated in principle, has been the topic of debate in the literature (Grinband et al., 2011a, 2011b; McKay et al., 2017; Ruggeri et al., 2019; Yeung et al., 2011). On the one hand it has been shown that activity in the dorsal medial prefrontal cortex is related to RT, suggesting that neural markers of conflict may in fact reflect time on task (Grinband et al., 2011b, 2011b; Ruggeri et al., 2019). On the contrary, other research has shown that enhanced prefrontal theta-band oscillations are found in conflicting trials even when controlling for RT (Cohen & van Gaal, 2014), or task-difficulty (McKay et al., 2017). The decoding results presented in this work likely reflect conflict processing, and not just task-difficulty, for two reasons. First, the spatial distribution and time-frequency dynamics of the congruency decoding results are comparable to those more commonly found in the literature on conflict processing, even in a study where conflicting signals were matched for RT (Cohen & van Gaal, 2014). Specifically, using the content discrimination task of Experiment 1 as example, we observe effects of conflict centered on the theta-band and ~230ms-610ms post-conflict presentation, with a clear medial frontal spatial profile (**Figure 2B, Figure 5 – Supplement 1A**). Second, auditory stimulus conflict was decodable from neural data for two tasks in which there were either no effects of conflict -or task difficulty- on behavioral performance (i.e. horizontal RDM task), or even increased behavioral performance on conflicting trials (i.e. volume oddball task). Therefore, we believe that the observed congruency decoding results presented here are mainly driven by the detection of conflicting sensory inputs and are not, or much less so, driven by task difficulty.

Conflict between features of a task-irrelevant stimulus versus conflict between stimuli

Contrary to the current study, previous studies using a variety of conflict-inducing paradigms and attentional manipulations reported conflict effects in behavior and electrophysiological recordings induced by unattended stimuli or stimulus features (Mao & Wang, 2008; Padrão et al., 2015; Zimmer et al., 2010). However, our study deviates from those studies in several crucial aspects. First, we explicitly separate task-relevant stimulus features that cause conflict and task-relevant features that do not, parsing the cognitive components that induce cognitive control in this context. Furthermore, in the RDM and volume oddball tasks we tested whether conflict between two task-irrelevant features could be detected by the brain. Specifically, we investigated if conflict between two task-irrelevant features would be detected in the presence or absence of object-based attention

(e.g. volume oddball task vs. vertical RDM task), also manipulating whether task-irrelevant conflicting features mapped onto the response or not (horizontal RDM task vs. vertical RDM task). This approach is crucially different from previous studies that exclusively tested whether a task-irrelevant or unattended stimulus (feature) could interfere with processing of a task-relevant feature (Mao & Wang, 2008; Padrão et al., 2015; Zimmer et al., 2010). Under such conditions, at least one source contributing to the generation of conflict (i.e. the task-relevant stimulus) is fully attended and therefore one cannot claim that under those circumstance conflict detection occurs outside the scope of attention.

It can be argued that in our horizontal RDM task, the task-irrelevant auditory features (location and content) that mapped onto the response of the primary task could interfere with the processing of horizontal dot-motion, i.e. the task-relevant feature. This is in fact true, as we found effects of auditory content-dot motion and auditory location-dot motion conflict in behavior (both on RTs and ERs). This highlights that a single feature of a task-irrelevant stimulus can interfere with the response to a task-relevant stimulus when there are overlapping feature-response mappings. This is different from two features of a task-irrelevant stimulus to produce inherent conflict (e.g. between auditory content and location), which is what we specifically investigated by always testing the presence of auditory content-location conflict only. A similar argument might be made for our vertical RDM and volume oddball tasks, because in those cases the auditorily presented stimuli could potentially conflict with responses that were exclusively made with the right hand, e.g. the spoken word “left” or the sound from left location may conflict generally more with a right hand response (independent of the up/down classification or oddball detection) than the spoken word “right” or the sound from right location. In the vertical RDM task, the auditorily presented stimuli were truly task-irrelevant, as both stimulus content and location in isolation did not affect behavior. In the volume oddball task sound content and location were task-irrelevant features, but these features were part of the attended stimulus and hence selected through object-based attention. In this task, the content of the auditory stimuli (e.g. “left”) did interfere with right-hand responses to the volume oddball task, resulting in longer RTs (compared to “right”). Moreover, in this task we did find behavioral and neural effects of conflict between two auditory features (**Figure 4**). The absence of conflict effects in the vertical RDM and presence of such effects in the volume oddball task and horizontal RDM indicates that at least one feature of the stimulus containing the conflicting features should be task-relevant or associated with a response in order for conflict to be detected. Summarizing, we show that the brain is not able to detect conflict that emerges between two features of a task-irrelevant stimulus in the absence of object-based attention.

Lastly, in other studies, conflicting stimuli were often task-irrelevant on one trial (e.g. because they were presented at an unattended location) but task-relevant on the next (e.g. because they were presented at the attended location) (e.g. Padrão et al., 2015; Zimmer et al., 2010). Such trial-by-trial fluctuations of task relevance allow for across-trial modulations to confound any current trial effects (e.g. conflict-adaptation effect) and also induce a “stand-by attentional mode” where participants never truly disengage to be able to determine if a stimulus is task-relevant. We prevented such confounding effects in the present study, where the (potentially) conflicting features or the auditory stimulus were task-irrelevant on every single trial in the vertical RDM, horizontal RDM and volume oddball task.

Differences between response conflict and perceptual conflict cannot account for absence of conflict detection in task-irrelevant sensory input

One difference between the content and location discrimination tasks on the one hand and the volume oddball and RDM tasks on the other, was the task relevance of the (conflicting) auditory features. Another major difference between these groups of tasks was, consequently, the origin of the conflict. When the auditory stimuli were task-relevant, the origin of conflict was found in the interference of a task-irrelevant feature on behavioral performance, whereas for the other tasks this was not the case. We argued that in the volume oddball and RDM tasks, salient auditory stimuli could be *intrinsically* conflicting. Intrinsic conflict is often referred to as perceptual

conflict, as opposed to the aforementioned behavioral conflict (Kornblum, 1994). Although perceptual conflict effects are usually weaker than response conflict effects, both in behavior and electrophysiology (Frühholz et al., 2011; van Veen et al., 2001; Wang et al., 2014), this difference in the origin of the conflict is unlikely to explain why we did not observe effects of conflict under task-irrelevant sensory input, as opposed to earlier studies.

First, several neurophysiological studies have previously reported electrophysiological modulations by perceptual conflict centered on the MFC (Jiang, Zhang, & Van Gaal, 2015; Nigbur et al., 2012; van Veen et al., 2001; Wang et al., 2014; Zhao et al., 2015). Second, an earlier study using a task similar to ours (but including only task-relevant stimuli) showed effects of perceptual conflict, i.e. unrelated to response-mapping, in both behavior and neural measures (Buzzell et al., 2013). Third, the prefrontal monitoring system has previously been observed to respond when participants view other people making errors (Jääskeläinen et al., 2016; Van Schie et al., 2004), suggesting that cognitive control can be triggered without the need to make a response. Fourth, in our volume oddball task, where conflict was perceptual in nature as well, we did observe conflict effects, both in behavior and neural data.

Inattentional deafness or genuine processing of stimulus features?

