# Frontal and parietal alpha oscillations reflect attentional modulation of cross-modal matching

Jonas Misselhorn[1*], Uwe Friese[1,2], & Andreas K. Engel[1]

[1]University Medical Center Hamburg-Eppendorf, Department of Neurophysiology and Pathophysiology, Hamburg, Germany
[2]University of Osnabrück, Institute of Cognitive Science, Osnabrück, Germany

[*]Correspondence: j.misselhorn@uke.de

## Abstract

Multisensory perception is characterised by attentional selection of relevant sensory inputs and exploitation of cross-modal similarities that promote cross-modal binding. Underlying mechanisms of both top-down and bottom-up modulations have been linked to changes in alpha/gamma dynamics in primary sensory cortices. Accordingly, it has been proposed that alpha oscillations provide pulsed inhibition for gamma activity and thereby dynamically route cortical information flow. In this study, we employed a recently introduced multisensory paradigm incorporating both bottom-up and top-down aspects of cross-modal attention in an EEG study. The same trimodal stimuli were presented in two distinct attentional conditions, focused on visual-tactile or audio-visual components, for which cross-modal congruence of amplitude changes had to be evaluated. Neither top-down nor bottom-up cross-modal attention modulated alpha or gamma power in primary sensory cortices. Instead, we found alpha band effects in bilateral frontal and right parietal cortex. We propose that frontal alpha oscillations reflect the origin of top-down control regulating perceptual gains and that parietal alpha relates to sensory re-orienting. Taken together, we suggest that the idea of selective cortical routing via alpha oscillations can be extended from sensory cortices to the fronto-parietal attention network.

## Acknowledgements

This work was supported by grants from the German Research Foundation (SFB 936/A3 and SFB TRR 169/B1 to A.K.E. as well as DFG FR-3366-1 to U.F.), the EU (ERC-2010-AdG-269716 to A.K.E.) and the Landesforschungsförderung Hamburg (CROSS FV 25 to A.K.E.). We thank Bettina Schwab for invaluable discussions on the manuscript and Nina Noverijan for assistance in data recording.

## Author contributions

J.M., U.F. and A.K.E. designed the experiment. J.M. recorded the data. J.M. analysed the data. J.M. wrote the main manuscript text. J.M., U.F. and A.K.E. reviewed the manuscript.

## Competing Interests

The authors declare that they have no competing interests.

## Data Availability

Behavioural and electrophysiological data will be made available upon request to the corresponding author.

# Introduction

Human perception is governed by constant influx of information through multiple sensory channels. The act of perceiving routes information flow by active engagement with the multisensory environment, causing sensory inputs to be constantly shaped by modulatory signals reflecting behavioural goals, contextual demands and structural properties of the environment. Spotting a singing bird in a tree, for instance, does not depend on tactile processing but on evaluating visual and auditory signals for temporo-spatial congruence. Lacing shoes, on the other hand, makes little use of audition but integrates vision and tactile perception in a goal-directed manner. These examples illustrate that multisensory perception is shaped by top-down and bottom-up modulation of sensory inputs. Attempts to understand multisensory perception accordingly need to address neural mechanisms underlying both selection of relevant sensory input and exploitation of cross-modal similarities that promote cross-modal binding.

A well-described mechanism of stimulus selection via attentional modulation is gain regulation of population responses in sensory regions[1,2]. In MEG and EEG studies, these gain regulations are likely reflected in alpha band dynamics[3]. Jensen and Mazaheri[4] propose that alpha band activity plays a general role in the up- and down-regulation of cortical processing capabilities ("gating by inhibition"). By pulsed inhibition, alpha oscillations could effectively gate gamma band activity related to active processing[5]. This has been shown repeatedly in the context of spatial attention[6,7,8] while evidence supporting its applicability to cross-modal attention is sparse[9,10]. Additionally, it is unclear whether pulsed inhibition regulates cortical processing beyond sensory cortices, for instance in cortical regions exerting top-down control.

Mechanisms underlying stimulus-driven cross-modal binding are less well understood. In fact, it remains a matter of debate at what stage of cortical processing such interactions take place[11,12,13]. While some evidence suggests that input to distinct modalities is processed in parallel and only converges later in regions of the temporal and parietal lobe[14,15,16], other evidence points out that interactions can already take place at the level of primary sensory regions[17,18,19]. The disparity of findings is not surprising given that factors driving cross-modal integration span from psychophysical (spatial/temporal congruence) to memory-dependent (semantic congruence and cross-modal correspondences). Yet, a linking observation is that low- and high-level integration have been associated with changes in gamma band activity[13,20,21].

In the EEG study reported here, we employed a recently introduced multisensory paradigm incorporating both bottom-up and top-down aspects of cross-modal attention[22]. This paradigm involved a trimodal stimulus consisting of a visual, an auditory and a tactile

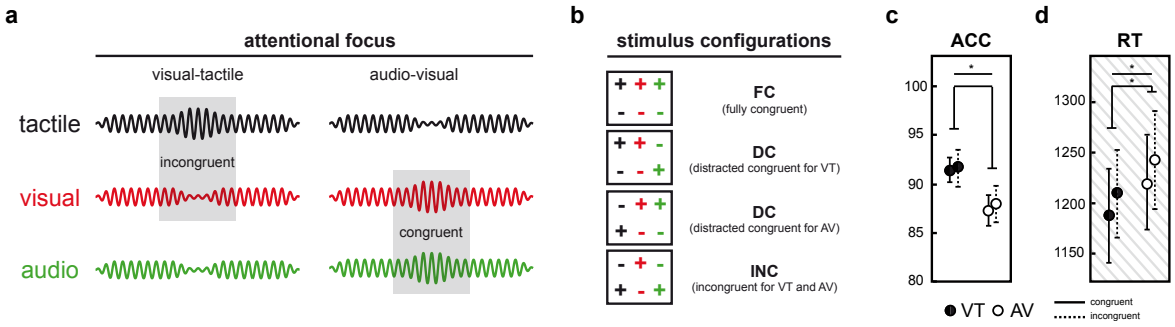



**Figure 1: Schematic of matching task and behavioural results**. **a)** Illustration of example trials in the two attentional conditions. Under visual-tactile focus (left), the changes in the target stimuli are incongruent, as the tactile stimulus (black) increases in intensity while the visual stimulus (red) decreases. Under audio-visual focus (right), target change is congruent, since visual and auditory (green) stimuli both increase in intensity. **b)** In each block, all possible stimulus configurations occurred with equal probability. Intensity changes are depicted by coloured plus/minus signs (colour coding as introduced in a)). **c)** Accuracy (ACC) in percentage correct. Error bars represent standard deviations. There was no effects of experimental conditions. **d)** Reactions times in milliseconds (RTs). Error bars represent standard deviations. *ATTENTION* as well as *CONGRUENCE* significantly affected the timing of responses, but not the accuracy of responding (Note: RTs were collected in a previous behavioural study[21] from the same sample of participants).

component that each underwent a brief increase or decrease in intensity. Participants had to attend two of the stimuli and had to decide whether the attended pair changed congruently or incongruently. In a similar study investigating audio-visual matching with MEG, changes in primary sensory alpha and gamma activity were more profound when participants attended presentations compared to when they were ignored[10]. In order to further investigate whether this modulation of alpha/gamma band dynamics holds in situations where attention is not holistic but rather modality-based, we presented the same trimodal stimuli in two distinct attentional conditions (top-down), focused on either visual-tactile (VT) or audio-visual (AV) components, for which cross-modal congruence (bottom-up) of amplitude changes had to be evaluated (Fig. 1 a). We expected top-down cross-modal attention to selectively enhance primary sensory alpha activity for irrelevant modalities and decrease alpha activity for attended modalities. This increase/decrease in alpha power might be accompanied by a decrease/increase in gamma band activity. As a bottom- up effect of cross-modal binding, gamma band activity in sensory cortices or temporal/parietal cortex is expected to be modulated by cross-modal congruence.

