

Serveur Académique Lausannois SERVAL serval.unil.ch

Author Manuscript

Faculty of Biology and Medicine Publication

This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Published in final edited form as:

Title: Sounds enhance visual completion processes.

Authors: Tivadar RI, Retsa C, Turoman N, Matusz PJ, Murray MM

Journal: NeuroImage

Year: 2018 Jun 26

DOI: [10.1016/j.neuroimage.2018.06.070](https://doi.org/10.1016/j.neuroimage.2018.06.070)

In the absence of a copyright statement, users should assume that standard copyright protection applies, unless the article contains an explicit statement to the contrary. In case of doubt, contact the journal publisher to verify the copyright status of an article.

Sounds Enhance Visual Completion Processes

**Ruxandra I. Tivadar^{1,2}, Chrysa Retsa¹, Nora Turoman¹, Pawel J. Matusz^{1,3},
and Micah M. Murray^{1,2,4,5,*}**

¹The LINE (Laboratory for Investigative Neurophysiology), Department of Radiology, University Hospital Center and University of Lausanne, 1011 Lausanne, Switzerland

²Department of Ophthalmology, University of Lausanne and Fondation Asile des Aveugles, 1003 Lausanne, Switzerland

³Information Systems Institute at the University of Applied Sciences Western Switzerland (HES-SO Valais), 3960 Sierre, Switzerland

⁴The EEG Brain Mapping Core, Center for Biomedical Imaging (CIBM), University Hospital Center and University of Lausanne, 1011 Lausanne, Switzerland

⁵Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN 37203-5721, USA

*Address correspondence to:

Prof. Micah Murray

Department of Radiology

University Hospital Center

Rue du Bugnon 46, RC07.04.021

1011 Lausanne Switzerland

Tel: +41 21 314 1321

Email: micah.murray@chuv.ch

Abstract

Everyday vision includes the detection of stimuli, figure-ground segregation, as well as object localization and recognition. Such processes must often surmount impoverished or noisy conditions; borders are perceived despite occlusion or absent contrast gradients. These illusory contours (ICs) are an example of so-called mid-level vision, with an event-related potential (ERP) correlate at ~100-150ms post-stimulus onset and originating within lateral-occipital cortices (the IC_{effect}). Presently, visual completion processes supporting IC perception are considered exclusively visual; any influence from other sensory modalities is currently unknown. It is now well-established that multisensory processes can influence both low-level vision (e.g. detection) as well as higher-level object recognition. By contrast, it is unknown if mid-level vision exhibits multisensory benefits and, if so, through what mechanisms. We hypothesized that sounds would impact the IC_{effect}. We recorded 128-channel ERPs from 17 healthy, sighted participants who viewed ICs or no-contour (NC) counterparts either in the presence or absence of task-irrelevant sounds. The IC_{effect} was enhanced by sounds and resulted in the recruitment of distinct configuration of active brain areas over the 70-170ms post-stimulus period. IC-related source-level activity within the lateral occipital cortex (LOC), inferior parietal lobe (IPL), as well as primary visual cortex (V1) were enhanced by sounds. Moreover, the activity in these regions was correlated when sounds were present, but not when absent. Results from a control experiment, which employed amodal variants of the stimuli, suggested that sounds impact the perceived brightness of the IC rather than shape formation per se. We provide the first demonstration that multisensory processes augment mid-level vision and everyday visual completion processes, and that one of the mechanisms is brightness enhancement. These results have important implications for the design of treatments and/or visual aids for low-vision patients.

Introduction

Illusory contours entail the perception of borders in regions of homogeneous contrast gradients (Leshner, 1995). The brain thus perceptually “fills in” missing borders. Illusory contours thus provide access to mid-level vision mechanisms that bridge sensation and perception (Marr, 1982), and have been used in laboratory settings to study processes of figure-ground segregation and border completion. Currently, illusory contour processes are considered exclusively visual (Leshner, 1995; Murray and Herrmann, 2013; Nieder and Wagner, 1999; Seghier and Vuilleumier, 2006); any influence from other sensory modalities is currently unknown.

The mechanisms of IC processes have been a subject of much debate (Murray and Herrmann, 2013). We and others have demonstrated that the human brain first exhibits sensitivity to the presence of ICs ~90ms post-stimulus onset via enhanced activity within the lateral occipital cortices; followed by feedback-mediated activity in V1/V2 (Knebel and Murray, 2012; Mendola et al., 1999; Murray et al., 2002; Pegna et al., 2002; Poscoliero and Girelli, 2017). Moreover, the same brain mechanism governs illusory contour processes induced by wide variations in low-level stimulus features (Murray and Herrmann, 2013) as well as when resulting in either modal or amodal completion (i.e. both with and without concomitant brightness enhancement of the IC; (Murray et al., 2004)). Sensitivity to illusory contours appears to be distinct from shape discrimination processes (Doniger et al., 2000; Murray et al., 2006). Additionally, these mechanisms seem to be conserved across species, from owls to cats and primates (Murray and Herrmann, 2013), and are present early in development (in humans within 6-8 months of age; (Csibra et al., 2000)). IC processes thus seem to be robust phenomena of visual perception.