The lack of conflict effects in the vertical RDM task might suggest a case of inattentional deafness, a phenomenon known to be induced by demanding visual tasks, which manifests itself in weakened early (~100ms) auditory evoked responses (Molloy et al., 2015). Interestingly, human speech seems to escape such load modulations and is still processed when unattended and task-irrelevant (Olguin et al., 2018; Röer et al., 2017; Zäske et al., 2016), potentially because of its inherent relevance, similar to (other) evolutionary relevant stimuli such as faces (Finkbeiner & Palermo, 2017; Lavie et al., 2003). Indeed, the results of our multivariate analyses demonstrate that spoken words are processed (at least to some extent) when they are task-irrelevant, as stimulus content (the words “left” and “right”, middle row **Figure 2B**, **Figure 4B** and **Figure 5B**) and stimulus location (whether the word was presented on the left or the right side”, bottom row **Figure 2B**, **Figure 4B** and **Figure 5B**) could be decoded from time-frequency data for all behavioral tasks. For all tasks, classification of stimulus content was present in the theta-band (4Hz-8Hz) which is in line with a previously proposed theoretical role for theta-oscillations in speech processing, namely that they track the acoustic-envelope of speech (Giraud & Poeppel, 2012). After this initial processing stage, further processing of stimulus content is reflected in more durable broadband activity for the content discrimination tasks, possibly related to higher-order processes (e.g. semantic) and response preparation (middle row and left column **Figure 2B**, **Figure 4B**). Similarly to processing of stimulus content, processing of stimulus location was most strongly reflected in the delta- to theta-range for all tasks (**Figure 4B**), which may relate to the auditory N1 ERP component, an ERP signal that is modulated by stimulus location (Fuentemilla et al., 2006; Lewald & Getzmann, 2011; Salminen et al., 2015). We also observed above-chance location decoding in the alpha-band for task-relevant auditory stimuli, convergent with the previously reported role of alpha-band oscillations in orienting and allocating audio-spatial attention (Weisz et al., 2014).

Thus, the characteristics of the early sensory processing of (task-irrelevant) auditory stimulus features in our study are in line with recent findings of auditory and speech processing. Moreover, our observations are in line with recent empirical findings that suggest a dominant role for late control operations, as opposed to early selection processes, in resolving conflict (Itthipuripat et al., 2019). Specifically, this work showed that in a Stroop-like paradigm both target and distractor information is analyzed fully, after which the conflicting input is resolved. Extending on this, we show preserved initial processing of task-irrelevant sensory input, but hampered late control operations necessary to detect conflict, at least in the current set-up.

Automatization of conflict processing does not promote detection of task-irrelevant conflict

Previous studies investigating conflict through masking procedures concluded that conflict detection by the MFC may happen automatically and is still operational under strongly reduced levels of stimulus visibility (D’Ostilio &

Garraux, 2012b; Jiang, Zhang, & van Gaal, 2015; van Gaal et al., 2008). Such automaticity can often be enhanced through training of the task. For example, training in a stop-signal paradigm in which stop-signals were rendered unconscious through masking, led to an increase in the strength of behavioral modulations of these stimuli (van Gaal et al., 2009). In order to see whether enhancing such automaticity could hypothetically increase the likelihood of conflict detection, we included extensive training sessions in the first experiment and we had measurements of the volume oddball task before and after exposure to conflicting tasks in the second experiment. In Experiment 1, we found no neural effects of conflict detection in the vertical RDM task, even when participants had been trained on the auditory task for 3600 trials (**Figure 2B**, **Figure 2 – Supplement 1D**). Training did result in a decrease of behavioral and neural conflict effects in content discrimination task I of Experiment 1, indicating that our training procedure was successful (**Figure 2 – Supplements 1A-C**) and suggesting more efficient functioning of conflict resolution mechanisms. In Experiment 2, participants performed the volume oddball task twice, once before and once after sound location and content had been mapped to responses. Again, we aimed to see if training on conflict tasks would enhance automaticity of conflict processing in a paradigm where the auditory conflicting features were task-irrelevant. We did not find any statistically reliable differences in behavioral conflict effects or accuracy of congruency decoding between the two runs (**Figure 4 - Supplement 1**). Therefore, it seems that the automaticity of conflict detection by the MFC and associated networks does not hold when the auditory stimulus is task-irrelevant (at least after the extent of training and exposure as studied here).

Increased detection performance on conflicting trials

Remarkably, we report increased behavioral performance on the volume oddball task for incongruent trials as compared to congruent trials (**Figure 4A**, **Figure 4 - Supplement 1B-D**). Speculatively, this increased behavioral performance (d') on incongruent trials may be due to attentional capture of conflicting stimuli. Attentional capture is the involuntary shift of attention towards salient stimuli (Awh et al., 2012; Theeuwes, 2010). Possibly, the detection of conflict between sound content and location is a salient event causing (re-)capture of attention towards the auditory stimulus, resulting in better processing of the stimulus information and ultimately better oddball detection performance. Thus, following the detection of conflict, frontal networks would have to exert control over attentional resources and direct them towards the source of the conflict. Interestingly, cases of frontal control over attentional processes have been demonstrated in the past, for example showing that task-irrelevant distractors that have been related to reward induce stronger attentional capture (Anderson et al., 2011) and that high working memory load increases the strength of attentional capture by distractors (Lavie & de Fockert, 2006). The present study was however not optimized to test directly which underlying mechanisms are associated with increased sensitivity of conflicting sensory input and this issue merits further experimentation.

Conclusion

Summarizing, high-level cognitive processes that require the integration of conflict inducing stimulus features are strongly hampered when none of the stimulus features of the conflict inducing stimulus are task-relevant and hence object-based attention is absent. This work nicely extends previous findings of perceptual processing outside the scope of attention (Peelen et al., 2009; Sand & Wiens, 2011; Schnuerch et al., 2016; Tusche et al., 2013), but suggest crucial limitations of the brain's capacity to process task-irrelevant "complex" cognitive control-initiating stimuli, indicative of an attentional bottleneck to detect conflict at high levels of information analysis. In contrast, the processing of more basic physical features of sensory input appears to be less deteriorated when input is task-irrelevant (Lachter et al., 2004).

Materials and methods

Participants

We performed two separate experiments, each containing multiple behavioral tasks. For each of these experiments we recruited 24 healthy human participants from the University of Amsterdam. None of the participants were took part in both experiments. So, in total 48 participants (37 female) aged 18 to 30, participated in this experiment for monetary compensation or participant credits. All participants had normal or corrected-to-normal vision and had no history of head injury or physical and mental illness. This study was approved by the local ethics committee of the University of Amsterdam and written informed consent was obtained from all participants after explanation of the experimental protocol. We will describe the experimental design and procedures for the two experiments separately. Data analyses and statistics were similar across experiments and will be discussed in the same section.

Experiment 1: design and procedures

Participants performed two tasks in which conflicting auditory stimuli were either task-relevant or task-irrelevant. In both tasks, conflict was elicited through a paradigm adapted from (Buzzell et al., 2013), in which spatial information and content of auditory stimuli could interfere. In content discrimination task I participants had to respond to the auditory stimuli, whereas in vertical RDM task participants had to perform a demanding RDM-task, while the auditory conflicting stimuli were still presented (**Figure 1A**). Participants performed both tasks on two experimental sessions of approximately 2.5 hours. In between these two experimental sessions, participants had two training sessions of one hour during which they only performed the task-relevant task (**Figure 1B**). On each experimental session, participants first performed a shortened version of the RDM-task to determine the appropriate coherency of the moving dots (73%-77% correct), followed by the actual task-irrelevant auditory task and finally the task-relevant auditory task. Participants were seated in a darkened, sound isolated room, 50cm from a 69x39cm screen (frequency: 120Hz, resolution: 1920 x 1080, RGB: 128, 128, 128). Both tasks were programmed in MATLAB (R2012b, The MathWorks, Inc.), using functionalities from Psychtoolbox (Kleiner et al., 2007).