# Results

## Psychophysics and behaviour

The trimodal stimulus material was designed such that the target amplitude changes in each modality were equally salient. This was achieved by estimating detection thresholds for each modality and change direction separately using a psychophysical staircase procedure[23]. Yet, a questionnaire that was completed during debriefing of a preceding behavioural study[22]

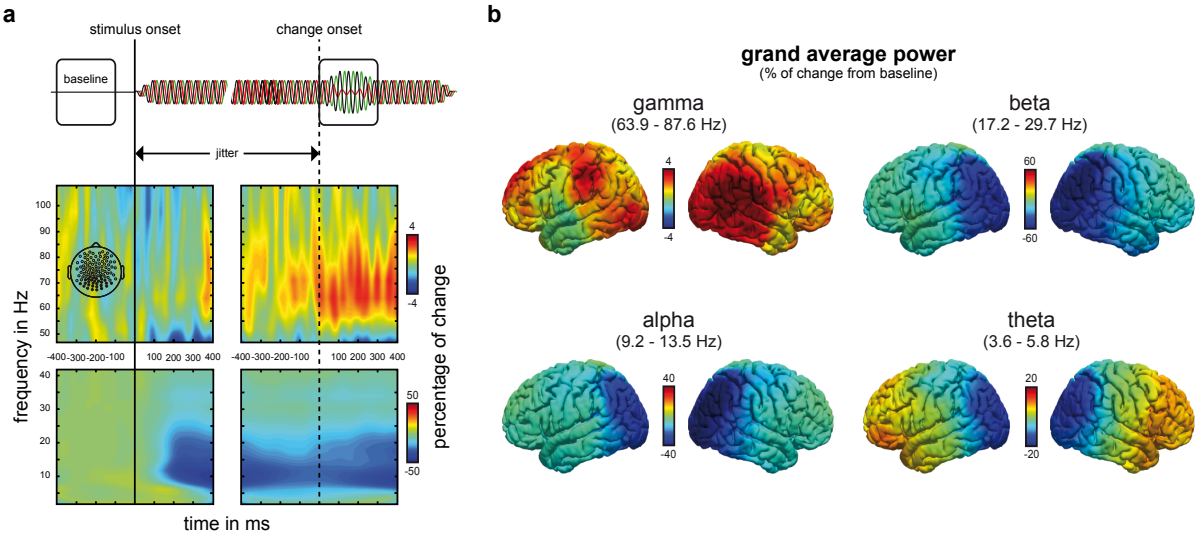



**Figure 2: Average time-frequency dynamics in sensor and source space. a)** Time-frequency dynamics in sensor space at occipito-parietal channels (topography shown in upper left panel), time locked to stimulus onset (left panels, black solid vertical line) and change onset (right panels, black dashed vertical line. ***Top:*** Schematic of temporal trial structure displaying the two relevant time-windows used for analysis: a baseline (-400:-100 ms) separated by a jitter (stimulus to change onset between 700 and 1000 ms) from the change interval (0:300 ms). ***Bottom:*** Time-frequency plots from posterior sensors. Values represent percentage of change from baseline. **b)** Distribution of band-limited power on the cortical surface in the change interval relative to baseline.

indicated that subjective salience of the sensory components was in fact not equal but strongest for the visual component. In particular, participants reported that the visual component was hardest to ignore when it was task irrelevant (pairwise Wilcoxon rank sum test, Bonferroni corrected, V-T: $p$ = .002, V-A: $p$ = .24, T-A: $p$ = .26). This should be kept in mind for the discussion of the effects of cross-modal attention.

The timing and accuracy of responding was analysed with a repeated measure analysis of variance (ANOVA) with factors *ATTENTION* (VT vs. AV) and *CONGRUENCE* (congruent vs. incongruent). Responses were faster but not more accurate when subjects attended cross-modally congruent pairs (Fig. 1 d; $p < .001$, $\eta_p^2 = 0.471$). When participants attended VT, timing as well as accuracy of responding was significantly better compared with the AV conditions (RT: $p$ = .036, $\eta_p^2$ = 0.201; ACC: $p$ = .005, $\eta_p^2$ = 0.327). No interaction effects between *ATTENTION* and *CONGRUENCE* were observed.

## ROI analysis

In Figure 2, we present an overview of time-frequency dynamics during the task as well as distributions of band-limited power for theta, alpha, beta and gamma bands in source space. In order to investigate power changes in oscillatory activity in primary sensory areas occurring after the stimulus increases or decreases, we conducted a regions of interest (ROI) analysis on source projected EEG data in primary visual, auditory and somatosensory cortex. Statistical evaluation was carried out with a repeated measures ANOVA with factors *ROI*, *ATTENTION*

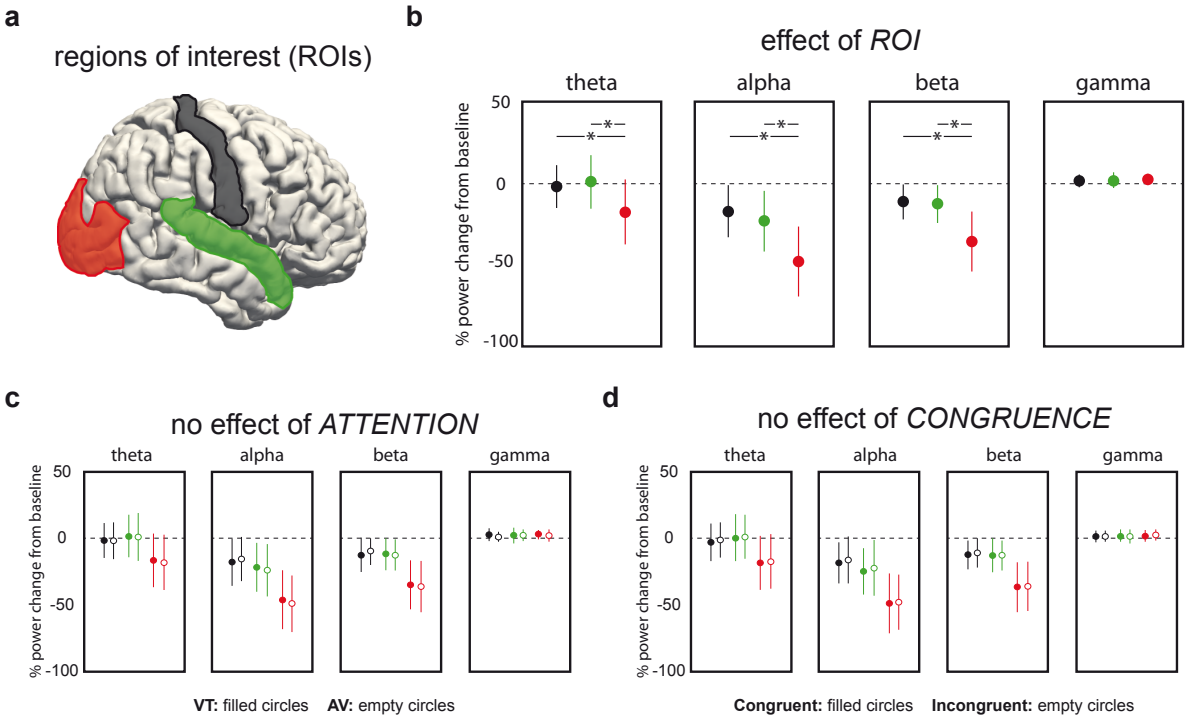



**Figure 3: Regions of interest (ROI) analysis of band-limited power during change epoch in primary sensory areas. a)** Primary sensory areas used for ROI ANOVA with factors *ATTENTION* (VT vs. AV), *CONGRUENCE* (congruent vs. incongruent) and *ROI* (visual vs. auditory vs. somatosensory); visual = red, auditory = green, somatosensory = black. **b)** Effect of *ROI* is significant for theta, alpha and beta bands, but not gamma band. Asterisk signifies significant comparisons ($p < .001$). **c)** No significant effect of *ATTENTION*. **d)** No significant effect of *CONGRUENCE*.

and *CONGRUENCE* separately for each frequency band (Fig. 3, see *Methods* for details). Significant effects of *ROI* were observed for theta, alpha and beta bands (for all, $p < .001$ and $\eta_p^2 > 0.5$; see Fig. 3 b). In the gamma band, *ROI* did not explain a significant amount of variance ($p = .689$). Simple effects analysis for the lower frequency bands showed that decreases in power were significantly stronger in visual compared with both auditory and somatosensory ROIs (for all comparisons, $p < .001$; Fig. 3 b). Power changes in auditory and somatosensory ROIs did not differ (for all, $p > .05$). The main effects and interactions of *ATTENTION* and *CONGRUENCE* were not significant (for all, $p > .05$; Fig. 3 c + d).