Nevertheless, there are instances where altered early-life experience can result in a breakdown of mid-level vision, including illusory contour sensitivity, even when vision is, in principle, surgically restored. This has recently been demonstrated in the case of children with bilateral congenital cataracts that were removed only at the age of ~9–11 years (McKyton et al., 2015). Such children, while able to perform low-level visual tasks such as color, size or shape discrimination, were nonetheless severely impaired on tasks requiring recognition of illusory contours or occlusion. Ameliorating such mid-level visual processes could potentially be possible with the aid of multisensory information, as multisensory benefits in perception,

cognition and behavior have been widely documented (Matusz et al., 2017; Murray et al., 2016b; Murray and Wallace, 2012; Stein and Meredith, 1993).

Mechanistically, the earliest stages of visual cortical processing are impacted by auditory information (Murray et al., 2016b). For example, the perceived brightness of visual stimuli can be increased when presented together with sounds (Marks, 1993; Noesselt et al., 2010; Stein et al., 1996). Similarly, the probability of perceiving a phosphene (as induced by single-pulse TMS over the occipital pole) can be substantially increased by presentation of sounds (Romei et al., 2013, 2009, 2007; Spierer et al., 2013). Even the number of perceived flashes can be dictated by the number of surrounding sounds (Shams et al., 2000). Finally, auditory stimuli presented alone can also modulate responsiveness of nominally “visual” cortices, including primary visual cortex (Matusz et al., 2016; Mercier et al., 2013; Schroeder and Foxe, 2002). These collective findings raise the possibility for multisensory information to impact the processing not only of physically present visual stimuli, but also illusory or perceptually completed visual information, such as illusory contours. One prior study indeed tested this possibility, by focusing primarily on the processing of the auditory stimulus (Fiebelkorn et al., 2010). They reported that illusory contour processes were not significantly affected by the presence or absence of a task-irrelevant tone. Instead, these authors proposed a sequential progression from the establishment of objects’ visuo-spatial boundaries to the cross-modal influence of attention.

Establishing that there are benefits of concomitant sounds on real-world processes, such as visual shape completion, would provide important novel evidence for the added value of multisensory processes in supporting (mid-level) vision rehabilitation. Due to the bottom-up nature of early multisensory processes (De Meo et al., 2015; ten Oever et al., 2016), the presentation of a sound simultaneously with illusory contour stimuli is predicted in the present study to enhance visual shape completion processes at a neurophysiologic level. We likewise tested whether the effects of sound on shape completion can extend beyond visual brightness enhancements, by including situations of both modal (involving perceived brightness enhancement concomitant to illusory contour perception) and amodal (no perceived brightness enhancement) completion.

Methods

Subjects

We tested twenty-one right-handed adults (women: 3, age range 21-39, mean age: 25.76, standard deviation of age: 4.54). All participants had normal or corrected-to-normal vision, and reported normal hearing capacities. No participant had a history of or current neurological or psychiatric illness. Data from four participants were excluded due to excessive EEG artifacts, technical issues during data acquisition, or failure to complete the task, thus leaving 17 in the final sample (15 men; aged 21–39).

Stimuli

Stimuli consisted of a set of 4 circular Kanizsa-type (Kanizsa, 1976) ‘pacman’ inducers that were arranged to either form an illusory contour or not (IC and NC conditions, respectively) (**Figure 1**). Inducers appeared black on a dark gray background (1.077 lux at a viewing distance of 80cm for all stimuli) and were presented for 100ms on any given trial. The IC shapes (and their equivalent NCs) were squares, resulting from the inducers being placed along the diagonals. The eccentricity from fixation to the center of an inducer was 2.5° of visual angle (5.31° center-to-center eccentricity for corresponding NCs). The NC conditions entailed 5 variations in the rotations of the inducers for each the circle and square ICs. These variations were included to prevent participants from selectively attending to particular regions of space as a strategy to successfully complete the task. ICs, when present, had a support ratio of 0.40, indicating that 40% of the shape’s borders were physically present as a contrast gradient. On half of the trials, a sound (1000Hz sinusoidal pure tone; 100ms duration; 44.1kHz sampling; 10ms fade-in/out to remove clicks) was synchronously presented with the array of inducers. A sound’s occurrence was randomized across trials. Sounds were presented via insert headphones (Etymotic model ER-4P; www.etymotic.com) at a sample rate of 48kHz, and the sound volume was adjusted to a level comfortable for each participant (53.3±0.2dB SPL as measured at the headphone using a CESVA SC-L sound pressure meter). Identical stimuli were presented in the amodal condition, except for the inclusion of a black outline encircling each inducer. The employed modal and amodal forms have been used in prior studies from our group, and are known to result in similar IC sensitivity (Murray et al., 2004). Each block of trials consisted of 240 visual stimuli with equal probability of IC, NC, modal and amodal, sound or no sound condition. This complete randomization of our experimental design controlled for

order effects, response bias and prevented participants from being able to predict the onset of IC/NC or modal/amodal conditions based on the presence of the sound. Therefore, the sound could not have acted as a cue or warning signal for the presence or absence of an illusory contour.