Experiment 1: behavioral tasks

Auditory content discrimination task I

In the task-relevant auditory conflict task, the spoken words "links" (i.e. "left" in Dutch) and "rechts" (i.e. "right" in Dutch) were presented through speakers located on both sides of the participant (**Figure 1A**). Auditory stimuli were matched in duration and sample rate (44 kHz) and were recorded by the same male voice. By presenting these stimuli through either the left or the right speaker, content-location conflict arose on specific trials (e.g. the word "left" through the right speaker). Trials were classified accordingly, as either congruent (i.e. location and content are the same) or incongruent (i.e. location and content are different). Participants were instructed to respond as fast and accurate as possible, by pressing left ("a") or right ("l") on a keyboard located in front of the participants, according to the stimulus content, ignoring stimulus location. Responses had to be made with the left or right index finger, respectively. The task was divided in twelve blocks of 100 trials each, allowing participants to rest in between blocks. After stimulus presentation, participants had a 2s period in which they could respond. A variable inter-trial-interval between 850ms-1250ms was initiated directly after the response. If no response was made, the subsequent trial would start after the 2s response period. Congruent and incongruent trials occurred equally often (i.e. 50% of all trials), as expectancy of conflict has been shown to affect conflict processing (Soutschek et al., 2015). Due to an error in the script, there was a disbalance in the amount of trials coming from the left (70%) vs. right (30%) speaker location for the first 14 participants. However, the amount of congruent vs. incongruent and “left” vs. “right” trials were equally distributed. For the upcoming analyses, all trial classes were balanced in trial count.

Vertical random-dot motion task

In the task-irrelevant auditory task, participants performed an RDM-task in which they had to discriminate the motion (up versus down) of white dots (n=603) presented on a black circle (RGB: 0, 0, 0; ~14° visual angle; **Figure 1A**). Onset of the visual stimulus was paired with the presentation of the auditory conflicting stimulus. Participants were instructed to respond according to the direction of the dots, by pressing the "up" or "down" key on a keyboard with their right hand as fast and accurate as possible. Again, participants could respond in a 2s time-interval which was terminated after responses and followed by an inter-trial-interval of 850-1250ms. Task difficulty, in terms of dot-motion coherency (i.e. proportion of dots moving in the same direction) was titrated between blocks to 73-77% correct of all trials within that block. Similar to content discrimination task I, the vertical RDM was divided in twelve blocks containing 100 trials each, separated by short breaks. Again, congruent and incongruent trials, with respect to the auditory stimuli, occurred equally often.

Experiment 2: design and procedures

In the second experiment, we wanted to investigate whether it is task-irrelevance of the auditory stimulus itself or task-irrelevance of the auditory features (i.e. content and location) that determine whether prefrontal control processes are hampered. Participants performed two tasks in which auditory stimuli were fully task-relevant (location-discrimination and content-discrimination), one task in which the auditory stimulus was relevant but the features auditory location and content were not (volume oddball) and one task in which the auditory stimulus itself was task-irrelevant, but its features location and content could potentially interfere with behavior (horizontal RDM). Participants came to the lab for one session, lasting three hours. Each session started with the volume oddball task, followed by (in a counterbalanced order) the location discrimination, content discrimination and horizontal RDM tasks, and ended with a block of the volume oddball task again.

We included the location discrimination and content discrimination tasks both to replicate the results of Experiment 1 and also to see if there would be differences in these results between the two tasks. Specifically, we investigated whether processing of auditory stimulus features -as indicated by multivariate classification accuracies- would differ between the two tasks. Participants were seated in a darkened, sound isolated room, 50cm from a 69x39cm screen (frequency: 120Hz, resolution: 1920 x 1080, RGB: 128, 128, 128). All four experiments were programmed in Python 2.7 using functionalities from PsychoPy (Peirce et al., 2019).

Experiment 2: behavioral tasks

Auditory content discrimination task (II)

The auditory content discrimination task from Experiment 2 is a (near identical) replication of the auditory content discrimination task from Experiment 1. Participants were fixating on a fixation mark in the center of the screen. Again, the spoken words "links" (i.e. "left" in Dutch) and "rechts" (i.e. "right" in Dutch) were presented through speakers located on both sides of the participant. Participants were instructed to respond according to the stimulus content by pressing left (“A”) or right (“L”) on a keyboard located in front of them, with their left and right index fingers respectively. Concurrently, on every trial, a black disk with randomly moving dots (coherence: 0) was presented to keep sensory input similar between tasks. After stimulus presentation, participants had an 800ms period in which they could respond. After a response, the response window would be terminated directly. A variable inter-trial-interval between 250ms-450ms was initiated directly after the response. If no response was made, the subsequent trial would start after the ITI. All stimulus features (i.e. sound content, location and congruency) were presented in a balanced manner (e.g. 50% congruent, 50% incongruent). The task was divided in six blocks of 100 trials each, allowing participants to rest in between blocks.

Auditory location discrimination task

The auditory location discrimination task was identical to the auditory content discrimination task II, with the exception that participants were now instructed to respond according to the location of the auditory stimulus.

Thus, participants had to press a left button (“A”) for sounds coming from a left speaker and right button (“L”) for sounds coming from a right speaker. Again, participants performed six blocks of 100 trials.

Volume oddball task

In the volume oddball task, the same auditory stimuli were presented. Again, on every trial, a black disk with randomly moving dots (coherence: 0) was presented to keep sensory input similar between tasks. Occasionally an auditory stimulus would be presented at a lower volume. The initial volume of the oddballs was set to 70% , but was staircased in between blocks to yield 83%-87% correct answers. If participants performance on the previous block was below or above this range the volume increased or decreased with 5% respectively. The odds of a trial being an oddball trial were 1/8 (drawn from a uniform distribution). Participants were instructed to detect these oddballs, by pressing the space-bar as fast as possible whenever they thought they heard a volume oddball. If they thought that the stimulus was presented at a normal volume, they were instructed to refrain from responding. The response interval was 800ms, which was terminated at response. A variable inter-trial-interval 150ms-350ms was initiated after this response interval. Participants performed two runs of this task, at the beginning of each session and at the end of each session. Each run contained five blocks of 100 trials each.

Horizontal RDM task

In the horizontal RDM task participants had to discriminate the motion (left versus right) of white dots (n=603) presented on a black circle (RGB: 0, 0, 0; ~14° visual angle). Onset of the visual stimulus was paired with the presentation of the auditory stimulus. Participants were instructed to respond according to the direction of the dots, by pressing a left key ("A"), or right key ("L") on a keyboard with left and right index fingers respectively as fast and accurate as possible. Participants could respond in an 800ms time-interval which was terminated after responses and followed by an inter-trial-interval of 250-450ms. Task difficulty, in terms of dot-motion coherency (i.e. proportion of dots moving in the same direction) was set to 0.3 in the first block. This value indicated an intermediate coherence, as every trial in that block could be the intermediate coherence, but also half (0.15) or twice this intermediate coherence (0.6). The intermediate coherence was titrated between blocks to 73-77% correct. If behavioral performance fell outside that range, 0.01 was added to or subtracted from the intermediate coherence. The horizontal RDM consisted of 10 blocks of 60 trials. All stimulus features (i.e. sound content, location and congruency) were presented in a balanced manner (e.g. 50% congruent, 50% incongruent).