## Cluster statistics

To complement the ROI analysis, we conducted a whole-brain analysis of task-related power changes in the interval after stimulus increases and decreases evaluated by means of nonparametric cluster-based permutation statistics (see *Methods* for details). For the *ATTENTION* contrast (VT *minus* AV), significant differences were found in the alpha band. In two roughly symmetric clusters, VT attention was associated with stronger power decrease of alpha oscillations when compared with AV attention (Fig. 4 a). In the left hemisphere, the cluster was situated in the border region of pre-central gyrus, middle frontal gyrus (MFG) and

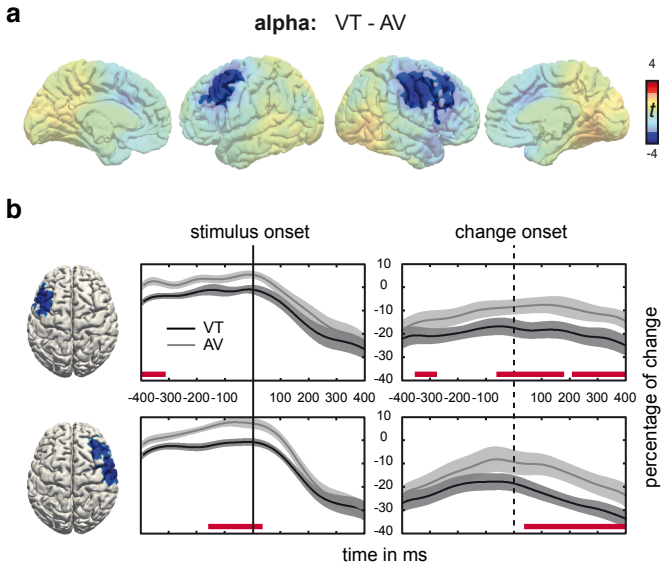



**Figure 4: Effect of *ATTENTION* (VT - AV). a)** Cluster-based permutation statistic of the contrast VT *minus* AV in the alpha band during change interval. T-values of the respective contrasts are depicted. Shaded voxels are non-significant. **b)** Time-course of alpha activity within significant clusters. Shading corresponds to the standard error of the mean. Red bars indicate temporal regions of significance as determined by permutation testing (see *Methods* for details).

superior frontal gyrus encompassing the frontal eye fields (FEF; $p$ = .031). In the right hemisphere, the cluster was situated similarly but expanded further into pre- and post-central gyrus ($p$ = .005). In a next step, we analysed time courses of alpha power modulations within these two clusters (Fig. 4 b, see *Methods*). Throughout the entire time course, alpha power was lower for VT than for AV attention. This difference was significant before stimulus onset in both hemispheres (left: [-400; -312] ms, right: [-165; 43] ms), in a short epoch prior to change onset in the left hemisphere ([-348; -286] ms) and throughout the whole change epoch in both hemispheres (left: [-69; 177] ms and [211; 400] ms, right: [43; 400] ms).

When evaluating the effect of *CONGRUENCE* (attended congruent *minus* attended incongruent), significant differences were found in alpha and theta bands (Fig. 5). Theta power was higher for congruent trials compared to incongruent trials in large parts of the medial wall of both hemispheres (Fig. 5 a). This effect was significant in a left hemispheric cluster stretching from posterior to anterior cingulate cortex. In a next step, we analysed the contributions of fully congruent (FC) and distracted congruent (DC) trials to the overall effect of congruence (see Fig. 1 b and *Methods* for details). It was driven by FC trials for which theta power was increased in cingulate cortex of both hemispheres, and in small clusters in left and right intraparietal sulcus (IPS) as well as left inferior frontal gyrus (IFG) and right MFG (Fig. 5 c). No significant theta power differences were seen when comparing distracted congruent with incongruent trials (Fig. 5 e).

In the alpha band, cross-modal congruence modulated power in large parts of medial and lateral cortex. In incongruent trials, stronger decrements of alpha power occurred in bilateral medial superior frontal cortex and left MFG (Fig. 5 b). Congruent trials were associated with stronger decrements in alpha power in bilateral medial occipito-parietal cortex and right

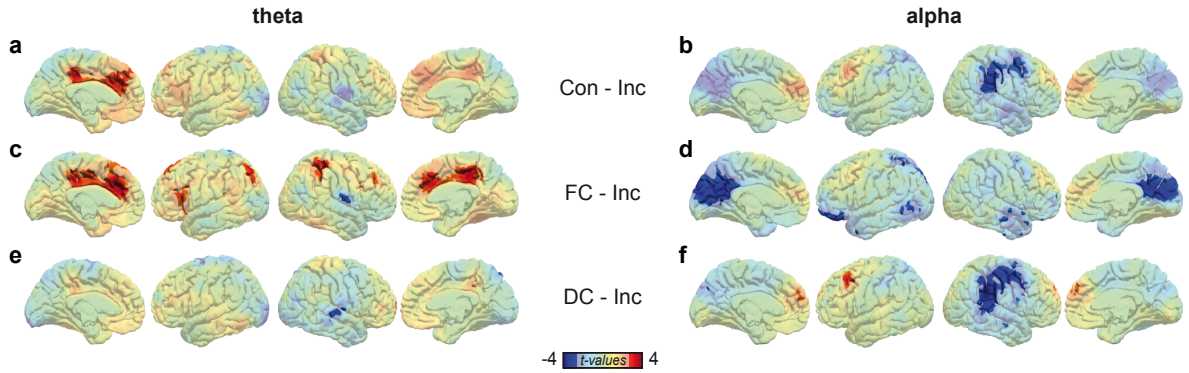



**Figure 5. Effect of *CONGRUENCE* (Congruent - Incongruent). a+b)** Cluster-based permutation statistic of the contrast *Congruent* - *Incongruent* in the alpha (a) and theta (b) band during change interval. All subplots depict t-values. Shaded voxels are non-significant. **c)** Contrast between theta power in *fully congruent* (FC) and *incongruent* trials. **d)** Contrast between alpha power in *fully congruent* (FC) and *incongruent* trials. **e)** Contrast between theta power in *distracted congruent* (DC) and *incongruent* trials. **f)** Contrast between alpha power in *distracted congruent* (DC) and *incongruent* trials.

inferior parietal and middle frontal cortex. The latter difference was significant in a cluster covering the right temporoparietal junction (TPJ), but stretching rostrally towards right MFG. Next, we disentangled contributions from FC and DC trials as before for theta power. While FC trials significantly drove the effect in medial occipito-parietal cortex (Fig. 5 d), DC trials contributed significantly to the effect in medial and left middle frontal cortex (Fig. 5 e). The effect in right TPJ and MFG was dominated by DC trials (Fig. 5 d).

# Discussion

We investigated bottom-up and top-down modulation of sensory processing in a cross-modal matching task involving visual, auditory and tactile perception. Contrary to our expectations, we did not find alpha/gamma oscillations in primary sensory areas to be modulated by bottom-up or top-down cross-modal attention. This finding is surprising given that both processes have often been noted to be accompanied by alpha/gamma modulations in sensory cortices[3,6,7,10,18,19]. Our explanation for the lack of primary sensory modulation is the nature of the task: many, if not most, multisensory studies employ detection tasks with near-threshold sensory stimulation. In these situations of low sensory drive, both bottom-up and top-down modulation of sensory input can be expected to have higher impact compared with situations of strong sensory drive. Stimulus-driven cross-modal enhancement by spatio-temporal congruence, for instance, is assumed to obey the law of inverse effectiveness, meaning that there is an inverse relationship between possible cross-modal enhancement and stimulus intensity[24]. Here, however, all stimulus intensities were clearly supra-threshold with superimposed amplitude changes. Top-down as well as bottom-up modulation of sensory processing might thus be subtle and hence not detectable by EEG. Instead, we found theta oscillations in cingulate cortex and most notably alpha oscillations in frontal and parietal cortex to be modulated. In the following, we propose that frontal alpha oscillations reflect the origin

of top-down control regulating perceptual gains and that parietal alpha oscillations relate to sensory (re-)orienting. Theta activity in cingulate cortex is finally discussed in the context of adaptive task-switching behaviour circumventing cross-modal matching.