Procedure and Task

Participants sat in a sound-attenuated darkened room (WhisperRoom MDL 102126E). A black central fixation dot remained on the computer screen throughout each block of trials. Stimuli were presented for 100ms duration. The participant's task was a two-alternative forced choice that required the discrimination between IC and NC presence on each trial using one hand for IC responses and the other for NC responses on a five-button SR-box (Psychology Software Tools, Inc.). Participants were instructed to respond as quickly and as accurately as possible. After responding, the next trial was initiated and was preceded by an inter-trial interval randomly ranging between 500ms and 1000ms. Each participant completed eight blocks of trials, making 1920 trials in total per participant (i.e. 240 per condition) During the experiment, participants took regular breaks between blocks of trials to maintain high concentration and prevent fatigue. Stimulus delivery and behavioral response collection were controlled by E-Prime software (Psychology Software Tools, Inc.). The accuracy of stimulus timing and synchrony was confirmed with an oscilloscope.

Behavioural Analysis

First, we excluded all trials with reaction times (RTs) shorter than 200 msec. We then excluded any remaining outliers on a single subject basis (i.e. for each subject and condition), applying a mean \pm 2 standard deviations criterion (Ratcliff, 1993; Field et al., 2012). On average, 2.7% of the trials were excluded from any condition. Accuracy was analyzed based on signal detection theory measures of sensitivity (d') and criterion (c) (Macmillan and Creelman, 2005). Hits were correctly identified ICs. False alarms were NC trials on which a participant reported the presence of an illusory contour. We compared d' and c with a 2x2 ANOVA with factors Inducer type (modal/amodal) and Sound (present/absent). RT data were analyzed with a 2x2x2 ANOVA with factors Inducer type (modal/amodal), Condition (IC/NC), and Sound (present/absent).

EEG acquisition and pre-processing

Continuous EEG was recorded at 1024Hz with a 128-channel BioSemi ActiveTwo AD-box (<http://biosemi.com>), referenced to the common mode sense (CMS, active electrode) and grounded to the driven right leg (DRL, passive electrode), which functions as a feedback loop driving the average potential across the electrode montage to the amplifier zero. Cartool (available at <http://www.fbmlab.com/cartool-software/>) (Brunet et al., 2011) was used for data pre-processing and statistical analyses. Prior to averaging, the EEG was filtered with a 2nd order Butterworth filter (-12dB/octave roll-off; 0.1Hz high-pass; 60Hz low-pass; 50Hz notch). The filters were computed linearly in both forward and backward directions to eliminate phase shifts. The continuous EEG was then segmented into peri-stimulus epochs spanning from 100ms pre-stimulus to 500ms post-stimulus onset (baseline was the -100ms to 0ms interval). Data quality was then controlled with a semi-automated artefact rejection criterion of $\pm 100\mu\text{V}$ at each channel as well as visual inspection to exclude any remaining transient noise, muscle artefacts and eye-movements (on average and as detailed below, 85% of trials were accepted per subject). Data from artefact-contaminated electrodes were interpolated (mean: 9.1, range: 4–12 channels) using three-dimensional splines (Perrin et al., 1987). Data were then baseline corrected and an average reference was applied before single trials were averaged to obtain the event-related potentials (ERPs). Given the near-ceiling performance of participants, no exclusion of EEG epochs based on behavioral accuracy was applied. ERPs were generated for each participant in response to the following 8 conditions (mean \pm s.e.m. of accepted epochs also indicated): modal ICs with sounds (MIC_snd; 198 ± 40), modal NCs with sounds (MNC_snd; 196 ± 43), modal ICs without sounds (MIC_nosnd; 202 ± 37), modal NCs without sounds (MNC_nosnd; 197 ± 42), amodal ICs with sounds (AIC_snd; 198 ± 41), amodal NCs with sounds (ANC_snd; 198 ± 43), amodal ICs without sounds (AIC_nosnd; 200 ± 38), and amodal NCs without sounds (ANC_nosnd; 195 ± 42).