Data analysis

We were primarily interested in the effects of congruency of the auditory stimuli on both behavioral and neural data. Therefore, we defined trial-congruency on the basis of these auditory stimuli, in all behavioral tasks of the two experiments. All behavioral analyses were programmed in MATLAB (R2017b, The MathWorks, Inc.).

Statistical analysis of behavioral data

In all tasks, trials with an RT <100ms or >1500ms were excluded from behavioral analyses. Missed trials were excluded in all tasks (except the volume oddball task) as well. In order to investigate whether current trial conflict effects were present under varying levels of task relevance and to inspect if training on/exposure to conflict inducing tasks modulated such conflict effects, we performed rm-ANOVAs on different behavioral measures. For all tasks, excluding the volume oddball task, the rm-ANOVAs were performed on the ER over all trials and RTs on correct trials. For the volume oddball task, perceptual sensitivity (d’; Green & Swets, 1966) and RTs of correct trials (i.e. ‘hit’ trials) were analyzed with rm-ANOVAs. If the assumption of sphericity was violated, we applied a Greenhouse-Geisser correction. For the tasks from Experiment 1, we performed these ANOVAs with task relevance, training (before vs. after) and current trial congruency as factors (2x2x2 factorial design). Additional post-hoc rm-ANOVAs, for content discrimination task I and vertical RDM task separately (2x2 factorial design), were used to inspect the origin of significant factors that were modulated by task relevance.

For content discrimination task I, the location discrimination task and horizontal RDM task we performed a rm-ANOVA with task and congruency as factors (3x2 factorial design). For the volume oddball task, we performed a rm-ANOVA with congruency and run number as factors (2x2 factorial design) on RTs, d' scores, hit rates and false alarm rates. We also applied paired sample t-tests comparing the difference in these variables between incongruent and congruent for all tasks, within each experimental session (vertical RDM, content discrimination task I) and run (volume oddball task).

To test interference of the *individual* auditory features, sound content and sound location, on performance on the vertical RDM task and volume oddball task we performed rm-ANOVAs on RTs with sound location and sound content as factors (2x2 factorial design). Additionally, for the horizontal RDM, we tested for auditory-visual conflict effects (i.e. conflict between sound content/location and dot direction) in RT and ER with paired sample t-tests comparing incongruent and congruent trials.

In case of null findings, we performed a Bayesian analysis (rm-ANOVA or paired sample t-test) with identical parameters and settings on the same data, to test if there was actual support of the null hypothesis (JASP Team, 2018).

Analysis of EEG data

EEG data were analyzed using custom-made software written in MATLAB, with support from the toolboxes EEGLAB (Delorme & Makeig, 2004) and ADAM (Fahrenfort et al., 2018) .

Acquisition and preprocessing

EEG-data were recorded with a 64-channel BioSemi apparatus (BioSemi B.V., Amsterdam, The Netherlands), at 512Hz. Vertical eye-movements were recorded with electrodes located above and below the left eye, horizontal eye-movements were recorded with electrodes located at the outer canthi of the left and the right eye. All EEG traces were re-referenced to the average of two electrodes located on the left and right earlobes (mastoidal reference for 1 participant in Experiment 1). The data were band-pass filtered offline, with cut-off frequencies of 0.01Hz-50Hz. Next, epochs were created by taking data from -1s to 2s around onset of stimulus presentation. We then rejected epochs containing blink artefacts and high voltage artefacts. Blinks were defined as VEOG data exceeding a threshold of ±100mV in a time-window of 0ms-800ms post-stimulus. This procedure resulted in the removal of 5.03% (SD = 9.71%) of epochs in Experiment 1 and 6.10% (*SD* = 7.54%) of epochs in Experiment 2. Subsequently, high-voltage artefacts were defined as events where voltage exceeded a threshold of ±300mV in a time-window of 0ms-800ms post-stimulus on any EEG-channel. With this second round of artefact rejection 3.65% (*SD* = 7.96%) of all epochs were removed in Experiment 1 and 4.16% (*SD* = 8.86%) in Experiment 2. In total, 8.68% (*SD* = 12.09%) of all trials were removed in Experiment 1 and 10.26% (*SD* = 13.92%) were removed in Experiment 2. Note that this procedure ensures the absence of blinks and high amplitude artefacts within the predefined ROI time-window of 100ms-700ms. The data of one participant in Experiment 2 contained many artefacts. Specifically, across all 5 tasks performed in Experiment 2, 64.46% (*SD* = 9.88%) of all epochs were rejected for this participant. This was more than 3 standard deviations from the average number of rejected trials across participants and files and left too few trials for the decoding analysis. Therefore this participant was removed from all EEG analyses.

Time-frequency domain multivariate pattern analysis (decoding)

We applied a multivariate backwards decoding model to EEG data that was transformed to the time-frequency domain. We used multivariate analyses both because its higher sensitivity in comparison with univariate analyses and to inspect if and to what extent different stimulus features (i.e. location and content) were processed in both tasks, without having to preselect spatial or time-frequency ROIs. The ADAM-toolbox was used for time-frequency decomposition and decoding (Fahrenfort et al., 2018). Single trial power spectra were

computed, by convolving the EEG data with a complex wavelet (wavelet size of 0.5s) after the application of a Hann taper (epochs: -100ms to 1000ms, 2Hz-30Hz in linear steps of 2Hz). Raw time-frequency data contained both induced and evoked power. Trials were classified according to current trial stimulus features (i.e. location and content) resulting in 4 trial types. As decoding algorithms are known to be time-consuming, epochs were resampled to 64Hz. Then, we applied a decoding algorithm to the data according to a 20-fold cross-validation scheme, using either stimulus location, stimulus content or congruency as stimulus class. Specifically, a linear discriminant analysis (LDA) was trained to discriminate between stimulus classes (e.g. left versus right speaker location etc.). Classification accuracy was computed as the area under the curve (AUC), a measure derived from Signal Detection Theory (Green & Swets, 1966).

The multivariate classifiers were on different subsets of trials, depending on the behavioral task. For the auditory tasks (content discrimination I and II, location discrimination and volume oddball detection), only correct trials were included in the analysis as errors tend to elicit a similar, albeit not identical, neural response as cognitive conflict and errors are more likely on incongruent trials (Cohen & van Gaal, 2014). For the volume oddball detection task we additionally excluded all oddball trials, thus only testing correct rejections, in order to prevent conflict arising between responses made exclusively with the right hand and sound content and location. For the two visual tasks (horizontal and vertical RDM) we trained the classifier on all trials.

Topographical maps of ROI decoding

Topographical maps were created in order to investigate the spatial sources of activity related to the processing of the auditory features (content, location and congruency). We first extracted classifier weights for each task and feature from the predefined ROI (100ms-700ms, 2Hz-8Hz) allowing us to directly compare the spatial distributions between features and tasks. However, raw classifier weights are not interpretable as neural sources of activity and therefore have to be reconstructed (Haufe et al., 2014). Thus, classifier weights were transformed to activation patterns by multiplying them with the covariance in the EEG data. The topographical activity maps of tasks and features with low decoding performance should be interpreted with caution, as activation patterns reconstructed from classifier weights may be unreliable when decoding performance is low (Haufe et al., 2014).