Reduction of alpha power in bilateral FEF and MFG as well as right pre-/post-central gyrus was stronger for VT cross-modal attention compared to AV attention. This difference was significant even before stimulus onset. Besides their role in oculomotor control, the FEF have been described as important structures in top-down attention[25]. In a study using TMS, Grosbas and Paus showed that disruption of activity in FEF shortly before the onset of the target in a visuospatial covert attention task facilitated responses[26]. Conversely, 10 Hz TMS over the right FEF was shown to impair visual search of unpredictable items with low salience[27]. Moore and Armstrong reconciled this conflicting evidence by suggesting that the FEF has a general role in regulating visual gain[28]. In their study, electric stimulation of FEF in the monkey either enhanced or inhibited responses to visual stimuli in V4 depending on whether retinopically corresponding sites were stimulated. Studies in humans supported this idea by showing that TMS over FEF could increase phosphene or contrast sensitivity of extrastriate cortex[29,30]. This top-down modulation of visual cortex was demonstrated even in the absence of sensory input[31]. Likewise, anticipatory alpha and stimulus related gamma activity in occipito-parietal cortex could be modulated by TMS over FEF[32]. The strength of modulation was shown to correlate with the strength of structural connectivity between frontal and parietal cortex via the superior longitudinal fasciculus[33]. Thus, animal and human studies jointly conclude that the FEF can dynamically modulate the gain of down-stream visual cortex independent of sensory input. In our study, FEF/MFG likely facilitated cross-modal matching by modulating visual gain to counter visual dominance. Although stimulus intensity was titrated to be balanced across modalities (see *Methods*), we have reason to assume that perceived salience was highest for the visual component. In a questionnaire that was completed during debriefing of the preceding behavioural study, we asked participants to rank the difficulty to ignore a given modality. Most participants reported that visual components were hardest and tactile components easiest to ignore. This finding is in line with a pattern of sensory dominance found for combinations of visual, auditory and somatosensory stimuli in a discrimination task[34]. Sensory dominance can be problematic under the assumption that cross-modal matching is not independent of perceptual gain. This is most likely the case for stimulus-driven aspects of multisensory integration – the idea of inverse effectiveness, after all, assumes multimodal stimuli of low but comparable intensity[24]. Consequentially, decreased power of alpha oscillations in FEF/MFG is taken as evidence for an increased down-regulation of visual gain in VT conditions to account for unequal subjective salience of the stimuli to be matched.

As discussed above, balancing perceptual gains across modalities by top-down modulation likely enables optimal use of stimulus-driven aspects of cross-modal matching. These bottom-up factors were ubiquitous in this task; on each trial, participants were simultaneously confronted with three salient events, that is, intensity changes in each modality. Although each change of intensity by itself possessed some degree of bottom-up salience, we suggest that cross-modal congruence amplified salience through cross-modal binding[13]. When cross-modal binding was enhanced between attended modalities, responses were facilitated. This was especially pronounced for fully congruent trials where conflict, and thus the need for actual matching, was absent. All other trials were either distracted congruent (attended modalities change congruently, but the distractor diverged) or attended incongruent (one of the attended modalities is congruent to the distractor). In these cases, cross-modal binding was always stronger between two given modalities compared to the respective third. When contrasting the EEG of these trials, we find alpha band effects in the right temporoparietal junction (rTPJ) and right MFG. Specifically, distracted congruent conditions were associated with decreased power of alpha oscillations in these regions compared with attended incongruent trials. In accordance with the "gating by inhibition" theory, we conclude that the rTPJ/rMFG were more strongly disinhibited when attended modalities had a stronger bottom-up drive for cross-modal binding. The TPJ receives inputs from visual, auditory and somatosensory cortex and is richly connected to temporal and frontal sites, making it an important hub for the interaction of multisensory integration and attention[35]. Accordingly, lesions to the right TPJ typically result in neglect[36,37]. A dominant interpretation of rTPJ's functional role is its involvement in (spatial) re-orienting based on stimulus salience[38]. In a model integrating goal-directed and stimulus-driven attention, it is suggested that a dorsal network comprising FEF and IPS instantiates attentional sets. As a counterpart, a ventral network comprising rTPJ and right ventral frontal gyrus mediates bottom-up signals acting as a circuit-breaker for the dorsal system. Studies employing multisensory paradigms have noted rTPJ's involvement in processing cross-modal congruence. In a study investigating visual-tactile pattern matching, pre-stimulus alpha and beta power in right supramarginal gyrus differentiated between detection and congruence-evaluation tasks[39]. Another study showed that alpha power in right posterior regions was more strongly supressed during congruent compared with incongruent audio-visual speech presentations[40]. Taken together with our results, we suggest that the rTPJ detects the increased salience of congruent cross-modal events. While each trial might, in principle, result in attentional capture by any of the three modalities, cross-modal binding by congruence might serve as a reliable "cue" for re-orienting towards the relevant modalities. Thereby, cross-modal binding between attended modalities might support modality-based re-orienting.

As pointed out above, fully congruent trials were characterised by the absence of cross-modal conflict. In the EEG, these highly salient trials were associated with stronger alpha power reductions in medial occipito-parietal cortex. In an event-related potentials study featuring visual, auditory and somatosensory stimuli, RT facilitation was correlated with the latency of the P300, which was localised in precuneus[40]. Other research suggests that alpha power reductions in occipito-parietal cortex and P300 dynamics are functionally coupled[41]. Here, enhanced involvement of medial occipito-parietal cortex is proposed to reflect increased bottom-up salience due to multisensory enhancement, i.e., increased perceptual gains of concurrent congruent sensory input to more than one modality[42]. In addition to mere bottom-up sensory salience, fully congruent stimuli occurred in only 25 % of all trials and were thereby salient. Actual cross-modal matching was required only in the remaining 75 % of trials where two modalities changed congruently while the third modality diverged. An efficient strategy would accordingly be to "switch" between these two tasks, i.e., between detecting highly salient events and cross-modal matching of conflicting input. In addition to the alpha band effect in precuneus, fully congruent trials were also associated with a relative increase in theta power in bilateral cingulate cortex. Theta band activity in cingulate cortex has previously been related to the adjustment of stimulus response mappings[43]. Together with insular cortex, cingulate cortex is part of a salience network which has importance for both bottom-up detection of salient events and switching between large-scale networks to adaptively control behaviour[44]. Here, it is suggested that reduced alpha power in medial occipito-parietal cortex related to multisensory enhancement acts as a salience signal detected by cingulate cortex which in turn initiates adaptive task-switching behaviour.

Taken together, we provide evidence that cross-modal matching in complex multisensory environments heavily relies on mechanisms of attention. Our results contrast with the majority of studies on multisensory integration concerned with stimulus detection where attentional load is typically low. Here, participants were confronted with a highly challenging multisensory setting. In order to counter the bias imposed by visual dominance, top-down regulation of perceptual gains likely supported an optimal exploitation of cross-modal similarities that promote perceptual binding. This was associated with decreased alpha band power in frontal cortices proposed to reflect the origin of top-down modulation. Likewise, bottom-up drive for cross-modal binding was related to changes in alpha power in right parietal cortex proposed to represent the bottom-up modulatory signal underlying sensory re-orienting. Both findings provide evidence for an extension of the idea that alpha/gamma dynamics indicate selective cortical routing beyond sensory cortex to the fronto-parietal attention network.

# Methods

## Participants

Twenty-one participants entered the study and received monetary compensation for their participation. They were on average 23.8 ± 2.5 years old and 11 of them were female (10 male). Vision, audition and tactile perception were normal and none of them had a history of neurological or psychiatric disorders. After an explanation of the experimental procedure, participants gave written consent. The ethics committee of the University Medical Center Hamburg-Eppendorf approved the study which was carried out in accordance with the declaration of Helsinki.