Differences were then calculated between the ERPs to the various IC and NC conditions in order to isolate brain activity associated with IC sensitivity (i.e. the IC_{effect} introduced by (Murray et al., 2002)) and to exclude that any effects of sound could be explained by multisensory interactions between the inducer stimuli and sounds. Hereafter, we refer to these differences for the modal inducers with sounds as $IC_{\text{effect_MSND}}$ (i.e. MIC_snd minus MNC_snd), for the modal inducers without sounds as $IC_{\text{effect_MNOSND}}$ (i.e. MIC_nosnd minus MNC_nosnd), for the amodal inducers with sounds as $IC_{\text{effect_ASND}}$ (i.e.

AIC_snd minus ANC_snd), and for the amodal inducers without sounds as IC_{effect}_ANOSND (AIC_nosnd minus ANC_nosnd). These ERPs were first analyzed following a 2×2 within-subject design. There were exclusively main effects of stimulus type (modal vs. amodal) starting at ~150ms post-stimulus onset and sound presence/absence over the ~100-150ms post-stimulus interval. Consequently, we focus the remainder of our analyses on the separate analysis of the IC_{effect} for modal and amodal inducer types, comparing responses when sounds were present vs. absent. Responses to modal and amodal stimuli were analysed separately to determine if influences of sounds on illusory contour processes were linked to the perceived brightness enhancement of the illusory contour, which occurs with modal but not amodal stimuli.

ERP analyses

We focused our analyses here on the initial 300ms post-stimulus onset, given our prior studies of the time course of visual completion processes (Murray and Herrmann, 2013) as well as multisensory processes (Murray et al., 2016a). ERPs were analyzed within an electrical neuroimaging framework that uses reference-independent measures of the electric field at the scalp (Koenig et al., 2014; Michel et al., 2004; Michel and Murray, 2012; Murray et al., 2008). Voltage waveforms were analyzed using univariate as well as multivariate analyses as a function of time. These analyses disentangle effects arising from either modulations in the strength of responses, alterations in the configuration of the active generators (inferred from the topography of the electric field at the scalp), or latency shifts in brain processes across experimental conditions.

The analyses used reference-free methods. Specifically, they entailed determining ERP strength with global field power (GFP), as well as ERP topographic differences via global map dissimilarity (GMD) (e.g. (Brunet et al., 2011; Murray et al., 2008)). GFP is computed as the root mean square of voltages across the whole electrode montage, whereas GMD is the root mean square of the squared difference between two GFP-normalized vectors across two conditions (here, the 128-channel EEG; (Lehmann and Skrandies, 1980)). GMD can range between 0 and 2, with 0 indicating no topographical differences and 2 indicating topographical inversion. GMD is also directly related to the Pearson product-moment correlation (i.e. spatial correlation equals $[(2-GMD^2)/2]$; (Brunet et al., 2011)). We plot the results in terms of spatial correlation, which ranges from -1 to 1, rather than GMD to facilitate legibility for readers less

accustomed to GMD. Topographic differences between conditions indicate a difference in the configuration of the underlying neural generators (Lehmann, 1987). GFP and GMD are orthogonal measures of distinct features of ERPs and can thus be analyzed independently. Statistical analyses of voltage waveforms, GFP, and GMD entailed non-parametric randomization tests. For analyses of voltage waveforms, significant differences were defined as those meeting both a temporal criterion of more than 15 contiguous time-frames (Guthrie and Buchwald, 1991) as well as a spatial criterion of more than 10% of the electrode montage (Matusz et al., 2015). For analyses of GFP and GMD a temporal criterion of more than 10 contiguous time-frames was applied (Sarmiento et al., 2016).

Next, to identify stable electric field topographies (i.e. ‘template maps’) and differences therein across the sound and no-sound conditions, we performed a hierarchical cluster analysis on the group-averaged ERPs concatenated across conditions (Murray et al., 2008). In this analysis, the group-averaged data from each condition are first normalized by their instantaneous GFP, which makes the clustering exclusively sensitive to topographic modulations. The application of a modified Krzanowski-Lai criterion (Murray et al., 2008) identifies the optimal number of temporally stable ERP clusters, i.e. the minimal number of stable maps accounting for the most amount of variance in the concatenated group-averaged data across conditions. We hereafter refer to these stable maps as template maps, because they are based on the group-averaged data and constitute a “template” pattern of ERP topographies as a function of time for each condition. The clustering makes no assumption regarding the orthogonality of the derived template maps (Koenig et al., 2014). Subsequently, we submitted a subset of the template maps derived from the group-average ERPs to a fitting procedure, wherein each time point of the single-subject ERP over a specified period of time based on the above hierarchical clustering is labelled according to the template maps that it best correlates with spatially in a winner takes all fashion (Murray et al., 2008). As an output, for each subject and condition, one obtains the number of time samples that a given template map is better correlated with spatially (see schematic of the clustering and fitting in **Supplementary Figure S1**). We then statistically tested the relative presence (in milliseconds) of each of these template maps in the instantaneous scalp topography of the ERP IC difference-waveform across the IC_{effect_MSND} and IC_{effect_MNOSND} conditions using a Wilcoxon signed rank test given that only 2 template maps were used in the fitting (see Results).