Time domain multivariate pattern analysis (decoding)

We applied a time-domain decoding analysis on EEG data, to inspect the possible effect of task relevance of a stimulus feature on sensory processing. For this analysis only EEG data from the tasks of Experiment 2 were used, as its parameters were comparable between tasks in this experiment (e.g. always visual stimulus present, no extensive training). For the analysis, we trained linear classifiers (LDA) to discriminate sound content ("left" versus "right") or sound location (left speaker versus right speaker). First, epochs (from -100ms to 1000ms around stimulus presentation) were resampled to 64Hz similar to the time-frequency decoding analyses. Then, the models were trained and tested according to a 20-fold cross-validation scheme. The AUC scores we obtained via multivariate analyses of our EEG data were tested per time-point with one-sided t-tests ($\bar{x} > 0.5$) across participants against chance-level (50%). These t-test were corrected for multiple comparisons over time, using cluster-based permutation tests ($p < .05$, 1000 iterations). For each decoded stimulus feature, we then compared the decoding accuracies of the behavioral task in which the feature was task-relevant, to all other tasks in a pairwise fashion (e.g. location decoding under location discrimination task versus horizontal RDM task), with cluster-corrected two-sided t-tests against 0.

Statistical analysis of EEG data

The AUC scores we obtained via multivariate analyses of our EEG data were tested per time-point and frequency with one-sided t-tests ($\bar{x} > 0.5$) across participants against chance-level (50%). These t-test were corrected for multiple comparisons over time and frequency, using cluster-based permutation tests ($p < .05$, 1000 iterations).

This procedure yields time-frequency clusters of significant above-chance classifier accuracy, indicative of information processing. Note that this procedure yields results that should be interpreted as fixed effects (Allefeld et al., 2016), but is nonetheless standard in the scientific community.

In addition to the cluster analysis, we performed hypothesis-driven analyses on classifier accuracies that were extracted from a pre-defined time-frequency ROI. All these analyses were performed in JASP (JASP Team, 2018). We then applied an ANCOVA on these accuracies with fixed effects being task and stimulus feature. Next, one-sample t-tests (one-sided, $\bar{x} > 0.5$) were performed on every task/feature combination to determine whether decoding accuracy of a specific feature within our pre-selected ROI was above chance during the various behavioral tasks. Additional Bayesian one-sample t-tests (one-sided, $\bar{x} > 0.5$, Cauchy scale = 0.71) were performed to inspect evidence in favor of the null-hypothesis that decoding accuracy was not above chance.

We performed the same analysis on different ROIs. The results of those analyses can be found in **Figure 5 – Supplement 1B-D**.

Supplementary Figures

Figure 2 - Supplement 1
Figure 4 - Supplement 1
Figure 4 - Supplement 2
Figure 5 - Supplement 1

Source Files

Figure 2

- Figure 2 source data 2.rtf
  Table with variable identifiers used in Figure 2 source data 1.
- Figure 2 source data 1.csv
  This csv table contains the raw data for Figure 2 panel A.

Figure 2 – Supplement 1

- Figure 2 – Supplement 1 source data 3.rtf
  Table with variable identifiers used in Figure 2 – Supplement 1 source data 1 and 2.
- Figure 2– Supplement 1 source data 1.csv
  This csv table contains the raw data results for Figure 2 – Supplement 1 panels A and B.
- Figure 2A source data 2.csv
  This csv table contains the statistical results for Figure 2 – Supplement 1 panels A and B.

Figure 4

- Figure 4 source data 2.rtf
  Table with variable identifiers used in Figure 4 source data 1.
- Figure 4 source data 1.csv
  This csv table contains the data for Figure 4 panel A.

Figure 4 – Supplement 1

- Figure 4 – Supplement 1 source data 3.rtf
  Table with variable identifiers used in Figure 4 – Supplement 1 source data 1 and 2.
- Figure 4 – Supplement 1 source data 1.csv
  This csv table contains the raw data for Figure 4 – Supplement 1 panels A, B, C and D.
- Figure 4 – Supplement 1 source data 2.csv
  This csv table contains the statistical results for Figure 4 – Supplement 1 panels A, B, C and D.

Figure 5

- Figure 5 source data 3.rtf

Table with variable identifiers used in Figure 5 source data 1 and 2.

- Figure 5 source data 1.csv

  This csv table contains the raw data that was statistically analyzed for Figure 5 panel B.

- Figure 5 source data 2.csv

  This csv table contains the data for Figure 5 panel B.

Literature

Alilović, J., Timmermans, B., Reteig, L. C., van Gaal, S., & Slagter, H. A. (2019). No Evidence that Predictions and Attention Modulate the First Feedforward Sweep of Cortical Information Processing. *Cerebral Cortex*, *29*(5), 2261–2278. https://doi.org/10.1093/cercor/bhz038

Allefeld, C., Görgen, K., & Haynes, J. D. (2016). Valid population inference for information-based imaging: From the second-level t-test to prevalence inference. *NeuroImage*, *141*, 378–392. https://doi.org/10.1016/j.neuroimage.2016.07.040

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, *108*(25), 10367–10371. https://doi.org/10.1073/pnas.1104047108

Appelbaum, L. G., Smith, D. V., Boehler, C. N., Chen, W. D., & Woldorff, M. G. (2011). Rapid Modulation of Sensory Processing Induced by Stimulus Conflict. *Journal of Cognitive Neuroscience*, *23*(9), 2620–2628. https://doi.org/10.1162/jocn.2010.21575

Atas, A., Desender, K., Gevers, W., & Cleeremans, A. (2016). Dissociating perception from action during conscious and unconscious conflict adaptation. *Journal of Experimental Psychology: Learning Memory and Cognition*. https://doi.org/10.1037/xlm0000206

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443. https://doi.org/10.1016/j.tics.2012.06.010

Broadbent, D. E. (1958). *Perception and Communication*. Oxford University Press.

Buzzell, G. A., Roberts, D. M., Baldwin, C. L., & McDonald, C. G. (2013). An electrophysiological correlate of conflict processing in an auditory spatial Stroop task: The effect of individual differences in navigational style. *International Journal of Psychophysiology*, *90*(2), 265–271. https://doi.org/10.1016/j.ijpsycho.2013.08.008

Canales-Johnson, A., Beerendonk, L., Blain, S., Kitaoka, S., Ezquerro-Nassar, A., Nuiten, S., Fahrenfort, J., Gaal, S. van, & Bekinschtein, T. A. (2020). Decreased Alertness Reconfigures Cognitive Control Networks. *Journal of Neuroscience*, *40*(37), 7142–7154. https://doi.org/10.1523/JNEUROSCI.0343-20.2020

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Chen, A., Tang, D., & Chen, X. (2013). Training Reveals the Sources of Stroop and Flanker Interference Effects. *PLoS ONE*, *8*(10), e76580. https://doi.org/10.1371/journal.pone.0076580

Chen, Z. (2012). Object-based attention: A tutorial review. *Attention, Perception, & Psychophysics*, *74*(5), 784–802. https://doi.org/10.3758/s13414-012-0322-z

Chen, Z., & Cave, K. R. (2006). When does visual attention select all features of a distractor? *Journal of Experimental Psychology: Human Perception and Performance*, *32*(6), 1452–1464. https://doi.org/10.1037/0096-1523.32.6.1452