## Experimental design

On each trial, we presented a trimodal stimulus consisting of a visual, an auditory and a tactile component that each underwent a brief increase or decrease in intensity (see *Stimulus Material* for details). Block-wise, participants attended either visual-tactile (VT) or audio-visual (AV) bimodal pairs and ignored the respective third component. The task was to decide whether the attended bimodal pairs changed congruently (i.e., in the same direction) or incongruently (i.e., in different directions; see Fig. 1 a). Verbal responses had to be withheld until stimulus offset to minimise myogenic artifacts. Therefore, reaction times (RTs) could not be evaluated. Instead, we present RT data of the same sample of participants from the behavioural study preceding the EEG study[22]. In each block, all possible eight stimulus configurations of increases and decreases across modalities were presented with equal probability (Fig. 1 b). VT and AV blocks containing 64 trials presented in randomised order were alternating. Data were collected on two separate days with identical experimental procedure so that EEG data of 1280 trials was collected from each participant. Prior to statistical analysis, trials were pooled without taking change direction into account. For instance, fully congruent trials were both trials where all modalities underwent decrements and trials where all modalities underwent increments (Fig. 1 b, pooling is indicated by boxes).

## Stimulus material

Visual contrast, auditory loudness and vibration strength were experimentally increased or decreased. The magnitudes of change per modality and direction were individually estimated prior to the experimental sessions using the same psychometric step function as described in Misselhorn et al. (QUEST)[22,23]. Intensity changes had a duration of 300 ms and onsets were jittered across trials between 700 and 1000 ms after stimulus onset (Fig. 2 a). In total, sensory stimulation had a fixed duration of 2 s. As visual stimulation, an expanding circular grating was centrally presented against a grey background on a CRT screen with a visual angle of 5°. The

auditory component consisted of a complex sinusoidal tone (13 sine waves: 64 Hz and its first 6 harmonics as well as 91 Hz and its first 5 harmonics, low-frequency modulator: 0.8 Hz) played back with audiometric insert earphones binaurally at 70 dB (E-A-RTONE 3A, 3M, USA). As tactile stimulation, high-frequency vibrations (250 Hz on C2 tactors, Engineering Acoustics Inc., USA) were delivered to the tips of both index fingers.

## EEG

EEG was recorded from 128 active electrodes (Easy Cap, Germany) including four ocular electrodes referenced to the nose. Data was sampled at 1000 Hz with an amplitude resolution of 0.1 µV using BRAINAMP MR amplifier (Brain Products, Germany) and digitised after analog filtering (low cutoff: 10 s, high cutoff: 1000 Hz). Offline, data was down-sampled to 500 Hz and digitally filtered (high-pass: 1 Hz, low-pass: 120 Hz, notch: 49-51 Hz, 99-101 Hz). Epochs of 2.5 s were cut from -500 ms relative to stimulus onset until stimulus offset and normalised to the pre-stimulus baseline. Next, data was re-referenced to the common average and linear trends were removed from all epochs. From the four ocular channels, two bipolar channels for horizontal and vertical eye movements were derived.

**Pre-processing.** Trials with incorrect answer and large non-stereotypical artifacts were excluded from further processing. Subsequently, independent component analysis (ICA) was performed separately for low and high frequency bands (low band: 1-30 Hz, high band: 30-120 Hz). Thereby, stereotypical low-frequency artifacts (for instance eye movements and heart beat) and high-frequency artifacts (i.e. myogenic activity) could be separated more reliably from neuronal activity. For both bands, principal components analysis was performed first to reduce data such that 99 % of variance is retained. Subsequently, ICA was performed on the rank-reduced data using the infomax algorithm[45]. Artifactual ICs were identified and rejected with respect to time course, spectrum and sensor topography[46]. For the high band, saccade-related transient potentials were removed additionally[47]. Finally, all epochs were visually inspected and epochs with remaining artifacts were rejected. Furthermore, a subset of 19 electrodes (i.e. most outer facial, temporal and neck electrodes) was excluded from further analysis due to poor signal-to-noise ratio. Lastly, data was stratified such that all conditions in the ensuing analysis hold the same amount of data within subjects. On average, 426 ± 89 trials per participant entered the analysis.

**Source reconstruction of band-limited signals.** Cleaned data in low and high bands were joined and epoched with respect to stimulus onset as well as change onset. Prior to filtering data into narrow bands by means of wavelet analysis, event related potentials were subtracted in order to remove phase-locked responses. A family of 40 complex Morlet wavelets $w$ with lengths of 2 s was constructed for logarithmically spaced frequencies between 2 and 120 Hz.

$$w(t, f_0) = A * e^{-\frac{t^2}{2\sigma_t^2}} * e^{2\pi f_0 t}$$

The number of cycles per wavelet ($m$) were logarithmically spaced between 3 and 10 and subsequently rounded off. Wavelets were normalised by factor $A = \sigma_t \sqrt{\pi}^{-\frac{1}{2}}$ with $\sigma_t = m/2\pi f_0$. Single trial data was convolved with the Morlet wavelets by multiplication in the frequency domain using fast fourier transformation with boxcar windows. Wavelet filtered single trial data was then reconstructed in source space using exact low-resolution brain electromagnetic tomography (eLORETA; regularisation: 0.05)[48]. Lead fields were computed for a three-shell head model[49]. The customised cortical grid was derived from a cortical surface provided by Freesurfer in MNI space by reducing the number of cortical nodes from 270000 to 10000[50]. Dipole directions at each node of the cortical grid were estimated by means of singular value decomposition of the trial averaged spectral power individually for all bands and kept constant for all trials of the given participant. Induced power was computed from these source reconstructed band-limited time domain signals. Power in the epoch after change onset was baseline corrected using the baseline of the mean over all conditions from -400 to -100 ms relative to stimulus onset (Fig. 2 a). By visual inspection of the resulting time-frequency landscapes, frequency bands in the theta, alpha, beta and gamma range were chosen individually for each participant (mean values and range in parentheses; theta: 4.7 [3.6; 5.8] Hz, alpha: 11.5 [9.2; 13.5] Hz, beta: 23.0 [17.2; 29.7] Hz, gamma: 78.9 [63.9; 87.6] Hz). Statistical analysis was carried out for the post-change interval ([0; 300] ms) only.

## Statistical analysis

**Behaviour.** Accuracy of responding (ACC) within experimental conditions was analysed using a repeated-measures analysis of variance (ANOVA) with factors *ATTENTION* (VT vs. AV) and *CONGRUENCE* (congruent vs. incongruent). The timing of verbal responses was not analysed because subjects were instructed to withhold responses until stimulus offset. Instead, data from the previous behavioural study was re-analysed for the sub-sample of participants enrolled in this EEG study[21]. The same ANOVA as described above for ACCs was evaluated.

***EEG: Regions of interest (ROIs) analysis.*** Primary cortical regions for vision, audition and tactile perception were chosen from the Freesurfer atlas which is constructed by gyral identification and parcellation based on anatomical landmarks[50]. For each frequency band, baseline-corrected, time and ROI averaged data in the post-change interval was evaluated by means of ANOVA with factors *ROI* (visual vs. auditory vs. somatosensory), *ATTENTION* (VT vs. AV) and *CONGRUENCE* (congruent vs. incongruent) and. Simple effects of significant ANOVA effects were assessed by paired-sample t-tests applying Bonferroni correction.

***EEG: Whole-brain permutation statistics.*** Complementing ROI analysis, a whole brain exploratory analysis of differences between experimental conditions was conducted and evaluated by means of nonparametric cluster-based permutation statistics[51]. A null distribution was computed by randomly drawing trials into two sets per subject (300000 permutations). For each node of the cortical grid, a paired-sample t-test was computed between averaged power of the two sets and statistical maps were thresholded ($p < .05$). Significant clusters were found and the size of the largest cluster was noted. This procedure was carried out separately for the four frequency bands. Contrasts corresponding to a 2 (*ATTENTION*) x 2 (*CONGRUENCE*) design were computed and evaluated against the aforementioned null-hypothesis (cut-off: 99th percentile). Reported *p*-values are Bonferroni-corrected.

Cluster statistics were complemented by post-hoc analyses that were designed (1) to detail on the time-course of the *ATTENTION* effect and (2) to disentangle the contributions of sub-conditions to the overall effect of *CONGRUENCE*.