Source estimations

To estimate the brain generators of the electrical activity recorded at the scalp level, we used a distributed linear inverse solution (minimum norm) combined with the LAURA (local autoregressive average) regularization approach. This approach uses biophysical laws as constraints (Grave de Peralta Menendez et al., 2004, 2001) and selects the source configuration that mimics the biophysical behavior of electric vector fields (i.e. activity at one solution point depends on the activity at neighboring points, according to electromagnetic laws). We used homogeneous regression coefficients in all directions and within the whole solution space. The solution space consisted of 4024 nodes, selected from a 6 x 6 x 6 grid equally distributed within the gray matter of the Montreal Neurological Institute's average brain model (<http://www.electrical-neuroimaging.ch>). The head model and lead field matrix were generated within the Spherical Model with Anatomical Constraints (SMAC; (Spinelli et al., 2000)), as implemented in Cartool. An issue of spurious (i.e. "ghost") sources may arise due to the fact that LAURA is a distributed source model. However, simulations and evaluations of empirical data (Michel et al., 2004), and discussions on this subject (Grave de Peralta Menendez et al., 2004; Martuzzi et al., 2009; Michel et al., 2004), have concluded that the localization precision approximately follows the grid size (i.e. 6mm). Moreover, we argue that by averaging source estimations across subjects and conditions and comparing these to each other, the likelihood of falsely accepting "ghost" sources is minimized, as it is less probable that such a source is consistently observed across individuals and conditions. The results of the described topographic analysis defined time periods for which intracranial sources were estimated and statistically compared across conditions. The ERP data were down-sampled and affine-transformed to a 111-channel montage prior to calculation of the inverse solutions. Source estimations were calculated after first averaging across time (i.e. 70-170ms post-stimulus onset, see Results section) for each subject and condition. LAURA provides current density measures as an output, and we statistically contrasted (paired t-test) their scalar values at each solution point. In addition to the 0.05 significance threshold at any given solution point, we also applied a spatial-extent criterion of >15 continuous solution points (see also (Knebel and Murray, 2012; Toepel et al., 2009) for a similar approach based on random field theory).

Correlation between estimated sources

A measure of non-directed interaction (Bastos and Schoffelen, 2016) was based on correlations between inter-regional source estimations at loci exhibiting significant differences in illusory contour sensitivity as a function of sound presence. As will be detailed below in the Results, three loci exhibited such differences: primary visual cortex (V1), lateral-occipital cortex (LOC), and parietal cortex (PC). Source estimation values at the maximum point (i.e. point with the smallest p-value) of each of these loci were extracted for each participant for the sound and no sound conditions ($IC_{\text{effect_MSND}}$ and $IC_{\text{effect_MNOSND}}$). We then calculated the Pearson's correlation coefficient, which is a measure of covariance, after first establishing the normality of the distribution of values across participants. The Holm-Bonferroni correction was applied to correct for multiple comparisons (Holm, 1979).

Results

Behavioral results

Mean accuracy for each condition was above 90% correct (range across conditions: 90%-98%), suggesting that participants could easily perform the task at near-ceiling levels. Values of d' and c were each submitted to a 2x2 ANOVA using within-subject factors of Inducer type (modal vs. amodal) and Sound (presence vs. absence) (**Supplementary Figure S2a-b**). For d' , which across all conditions was >3.4 , neither main effect was statistically reliable (both p 's >0.12) nor was the interaction between these factors ($p>0.37$). For c , there was a main effect of Inducer type ($F_{(1,16)}=12.43$; $p=0.003$; $\eta_p^2=0.44$), with a generally larger bias to respond "IC present" with amodal than with modal inducers. By contrast, neither the main effect of Sound ($p=0.06$) nor the interaction between factors was statistically reliable ($p=0.37$). The 2x2x2 ANOVA for RTs entailed within subject factors of Inducer type (modal vs. amodal), Condition (IC vs. NC) and Sound (presence vs. absence) (**Supplementary Figure S2c**). Each main effect was statistically reliable (all p 's <0.001), and there was also a significant stimulus type \times contour presence interaction ($p<0.002$). No other interaction was reliable (all p 's >0.26). Overall, no behavioural measure provided evidence that sounds differentially affected IC vs. NC processing.