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with 2 ears. *Journal of the Acoustical Society of America*. https://doi.org/10.1121/1.1907229

Cohen, M. X. (2014). A neural microcircuit for cognitive conflict detection and signaling. *Trends in Neurosciences*, *37*(9), 480–490. https://doi.org/10.1016/j.tins.2014.06.004

Cohen, M. X., & Cavanagh, J. F. (2011). Single-Trial Regression Elucidates the Role of Prefrontal Theta Oscillations in Response Conflict. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00030

Cohen, M. X., & van Gaal, S. (2014). Subthreshold muscle twitches dissociate oscillatory neural signatures of conflicts from errors. *NeuroImage*, *86*, 503–513. https://doi.org/10.1016/j.neuroimage.2013.10.033

Cosman, J. D., & Vecera, S. P. (2012). Object-based attention overrides perceptual load to modulate visual distraction. *Journal of Experimental Psychology. Human Perception and Performance*, *38*(3), 576–579. https://doi.org/10.1037/a0027406

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204–211. https://doi.org/10.1016/j.tics.2006.03.007

Dehaene, S., & Changeux, J.-P. (2011). Experimental and Theoretical Approaches to Conscious Processing. *Neuron*, *70*(2), 200–227. https://doi.org/10.1016/j.neuron.2011.03.018

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, *70*. https://doi.org/10.1037/h0042712

D'Ostilio, K., & Garraux, G. (2012a). Brain mechanisms underlying automatic and unconscious control of motor action. *Frontiers in Human Neuroscience*, *6*(September), 1–5. https://doi.org/10.3389/fnhum.2012.00265

D'Ostilio, K., & Garraux, G. (2012b). Dissociation between unconscious motor response facilitation and conflict in medial frontal areas. *European Journal of Neuroscience*, *35*(2), 332–340. https://doi.org/10.1111/j.1460-9568.2011.07941.x

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, *8*(12), 1784–1790. https://doi.org/10.1038/nn1594

Fahrenfort, J. J., Van Driel, J., van Gaal, S., & Olivers, C. N. L. (2018). From ERPs to MVPA using the Amsterdam Decoding and Modeling toolbox (ADAM). *Frontiers in Neuroscience - Brain Imaging Methods*, *12*(July). https://doi.org/10.3389/fnins.2018.00368

Fahrenfort, J. J., van Leeuwen, J., Olivers, C. N. L., & Hogendoorn, H. (2017). Perceptual integration without conscious access. *Proceedings of the National Academy of Sciences*, *114*(14), 3744–3749. https://doi.org/10.1073/pnas.1617268114

Finkbeiner, M., & Palermo, R. (2017). The Role of Spatial Attention in Nonconscious Processing A Comparison of Face and Nonface Stimuli. *Psychological Science*, *20*(1), 42–51.

Frühholz, S., Godde, B., Finke, M., & Herrmann, M. (2011). Spatio-temporal brain dynamics in a combined stimulus-stimulus and stimulus-response conflict task. *NeuroImage*, *54*(1), 622–634. https://doi.org/10.1016/j.neuroimage.2010.07.071

Fuentemilla, L., Marco-Pallarés, J., & Grau, C. (2006). Modulation of spectral power and of phase resetting of EEG contributes differentially to the generation of auditory event-related potentials. *NeuroImage*, *30*(3), 909–916. https://doi.org/10.1016/j.neuroimage.2005.10.036

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517. https://doi.org/10.1038/nn.3063

Goldman-Rakic. (1995). Architecture of the prefrontal cortex and the central executive. In *Structure and functions of the human prefrontal cortex* (pp. 71–83). New York Academy of Sciences.

Goldman-Rakic. (1996). The prefrontal landscape: Implications of functional architecture for understanding human mentation and the central executive. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1346), 1445–1453. https://doi.org/10.1098/rstb.1996.0129

Green, D. M., & Swets, J. A. (1966). Signal detection theory and psychophysics. *Society*. https://doi.org/10.1901/jeab.1969.12-475

Grinband, J., Savitskaya, J., Wager, T. D., Teichert, T., Ferrera, V. P., & Hirsch, J. (2011a). Conflict, error likelihood, and RT: Response to Brown & Yeung et al. *NeuroImage*, *57*(2), 320–322. https://doi.org/10.1016/j.neuroimage.2011.04.027

Grinband, J., Savitskaya, J., Wager, T. D., Teichert, T., Ferrera, V. P., & Hirsch, J. (2011b). The Dorsal Medial Frontal Cortex is Sensitive to Time on Task, Not Response Conflict or Error Likelihood. *NeuroImage*, *57*(2), 303–311. https://doi.org/10.1016/j.neuroimage.2010.12.027

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding Dynamic Brain Patterns from Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data. *Journal of Cognitive Neuroscience*, *29*(4), 677–697. https://doi.org/10.1162/jocn_a_01068

Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., & Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage*, *87*, 96–110. https://doi.org/10.1016/j.neuroimage.2013.10.067

Haxby, J., Gobbini, M., Furey, M., Ishai, A., Schouten, J., & Pietrini, P. (2001). Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. *Science (New York, N.Y.)*, *293*, 2425–2430. https://doi.org/10.1126/science.1063736

Hebart, M. N., & Baker, C. I. (2017). Deconstructing multivariate decoding for the study of brain function. *NeuroImage*, *August*, 1–15. https://doi.org/10.1016/j.neuroimage.2017.08.005

Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychologica*, *136*(2), 189–202. https://doi.org/10.1016/j.actpsy.2010.04.011

Huber-Huber, C., & Ansorge, U. (2018). Unconscious conflict adaptation without feature-repetitions and response time carry-over. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(2), 169–175. https://doi.org/10.1037/xhp0000450

Itthipuripat, S., Deering, S., & Serences, J. T. (2019). When Conflict Cannot be Avoided: Relative Contributions of Early Selection and Frontal Executive Control in Mitigating Stroop Conflict. *Cerebral Cortex*, 1–12. https://doi.org/10.1093/cercor/bhz042

Jääskeläinen, I. P., Tapani, K., Lahnakoski, J. M., Manoach, D. S., Sams, M., Ahveninen, J., Halme, H.-L., Agam, Y., & Glerean, E. (2016). Neural mechanisms supporting evaluation of others' errors in real-life like conditions. *Scientific Reports*, *6*(1), 1–10. https://doi.org/10.1038/srep18714

Janssens, C., De Loof, E., Boehler, C. N., Pourtois, G., & Verguts, T. (2018). Occipital alpha power reveals fast attentional inhibition of incongruent distractors. *Psychophysiology*, *55*(3). https://doi.org/10.1111/psyp.13011

JASP Team. (2018). JASP (Version 0.8.6.0). In *[Computer software].*

Jehee, J. F. M., Brady, D. K., & Tong, F. (2011). Attention Improves Encoding of Task-Relevant Features in the Human Visual Cortex. *Journal of Neuroscience*, *31*(22), 8210–8219. https://doi.org/10.1523/JNEUROSCI.6153-09.2011

Jiang, J., Correa, C. M., Geerts, J., & van Gaal, S. (2018). The relationship between conflict awareness and behavioral and oscillatory signatures of immediate and delayed cognitive control. *NeuroImage*, *177*(May), 11–19. https://doi.org/10.1016/j.neuroimage.2018.05.007