(1) For clusters showing a significant effect of *ATTENTION*, we computed the time course of average within cluster spectral power separately for visual-tactile and audio-visual conditions. Significance of the difference between time courses was evaluated using nonparametric cluster-based permutation statistics (300000 permutations). For each permutation, time courses were shuffled and paired-sample t-tests between VT and AV were computed for each sample. The number of samples included in the longest temporally continuous cluster of significant difference was noted to form the maximum statistic null distribution. In the original data, periods of significant difference between attentional conditions were considered significant in the temporal domain when they held more samples than the 99th percentile of the null distribution.

(2) For this analysis we differentiated according to whether attended stimulus components were "fully congruent" or "distracted congruent". Fully congruent (FC) means that all stimulus components, including the distracting modality, change congruently (that is, all components increased or decreased in intensity; Fig. 1 b, top box). Distracted congruent (DC) means that the distractor's change direction deviates from the change direction in the attended modalities (Fig. 1 b, middle boxes). In this case, the participant has to resolve the conflict between attended congruence and unattended incongruence. In order to disentangle these two scenarios, we computed contrasts of FC respectively DC against attended incongruent conditions.

# References


1. Pessoa L., Kastner S. & Ungerleider L. G. Neuroimaging studies of attention: From modulation of sensory processing to top-down control. J Neurosci. 23, 3990-3998 (2003).
2. Ling S., Liu T. & Carrascoa M. How spatial and feature-based attention affect the gain and tuning of population responses. Vision Res. 49, 1194-1204 (2009).
3. Chaumon M. & Busch N. A. Prestimulus neural oscillations inhibit visual perception via modulation of response gain. J Cogn Neurosci. 26, 2514-29 (2014).
4. Jensen O. & Mazaheri A. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. Front Hum Neurosci. 4, 186 (2010).
5. Osipova D., Hermes D. & Jensen O. Gamma power is phase-locked to posterior alpha activity. PLoS One. e3990 (2008).
6. Worden M. S., Foxe J. J., Wang N. & Simpson G. V. Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. J Neurosci. 20, RC63 (2000).
7. Siegel M., Donner T. H., Oostenveld R., Fries P., & Engel A. K. Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention. Neuron. 60, 709-19 (2008).
8. Foster J. J., Sutterer D. W., Serences J. T., Vogel E. K. & Awh E. Alpha-band oscillations enable spatially and temporally resolved tracking of covert spatial attention. Psychol Sci. 28, 929-941 (2017).
9. Mazaheri A., van Schouwenburg M. R., Dimitrijevic A., Denys D., Cools R. & Jensen O. Region-specific modulations in oscillatory alpha activity serve to facilitate processing in the visual and auditory modalities. Neuroimage. 87, 356-62 (2014).
10. Friese U., Daume J1., Göschl F., König P., Wang P. & Engel A. K. Oscillatory brain activity during multisensory attention reflects activation, disinhibition, and cognitive control. Sci Rep. 6, 32775 (2016).
11. Ghazanfar A. A. & Schroeder C. E. Is neocortex essentially multisensory? Trends Cogn Sci. 10, 278-85 (2006).
12. Koelewijn T., Bronkhorst A. & Theeuwesa J. Attention and the multiple stages of multisensory integration: A review of audiovisual studies. Acta Psychol. 134, 372-384 (2010).
13. Senkowski D., Schneider T. R., Foxe J. J., & Engel A. K. Crossmodal binding through neural coherence: implications for multisensory processing. Trends in Neurosciences. 31, 401-409 (2008).
14. Bien N., ten Oever S., Goebel R. & Sack A. T. The sound of size: Cross-modal binding in pitch-size synesthesia: A combined TMS, EEG and psychophysics study. Neuroimage 59, 663-672 (2012).
15. Schneider T. R., Debener S., Oostenveld R. & Engel A. K. Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. Neuroimage 42, 1244-54 (2008).
16. Konen C. S. & Haggard P. Multisensory parietal cortex contributes to visual enhancement of touch in humans: A single-pulse TMS study. Cereb Cortex 24, 501–507 (2014).

17. Cappe C. & Barone P. Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. Eur J Neurosci. 22, 2886-2902 (2005).

18. Senkowski D., Talsma D., Herrmann C. S. & Woldorff M. G. Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. Exp Brain Res. 166, 411-426 (2005).

19. Lakatos P., O'Connell M. N., Barczak A., Mills A., Javitt D. C. & Schroeder C. E. The leading sense: Supramodal control of neurophysiological context by attention. Neuron 64, 419-30 (2009).

20. van Atteveldt N., Murray M. M., Thut G. & Schroeder C. Multisensory integration: flexible use of general operations. Neuron 81, 1240–1253 (2014).

21. Keil J. & Senkowski D. Neural oscillations orchestrate multisensory processing. The Neuroscientist. 1-18 (2018).

22. Misselhorn J., Daume J., Engel A. K. & Friese U. A matter of attention: Cross-modal congruence enhances and impairs performance in a novel trimodal matching paradigm. Neuropsychologia 88, 113-122 (2015).

23. Watson A. B. & Pelli D. G. QUEST: a Bayesian adaptive psychometric method. Perception & Psychophysics 33, 113–120 (1983).

24. Meredith M. A. & Stein B. E. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. J Neurophysiol. 56, 640-62 (1986).

25. Corbetta M., Akbudak E., Conturo T. E., Snyder A. Z., Ollinger J. M., Drury H. A. et al. A common network of functional areas for attention and eye movements. Neuron 21, 761-773 (1998).

26. Grosbas M. H. & Paus T. Transcranial magnetic stimulation of the human frontal eye field: effects on visual perception and attention. J Cogn Neurosci. 14, 1109-20 (2002).

27. Muggleton N. G., Juan C. H., Cowey A. & Walsh V. Human frontal eye fields and visual search. J Neurophysiol. 89, 3340-3 (2003).

28. Moore T. & Armstrong K. M. Selective gating of visual signals by microstimulation of frontal cortex. Nature 421, 370-373 (2003).

29. Silvanto J., Lavie N. & Walsh V. Stimulation of the human frontal eye fields modulates sensitivity of extrastriate visual cortex. J Neurophysol. 96, 941-5 (2006).

30. Quentin R., Chanes L., Vernet M. & Valero-Cabré A. Fronto-parietal anatomical connections influence the modulation of conscious visual perception by high-beta frontal oscillatory activity. Cereb Cortex 25, 2095-2101 (2015).

31. Ruff C. C., Blankenburg F., Bjoertomt O., Bestmann S., Freeman E., Haynes J. D. et al. Concurrent TMS-fMRI and psychophysics reveal frontal influences on human retinotopic visual cortex. Curr Biol. 16, 1479-1488 (2006).

32. Marshall T. R., O'Shea J., Jensen O. & Bergmann T. O. Frontal eye fields control attentional modulation of alpha and gamma oscillations in contralateral occipitoparietal cortex. J Neurosci. 35, 1638-1647 (2015).

33. Marshall T. R., Bergmann T. O. & Jensen O. Frontoparietal structural connectivity mediates the top-down control of neuronal synchronization associated with selective attention. PLoS Biol. 13, e1002272 (2015).

34. Hecht D. & Reiner M. Sensory dominance in combinations of audio, visual and haptic stimuli. Exp Brain Res. 193, 307-14 (2009).

519 35. Donaldson P. H., Rineheart N. J. & Enticott P. G. Nonivasive stimulation of the
520 temporoparietal junction: A systematic review. Neurosci Biobehav Rev. 55, 547-572
521 (2015).

522 36. Karnath H. New insights into the functions of the superior temporal cortex. Nat Rev
523 Neurosci. 2, 568-576 (2001).

524 37. Friedrich F. J., Egly R., Rafal R. D. & Beck D. Spatial attention deficits in humans: A
525 comparison of superior parietal and temporal-parietal junction lesions.
526 Neuropsychology 12, 193-207 (1998).

527 38. Corbetta M. & Shulmann G. L. Control of goal-directed and stimulus-driven attention
528 in the brain. Nat Rev Neurosci. 3, 201-15 (2002).

529 39. Göschl F., Friese U., Daume J., König P. & Engel A. K. Oscillatory signatures of
530 crossmodal congruence effects: An EEG investigation employing a visuotactile pattern
531 matching paradigm. Neuroimage 116, 177-186 (2015).