ERP results

The ERP analysis focused on determining the impact of sound on illusory shape completion (i.e. the IC_{effect} with and without presentation of sounds). For illustrative purposes we display unsubtracted ERP waveforms in response to modal inducers from an exemplar posterior electrode site (**Figure 2**). Here, it can be observed that responses to IC stimuli were enhanced by sound presentation. For all analyses, we isolated IC-specific activity by subtracting the ERPs in the NC condition from ERPs in the IC condition, which ensured that we were only measuring the effects of sound vs. no sound on illusory contour sensitivity. This subtraction is important, as it excludes all effects of sound on the processing of the ‘pac-man’ inducers themselves, independently of their orientation to form or not an illusory contour. **Figure 3A** displays the $IC_{\text{effect_MSND}}$ and $IC_{\text{effect_MNOSND}}$ waveforms at an exemplar electrode site. **Figure 3B** displays the percentage of the electrode montage exhibiting a significant difference as a function of time (based on a non-parametric randomization test). Reliable effects were first observed over the 64-88ms post-stimulus period, with effects also evident over the 142-173ms and 203-231ms periods. There was also a short-lived (<15 contiguous time frames) difference observed during the baseline. However, we would remind readers of the reference-dependent nature of analyses of voltage waveforms. Also, we would emphasize that baseline differences were not observed in analyses of global measures of the electric field at the scalp. No reliable differences in response strength were observed (i.e. GFP; see **Supplementary Figure S3**). By contrast, **Figure 3C** displays the spatial correlation between these responses as a function of time. Reliable topographic differences were observed over the 68-93ms post-stimulus period. The group-averaged data were then submitted to topographic clustering in order to better characterize the basis for these topographic differences. A total of 9 template maps, explaining 90.1% of the total global variance in the ERPs was observed. To iterate, a template map refers to a stable ERP topography observed in the group-average data. In general, the same pattern of maps characterized the $IC_{\text{effect_MSND}}$ and $IC_{\text{effect_MNOSND}}$ responses with the exception of the 70-170ms post-stimulus time-period (**Supplementary Figure S4**). Here, two template maps were observed (blue- and green-framed maps in **Figure 4A**). This pattern was then statistically assessed by calculating the spatial correlation of each of these two template maps with the responses from individual participants from both $IC_{\text{effect_MSND}}$ and $IC_{\text{effect_MNOSND}}$ conditions over the 70-170ms period. This yields the amount of time each template map had a higher spatial correlation with each condition (**Figure 4B**). These values were then submitted to a set of Wilcoxon signed rank

tests. The manner in which each template map characterized responses in each condition significantly differed ($p < 0.001$). Specifically, one map predominantly characterized responses to the $IC_{\text{effect_MNOSND}}$ condition ($p < 0.001$), whereas both template maps equally characterized responses to the $IC_{\text{effect_MSND}}$ condition ($p = 0.868$).

LAURA source estimates were calculated to provide an estimate of the likely sources of brain activity over this time window. Three clusters were identified from the statistical contrast of source estimations from the $IC_{\text{effect_MSND}}$ and $IC_{\text{effect_MNOSND}}$ conditions (**Figure 5A**). These clusters were located in the right primary visual cortex (V1), right lateral occipital cortex (LOC), and the right inferior parietal lobe (IPL). The Talairach and Tournoux (Talairach and Tournoux, 1988) coordinates of local statistical maxima were respectively (5, -81, 5mm), (47, -69, 16mm), and (29, -68, 38mm). In a final step, we calculated the Pearson correlation between source estimations in the local maxima within these clusters separately for the $IC_{\text{effect_MSND}}$ and $IC_{\text{effect_MNOSND}}$ conditions, in order to provide a measure of non-directed interaction between activities across the clusters in the presence vs. absence of sounds. Source estimation magnitudes were significantly correlated across these clusters for the $IC_{\text{effect_MSND}}$ condition after Holm-Bonferroni correction for multiple comparisons, but not for the $IC_{\text{effect_MNOSND}}$ condition (**Figure 5B**). However, these correlation coefficients did not themselves reliably differ when tested using Fisher's r -to- z transformation (LOC-IPL: $z = 0.46$; $p = 0.32$; IPL-V1: $z = 1.1$; $p = 0.14$; LOC-V1: $z = 1.16$; $p = 0.12$; all tests 1-tailed).

With modal completion stimuli, it cannot be discerned if sounds are impacting either or both form completion as well as perceptual brightness enhancement of the bound figure. For this reason, we also presented participants with stimuli resulting in amodal completion as well as rotated inducer arrays preventing such perceptions (AIC and ANC, respectively). These stimuli result in the perception of a bound form without concomitant perceptual brightness enhancement. As above, we analyzed the difference between these conditions both when sounds were presented and not ($IC_{\text{effect_ASND}}$ and $IC_{\text{effect_ANOSND}}$ conditions, respectively). There was no evidence for reliable significant differences between these conditions at the level of ERP voltage waveforms, GFP, or GMD. A short-lived GFP difference was observed over the 90-98ms post-stimulus period, which did not meet our 10TF temporal criterion.

Discussion

We provide the first demonstration of sounds benefiting mid-level vision. Sounds enhanced the initial stages of illusory contour processes by changing the network of brain regions involved to include not only the LOC and inferior parietal lobule (as in prior visual-only studies), but also V1. Moreover, activity across this extended network was correlated when sounds were present, but not when absent. Importantly, while sounds affected ERPs to stimuli resulting in modal completion, this was not the case with stimuli resulting in amodal completion. This collective pattern of results not only extends our models of both visual functions and multisensory processes, but also opens new avenues for visual rehabilitation.