Jiang, J., Zhang, Q., & van Gaal, S. (2015). Conflict awareness dissociates theta-band neural dynamics of the medial frontal and lateral frontal cortex during trial-by-trial cognitive control. *NeuroImage*, *116*, 102–111. https://doi.org/10.1016/j.neuroimage.2015.04.062

Jiang, J., Zhang, Q., & Van Gaal, S. (2015). EEG neural oscillatory dynamics reveal semantic and response conflict at difference levels of conflict awareness. *Scientific Reports*, *5*(July), 1–11. https://doi.org/10.1038/srep12008

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219. https://doi.org/10.1016/0010-0285(92)90007-O

Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). What's new in Psychtoobox-3? *Perception*.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: Two distinct brain processes. *Trends in Cognitive Sciences*, *11*(1), 16–22. https://doi.org/10.1016/j.tics.2006.10.012

Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychologica*, *134*(3), 372–384. https://doi.org/10.1016/j.actpsy.2010.03.010

Kok, P., Rahnev, D., Jehee, J. F. M., Lau, H. C., & de Lange, F. P. (2012). Attention Reverses the Effect of Prediction in Silencing Sensory Signals. *Cerebral Cortex*, *22*(9), 2197–2206. https://doi.org/10.1093/cercor/bhr310

Kornblum, S. (1994). The way irrelevant dimensions are processed depends on what they overlap with: The case of Stroop- and Simon-like stimuli. *Psychological Research*, *56*(3), 130–135. https://doi.org/10.1007/BF00419699

Kouider, S., Barbot, A., Madsen, K. H., Lehericy, S., & Summerfield, C. (2016). Task relevance differentially shapes ventral visual stream sensitivity to visible and invisible faces. *Neuroscience of Consciousness*, *2016*(niw021). https://doi.org/10.1093/nc/niw021

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after broadbent (1958): Still no identification without attention. In *Psychological Review*. https://doi.org/10.1037/0033-295X.111.4.880

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, *7*(1), 12–18. https://doi.org/10.1016/S1364-6613(02)00013-X

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, 1–9. https://doi.org/10.10162236

Lavie, N., & de Fockert, J. (2006). Frontal control of attentional capture in visual search. *Visual Cognition*, *14*(4–8), 863–876. https://doi.org/10.1080/13506280500195953

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load Theory of Selective Attention and Cognitive Control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Lavie, N., Ro, T., & Russell, C. (2003). The role of perceptual load in processing distractor faces. *Psychological Science*, *14*(5), 510–515. https://doi.org/10.1111/1467-9280.03453

Lewald, J., & Getzmann, S. (2011). When and where of auditory spatial processing in cortex: A novel approach using electrotomography. *PLoS ONE*, *6*(9). https://doi.org/10.1371/journal.pone.0025146

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences*, *99*(14), 9596–9601. https://doi.org/10.1073/pnas.092277599

MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(1), 126–135. https://doi.org/10.1037/0278-7393.14.1.126

Mao, W., & Wang, Y. (2008). The active inhibition for the processing of visual irrelevant conflict information. *International Journal of Psychophysiology*, *67*(1), 47–53. https://doi.org/10.1016/j.ijpsycho.2007.10.003

McKay, C. C., van den Berg, B., & Woldorff, M. G. (2017). Neural cascade of conflict processing: Not just time-on-task. *Neuropsychologia*, *96*, 184–191. https://doi.org/10.1016/j.neuropsychologia.2016.12.022

Molloy, K., Griffiths, T. D., Chait, M., & Lavie, N. (2015). Inattentional Deafness: Visual Load Leads to Time-Specific Suppression of Auditory Evoked Responses. *Journal of Neuroscience*, *35*(49), 16046–16054. https://doi.org/10.1523/JNEUROSCI.2931-15.2015

Moray, N. (1959). Attention in Dichotic Listening: Affective Cues and the Influence of Instructions. *Quarterly Journal of Experimental Psychology*, *11*(1). https://doi.org/10.1080/17470215908416289

Nigbur, R., Cohen, M. X., Ridderinkhof, K. R., & Stürmer, B. (2012). Theta Dynamics Reveal Domain-specific Control over Stimulus and Response Conflict. *Journal of Cognitive Neuroscience*, *24*(5), 1264–1274. https://doi.org/10.1162/jocn_a_00128

O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). FMRI evidence for objects as the units of attentional selection. *Nature*, *401*(6753), 584–587. https://doi.org/10.1038/44134

Olguin, A., Bekinschtein, T. A., & Bozic, M. (2018). Neural encoding of attended continuous speech under different types of interference. *Journal of Cognitive Neuroscience*, *30*(11), 1606–1619. https://doi.org/10.1162/jocn_a_01303

Padrão, G., Rodriguez-Herreros, B., Pérez Zapata, L., & Rodriguez-Fornells, A. (2015). Exogenous capture of medial-frontal oscillatory mechanisms by unattended conflicting information. *Neuropsychologia*, *75*, 458–468. https://doi.org/10.1016/j.neuropsychologia.2015.07.004

Peelen, M. V., Fei-Fei, L., & Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature*, *460*(7251), 94–97. https://doi.org/10.1038/nature08103

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Polk, T. A., Drake, R. M., Jonides, J. J., Smith, M. R., & Smith, E. E. (2008). Attention Enhances the Neural Processing of Relevant Features and Suppresses the Processing of Irrelevant Features in Humans: A Functional Magnetic Resonance Imaging Study of the Stroop Task. *The Journal of Neuroscience*, *28*(51), 13786–13792. https://doi.org/10.1523/JNEUROSCI.1026-08.2008

Rahnev, D. A., Huang, E., & Lau, H. C. (2012). Subliminal stimuli in the near absence of attention influence top-down cognitive control. *Attention, Perception, and Psychophysics*, *74*(3), 521–532. https://doi.org/10.3758/s13414-011-0246-z

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*(5695), 443–447. https://doi.org/10.1126/science.1100301

Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience*, *29*(1), 203–227. https://doi.org/10.1146/annurev.neuro.29.051605.112939

Roelfsema, P. R., & Houtkamp, R. (2011). Incremental grouping of image elements in vision. *Attention, Perception & Psychophysics*, *73*(8), 2542–2572. https://doi.org/10.3758/s13414-011-0200-0

Röer, J. P., Körner, U., Buchner, A., & Bell, R. (2017). Semantic priming by irrelevant speech. *Psychonomic Bulletin and Review*, *24*(4), 1205–1210. https://doi.org/10.3758/s13423-016-1186-3

Rousselet, G. A., Thorpe, S. J., & Fabre-Thorpe, M. (2004). How parallel is visual processing in the ventral pathway? *Trends in Cognitive Sciences*, *8*(8), 363–370. https://doi.org/10.1016/j.tics.2004.06.003

Ruggeri, P., Meziane, H. B., Koenig, T., & Brandner, C. (2019). A fine-grained time course investigation of brain dynamics during conflict monitoring. *Scientific Reports*, *9*(1), 3667. https://doi.org/10.1038/s41598-019-40277-3

Salminen, N. H., Takanen, M., Santala, O., Lamminsalo, J., Altoè, A., & Pulkki, V. (2015). Integrated processing of spatial cues in human auditory cortex. *Hearing Research*, *327*, 143–152. https://doi.org/10.1016/j.heares.2015.06.006

Sand, A., & Wiens, S. (2011). Processing of unattended, simple negative pictures resists perceptual load. *NeuroReport*, *22*(7), 348–352. https://doi.org/10.1097/WNR.0b013e3283463cb1