532 40. Romero Y. R., Keil J., Balz J., Niedeggen M., Gallinat J., & Senkowski D. Alpha-Band
533 Oscillations Reflect Altered Multisensory Processing of the McGurk Illusion in
534 Schizophrenia. Front Hum Neurosci. 10, 41 (2016).

535 41. Wang W., Hu L., Cui H., Xie X. & Hu Y. Spatio-temporal measures of
536 electrophysiological correlates for behavioral multisensory enhancement during visual,
537 auditory and somatosensory stimulation: A behavioral and ERP study. Neurosci Bull.
538 29, 715–724 (2013).

539 42. Peng W., Hu L., Zhang Z. & Hu Y. Causality in the association between P300 and Alpha
540 event-related desynchronization. PLoS One 7, e34163 (2012).

541 43 Womelsdorf T., Johnston K., Vinck M. & Everling S. Theta-activity in anterior
542 cingulate cortex predicts task rules and their adjustments following errors. Proc
543 NatlAcad Sci USA 107, 5248-53 (2010).

544 44. Menon V. & Uddin L. Q. Saliency, switching, attention and control: a network model
545 of insula function. Brain Struct Funct. 214, 655-667 (2010).

546 45. Delorme A, & Makeig S. EEGLAB: an open source toolbox for analysis of single-trial
547 EEG dynamics. J Neurosci Methods. 134, 9-21 (2004).

548 46. Hipp J. & Siegel M. Dissociating neuronal gamma-band activity from cranial and ocular
549 muscle activity in EEG. Front Hum Neurosci. 7, 338 (2013).

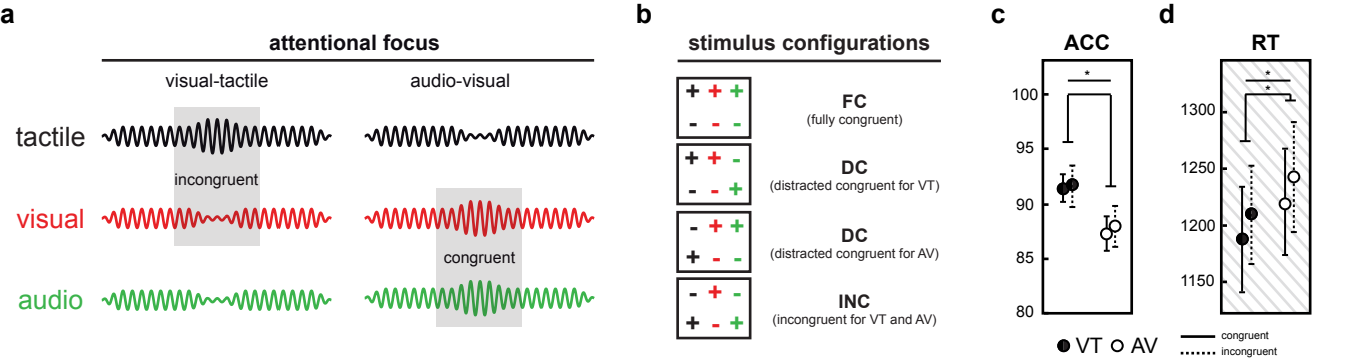
550 47. Hassler U., Trujillo-Barreto N. & Gruber T. Induced gamma band responses in human
551 EEG after the control of miniature saccadic artifacts. Neuroimage 57, 1411-1421
552 (2011).

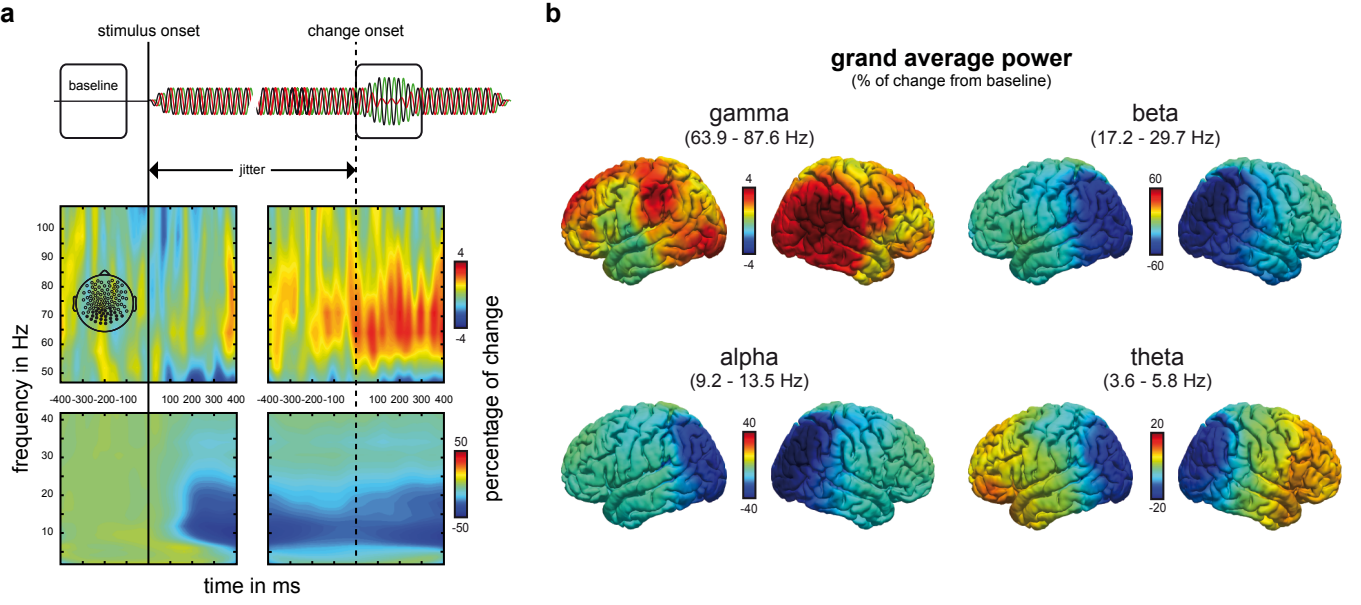
553 48. Pascual-Marqui R. D., Lehmann D., Koukkou M., Kochi K., Anderer P., Saletu B. et al.
554 Assessing interactions in the brain with exact low-resolution electromagnetic
555 tomography. Philos Trans A Math Phys Eng Sci. 369, 3768-84 (2011).

556 49. Nolte G. & Dassios G. Analytic expansion of the EEG lead field for realistic volume
557 conductors. Phys Med Biol. 50, 3807-23 (2005).

558 50. Desikan R. S., Segonne F., Fischl B., Quinn B. T., Dickerson B. C., Blacker D. et al. An
559 automated labeling system for subdividing the human cerebral cortex on MRI scans into
560 gyral based regions of interest. Neuroimage 31, 968-980 (2006).

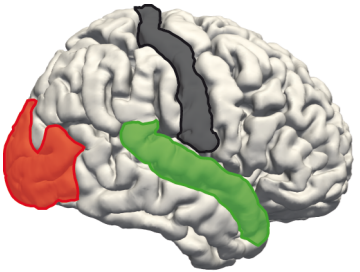
561 51. Maris E. & Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. J
562 Neurosci Methods 164, 177-90 (2007).


a
attentional focus
visual-tactile
audio-visual
tactile
visual
audio
incongruent
congruent
b
stimulus configurations
FC
(fully congruent)
DC
(distracted congruent for VT)
DC
(distracted congruent for AV)
INC
(incongruent for VT and AV)
c
ACC
100
95
90
85
80
VT
AV
d
RT
1300
1250
1200
1150
congruent
incongruent


a
stimulus onset
change onset
baseline
jitter
frequency in Hz
time in ms
percentage of change
b
grand average power
(% of change from baseline)
gamma
(63.9 - 87.6 Hz)
beta
(17.2 - 29.7 Hz)
alpha
(9.2 - 13.5 Hz)
theta
(3.6 - 5.8 Hz)