Sounds influenced the initial stages of IC processing. The timing of our effects at 70–177ms post-stimulus onset corresponds to when previous research has observed IC sensitivity, where ERP modulations appeared ~90–150ms post-stimulus onset (Anken et al., 2016; Murray and Herrmann, 2013). Aside from its timing, the localization of the $IC_{\text{effect_MNOSND}}$ is also highly consistent with previous visual-only studies implicating the LOC and IPL during early stages of IC processing. The present study extends these findings. First, we show that the ERP difference between the $IC_{\text{effect_MNOSND}}$ and $IC_{\text{effect_MSND}}$ is topographic in nature and thus follows from changes in the configuration of active intracranial sources rather than a straightforward gain modulation, which would have resulted in GFP effects. Additionally, the source estimation analyses indicate that IC sensitivity with sounds includes not only the LOC and IPL as above, but also V1. While the LOC (and to a lesser extent IPL) has been reliably implicated in many human neuroimaging studies of IC sensitivity, the involvement of V1 has been more elusive (Murray and Herrmann, 2013). Those who have observed V1 responses in humans did so using haemodynamic imaging methods (Hirsch et al., 1995; Kok et al., 2016; Larsson et al., 1999; Maertens et al., 2008; Seghier et al., 2000). The absence of temporal information therefore could not preclude that these activations were later and the result of feedback processes from higher-tier areas; a model indeed borne out in our and others' prior empirical works both in humans (Anken et al., 2016; Knebel and Murray, 2012; Murray and Herrmann, 2013) and non-human primates (Lee and Nguyen, 2001), as well as computational models (Dura-Bernal et al., 2011). The topographic clustering analyses of the present study identified temporally stable ERP topographies over the 70–170ms post-stimulus period. By extension, the implication is that the underlying source configuration was similarly stable over this time period, though we cannot unequivocally exclude the possibility that the source configuration is rapidly changing but on average

appears as temporally stable. Such notwithstanding, the present results show for the first time that a network of LOC, IPL, and V1 is *concurrently* active, and moreover in a *correlated* manner, when IC processes are facilitated by sounds and the multisensory as well as cross-modal processes they impart.

We posit that these cross-modal processes enhanced, in a largely bottom-up fashion, the excitability of neuronal populations within V1 that are responsive to the illusory contours that themselves have been claimed to be the locus of IC sensitivity and depend on feedback inputs from LOC (and IPL). In order to isolate cross-modal effects on the neuronal representation of the illusory contour (and by extension mid-level vision), it was therefore necessary to exclude any multisensory influences on the responses to the inducers themselves; hence our focus on the IC_{effect} waveforms. Operationally, we consider as multisensory those processes between responses to physically-presented external stimuli (e.g. sounds enhancing perceived brightness; (Stein et al., 1996)), and as cross-modal those processes whereby sounds impact the responses within nominally visual cortices even in the absence of external visual inputs. However, the distinction between these operational definitions is neither absolute nor mutually exclusive. Such notwithstanding, the effects presented here are thus distinct from effects characterized by Stein et al. (Stein et al., 1996), as well as many others, where multisensory processes result in perceived brightness enhancement of physically-presented visual stimuli. Such effects are consistent with a cross-modal influence on low-level vision. Rather, we here show for the first time that cross-modal processes impact visual completion of contours in the absence of a luminance gradient. It is now well established that multisensory processes occur both early in time and within low-level visual cortices in species ranging from mice to humans (Campus et al., 2017; De Meo et al., 2015; Ghazanfar and Schroeder, 2006; Lakatos et al., 2008; Meijer et al., 2017; Murray et al., 2016b). What we instead demonstrate for the first time here is the cross-modal influence of sound on visual completion and figure-ground segregation – i.e. on mid-level visual functions. This demonstrates that a function commonly conceived as exclusively visual is instead impacted by information from another sensory modality. Importantly, our results highlight the limitations of the existing models of visual functions, as the mechanisms governing them, whether it is detection, object recognition, or mid-level vision, differ to those operating in purely visual, laboratory settings. Our finding thus call for more systematic study

of the effects and the mechanisms of mid-level vision in naturalistic, multisensory settings (see also Matusz et al., 2018).