Schnuerch, R., Kreitz, C., Gibbons, H., & Memmert, D. (2016). Not quite so blind: Semantic processing despite inattentional blindness. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(4), 459–463. https://doi.org/10.1037/xhp0000205

Schoenfeld, M. A., Hopf, J.-M., Merkel, C., Heinze, H.-J., & Hillyard, S. A. (2014). Object-based attention involves the sequential activation of feature-specific cortical modules. *Nature Neuroscience*, *17*(4), 619–624. https://doi.org/10.1038/nn.3656

Sigman, M., & Dehaene, S. (2006). Dynamics of the Central Bottleneck: Dual-Task and Task Uncertainty. *PLoS Biology*, *4*(7), e220. https://doi.org/10.1371/journal.pbio.0040220

Simons, D. J., & Chabris, C. F. (1999). Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events. *Perception*, *28*(9), 1059–1074. https://doi.org/10.1068/p281059

Soutschek, A., Stelzel, C., Paschke, L., Walter, H., & Schubert, T. (2015). Dissociable Effects of Motivation and Expectancy on Conflict Processing: An fMRI Study. *J Cogn Neurosci*, *27*(2), 409–423. https://doi.org/10.1162/jocn_a_00712

Stefanics, G., Csukly, G., Komlósi, S., Czobor, P., & Czigler, I. (2012). Processing of unattended facial emotions: A visual mismatch negativity study. *NeuroImage*, *59*(3), 3042–3049. https://doi.org/10.1016/j.neuroimage.2011.10.041

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. https://doi.org/10.1016/j.actpsy.2010.02.006

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, *76*(3), 282–299. https://doi.org/10.1037/h0027242

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Turatto, M., Mazza, V., & Umiltà, C. (2005). Crossmodal object-based attention: Auditory objects affect visual processing. *Cognition*, *96*(2), B55–B64. https://doi.org/10.1016/j.cognition.2004.12.001

Tusche, A., Kahnt, T., Wisniewski, D., & Haynes, J. D. (2013). Automatic processing of political preferences in the human brain. *NeuroImage*, *72*, 174–182. https://doi.org/10.1016/j.neuroimage.2013.01.020

van Gaal, S., de Lange, F. P., & Cohen, M. X. (2012). The role of consciousness in cognitive control and decision making. *Frontiers in Human Neuroscience*, *6*(May), 1–15. https://doi.org/10.3389/fnhum.2012.00121

van Gaal, S., Lamme, V. A. F., Fahrenfort, J. J., & Ridderinkhof, K. R. (2010). Dissociable Brain Mechanisms Underlying the Conscious and Unconscious Control of Behavior. *J.Cogn Neurosci.*, *0898-929X (Linking)*, 91–105. https://doi.org/10.1162/jocn.2010.21431

van Gaal, S., Naccache, L., Meuwese, J. D. I., van Loon, A. M., Leighton, A. H., Cohen, L., & Dehaene, S. (2014). Can the meaning of multiple words be integrated unconsciously? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1641), 20130212–20130212. https://doi.org/10.1098/rstb.2013.0212

van Gaal, S., Ridderinkhof, K. R., Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. F. (2008). Frontal Cortex Mediates Unconsciously Triggered Inhibitory Control. *Journal of Neuroscience*, *28*(32), 8053–8062. https://doi.org/10.1523/JNEUROSCI.1278-08.2008

van Gaal, S., Ridderinkhof, K. R., van den Wildenberg, W. P. M., & Lamme, V. A. F. (2009). Dissociating Consciousness From Inhibitory Control: Evidence for Unconsciously Triggered Response Inhibition in the Stop-Signal Task. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(4), 1129–1139. https://doi.org/10.1037/a0013551

Van Schie, H. T., Mars, R. B., Coles, M. G. H., & Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience*, *7*(5), 549–554. https://doi.org/10.1038/nn1239

van Veen, V., Cohen, J. D., Botvinick, M. M., Stenger, V. A., & Carter, C. S. (2001). Anterior Cingulate Cortex, Conflict Monitoring, and Levels of Processing. *NeuroImage*, *14*(6), 1302–1308. https://doi.org/10.1006/nimg.2001.0923

VanRullen, R. (2007). The power of the feed-forward sweep. *Advances in Cognitive Psychology*, *3*(1–2), 167–176. https://doi.org/10.2478/v10053-008-0022-3

Wang, K., Li, Q., Zheng, Y., Wang, H., & Liu, X. (2014). Temporal and spectral profiles of stimulus-stimulus and stimulus-response conflict processing. *NeuroImage*, *89*, 280–288. https://doi.org/10.1016/j.neuroimage.2013.11.045

Wegener, D., Galashan, F. O., Aurich, M. K., & Kreiter, A. K. (2014). Attentional spreading to task-irrelevant object features: Experimental support and a 3-step model of attention for object-based selection and feature-based processing modulation. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00414

Weisz, N., Müller, N., Jatzev, S., & Bertrand, O. (2014). Oscillatory alpha modulations in right auditory regions reflect the validity of acoustic cues in an auditory spatial attention task. *Cerebral Cortex*, *24*(10), 2579–2590. https://doi.org/10.1093/cercor/bht113

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. https://doi.org/10.1080/14639220210123806

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences*, *90*(18), 8722–8726. https://doi.org/10.1073/pnas.90.18.8722

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? In *Nature Reviews Neuroscience*. https://doi.org/10.1038/nrn1411

Xu, Y. (2010). The Neural Fate of Task-Irrelevant Features in Object-Based Processing. *Journal of Neuroscience*, *30*(42), 14020–14028. https://doi.org/10.1523/JNEUROSCI.3011-10.2010

Yeung, N., Cohen, J. D., & Botvinick, M. M. (2011). Errors of interpretation and modeling: A reply to Grinband et al. *NeuroImage*, *57*(2), 316–319. https://doi.org/10.1016/j.neuroimage.2011.04.029

Zäske, R., Perlich, M. C., & Schweinberger, S. R. (2016). To hear or not to hear: Voice processing under visual load. *Attention, Perception, and Psychophysics*, *78*(5), 1488–1495. https://doi.org/10.3758/s13414-016-1119-2

Zhao, J., Liang, W. K., Juan, C. H., Wang, L., Wang, S., & Zhu, Z. (2015). Dissociated stimulus and response conflict effect in the Stroop task: Evidence from evoked brain potentials and brain oscillations. *Biological Psychology*, *104*, 130–138. https://doi.org/10.1016/j.biopsycho.2014.12.001

Zimmer, U., Itthipanyanan, S., Grent-'T-Jong, T., & Woldorff, M. G. (2010). The electrophysiological time course of the interaction of stimulus conflict and the multisensory spread of attention. *European Journal of Neuroscience*, *31*(10), 1744–1754. https://doi.org/10.1111/j.1460-9568.2010.07229.x

Zylberberg, A., Dehaene, S., Roelfsema, P., & Sigman, M. (2011). The human Turing machine: A neural framework for mental programs. *Trends in Cognitive Sciences*, *15*, 293–300. https://doi.org/10.1016/j.tics.2011.05.007

Zylberberg, A., Fernández Slezak, D., Roelfsema, P. R., Dehaene, S., & Sigman, M. (2010). The Brain's Router: A Cortical Network Model of Serial Processing in the Primate Brain. *PLoS Computational Biology*, *6*(4). https://doi.org/10.1371/journal.pcbi.1000765