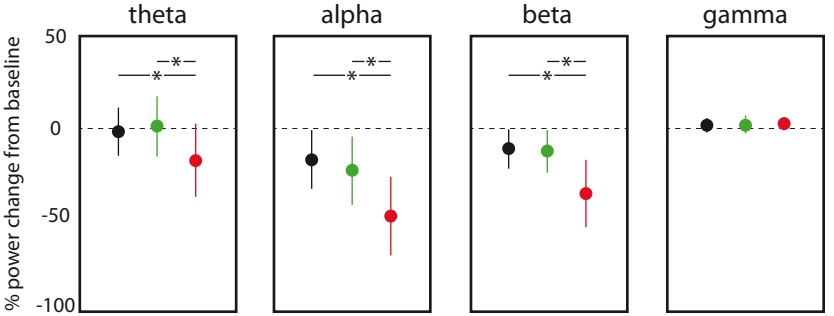
**a**

regions of interest (ROIs)



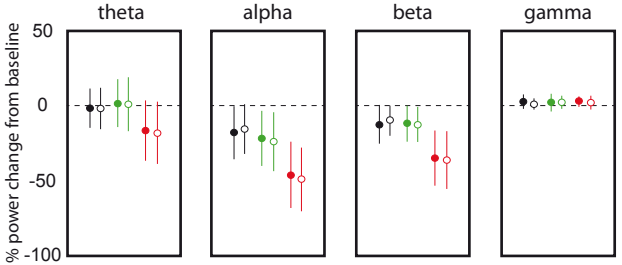
**b**

effect of *ROI*




**c**

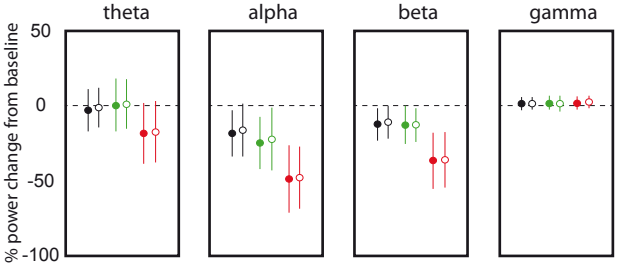
no effect of *ATTENTION*



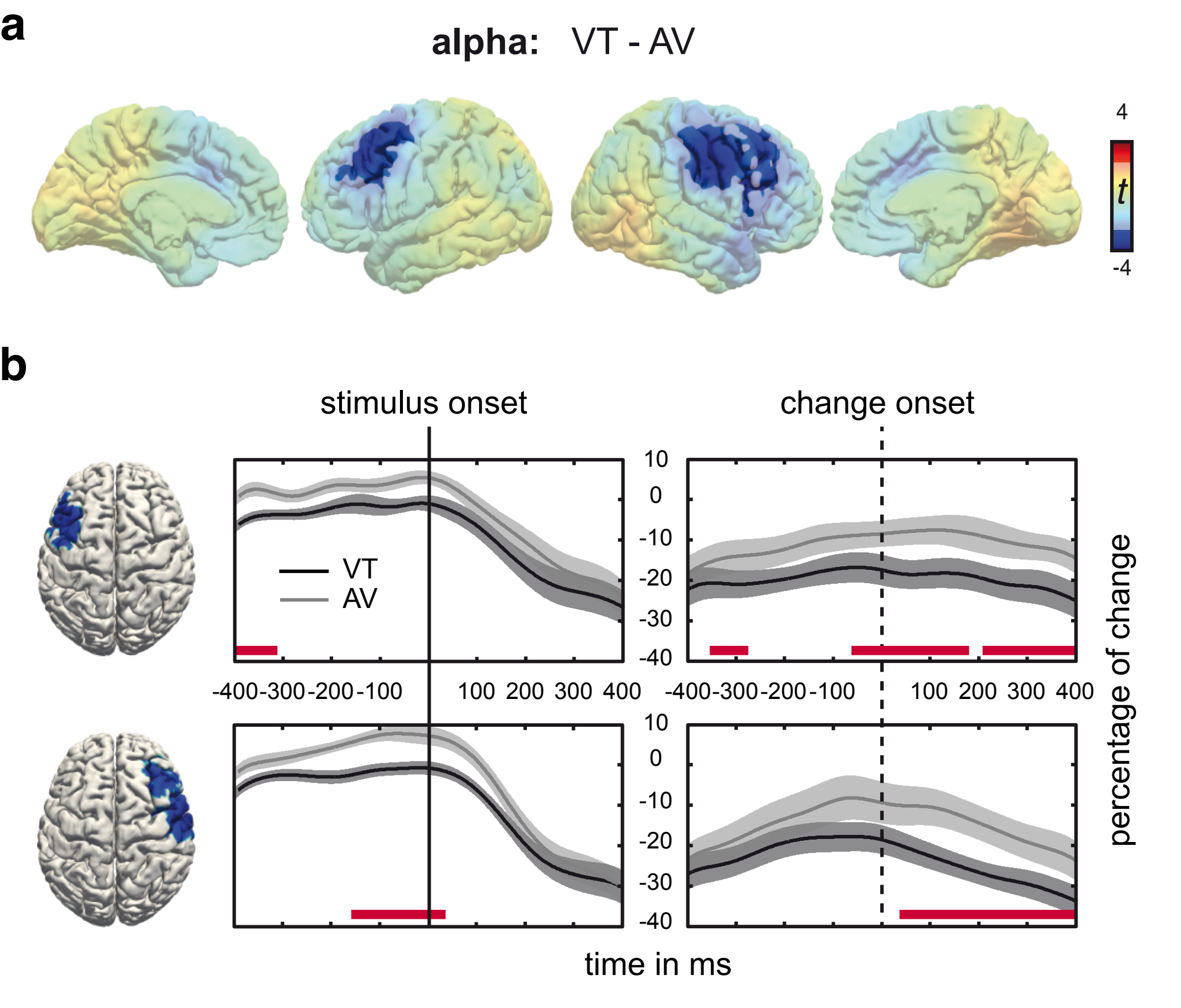

**VT:** filled circles **AV:** empty circles

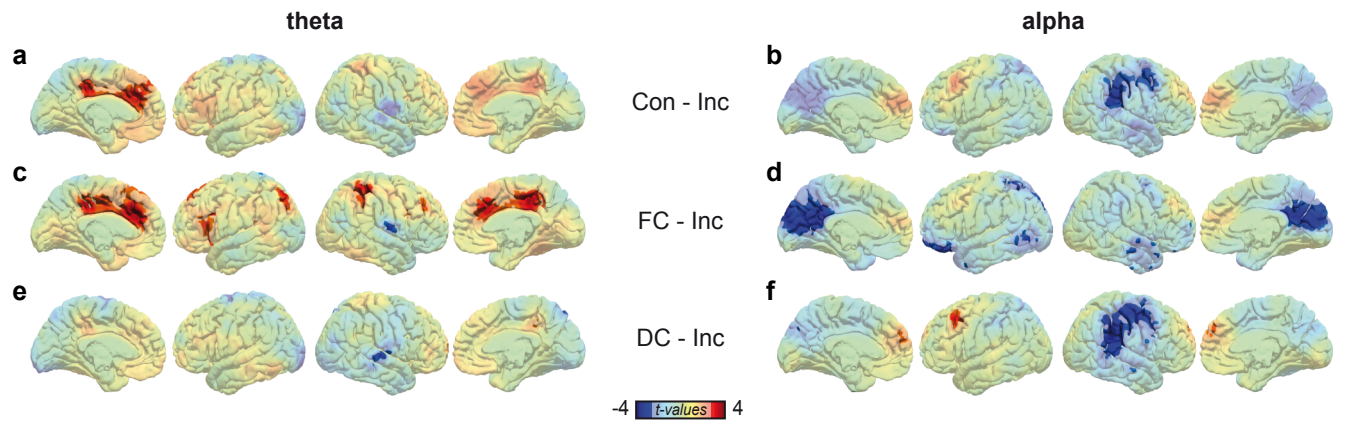
**d**

no effect of *CONGRUENCE*




**Congruent:** filled circles **Incongruent:** empty circles


a
alpha: VT - AV
4
t
-4
b
stimulus onset
change onset
VT
AV
10
0
-10
-20
-30
-40
-400-300-200-100
100 200 300 400
percentage of change
time in ms


theta
alpha
a
b
Con - Inc
c
d
FC - Inc
e
f
DC - Inc
-4
t-values
4