Excitability changes in human V1 have been measured as significantly decreased thresholds for phosphene perception following single-pulse transcranial magnetic stimulation in combination with presentation of external sounds (Bolognini et al., 2010; Romei et al., 2013, 2012, 2009, 2007; Spierer et al., 2013). Sound-induced V1 excitability changes occurred when the sound preceded the TMS pulse by 75-120ms (Romei et al., 2007), thereby overlapping with the present timing of sounds enhancing the $I_{C_{effect}}$. These excitability increases follow an alpha (~10Hz) cycle that persists even after sound offset (Romei et al., 2013) and may moreover coincide with phase-resetting of ongoing alpha activity (Romei et al., 2012), which may be a neurophysiologic mechanism contributing to these excitability increases (Ohshiro et al., 2017; van Atteveldt et al., 2014). In the case of IC processing, there thus appears to be a mechanism by which sounds can put V1 neuronal ensembles in a more excitable state that enhances their responsiveness to feedback signals, originating in the LOC and IPL, which are producing illusory contours.

Increased V1 activity in response to illusory contours has been previously reported in the absence of sounds. Maertens et al. (Maertens et al., 2008) showed with functional magnetic resonance imaging (fMRI) that responses specifically increase at the retinotopic location of the perceived illusory border within V1. Studies in animal models have likewise documented increased neural firing rates within V1 (as well as V2) (Grosf et al., 1993; Redies et al., 1986; Sheth et al., 1996). However, it remains controversial whether there is a 1:1 relationship between neuronal sensitivity to luminance-defined contours and that to illusory contours (Ramsden et al., 2001) as well as the extent to which such signals rely on feedback inputs from higher-level regions (Lee and Nguyen, 2001; Murray and Herrmann, 2013). It is likewise unresolved to what extent V1 responses to ICs (in the absence of sounds) in humans are driven by top-down attentional control processes, implicated under demanding task settings. For example, Maertens et al. (Maertens et al., 2008) had participants perform a curvature discrimination on illusory contours and observed effects within V1, whereas Mendola et al. (Mendola et al., 1999) observed no such effects in V1 when requiring fixation but no fine-grained (or even explicit) discrimination. In a similar manner, Fiebelkorn et al. (Fiebelkorn et al., 2010) failed to observe effects of sounds on illusory contour processes when their participants were selectively attending to the occasional flicker of one of the inducers in

the stimulus array, which may have minimized any susceptibility of illusory contour processes to cross-modal influences. In the present study, participants discriminated the presence from absence of illusory contour shapes on a trial-by-trial basis, while sounds were completely uninformative and equally likely on any given trial. There is no evidence from our behavioral measures for sounds resulting in more demanding performance (though performance was at ceiling here and will thus need to be parametrically varied in future work). Thus, we consider an account based on top-down attentional control unlikely.

Our results are probably most exciting from the neurorehabilitation standpoint, as they suggest promising avenues for rehabilitation, and mitigation of visual impairments through multisensory regimes. For example, patients that suffer from congenital cataracts can, after cataract removal, perform as well as their age-matched controls in low-level visual discrimination (i.e. color, size, and shape discrimination). However, they continue to demonstrate long-lasting impairments of mid-level visual functions (e.g. illusory contour processing, stereoscopic depth discrimination, shading, and occlusion, (Hadad et al., 2017; McKyton et al., 2015)). As the present study has clearly demonstrated that multisensory processes can influence mid-level visual functions, it stands to reason that multisensory stimulation could provide more effective and more efficient rehabilitation in patients with impaired mid-level vision (see also Murray et al., 2015). While practical limitations remain an issue often impeding the widespread use of multisensory technologies in clinical practice (Gori et al., 2016), efforts are improving the accessibility and are already demonstrating the utility of multisensory rehabilitation in visually deprived children (Cappagli et al., 2017) as well as in adults after stroke (Tinga et al., 2016).

Our results collectively emphasize the importance of multisensory and cross-modal audiovisual processes in aiding visual functions. Our present findings significantly extend existing models of visual functions by demonstrating that sounds influence not only simple perception processes, such as detection or localization, and higher-level visual processes, such as attentional selection of visual objects, but also mid-level visual processes related to object grouping. Multisensory information might thus potentially be useful to aid the rehabilitation of these functions in the visually impaired. What should be established now is the extent to which multisensory processes can facilitate the restoration of visual functions across different levels; an exciting endeavor at the focus of our ongoing research.

Competing Interests

The authors declare no competing financial interests.

Acknowledgements

MMM is funded by the Swiss National Science Foundation (grants 320030-149982 and 320030-169206 as well as National Centre of Competence in Research project 'SYNAPSY, The Synaptic Bases of Mental Disease' [project 51NF40-158776]) as well as by a generous grantor advised by Carigest SA. PJM is funded by the Pierre Mercier Foundation as well as the Swiss National Science Foundation (grant PZ00P1_174150). MMM and PJM are both funded by a grant from the Fondation Asile des Aveugles. We thank Jean-François Knebel for his technical assistance.

References

- Anken, J., Knebel, J.-F., Crottaz-Herbette, S., Matusz, P.J., Lefebvre, J., Murray, M.M., 2016. Cue-dependent circuits for illusory contours in humans. *Neuroimage* 129. <https://doi.org/10.1016/j.neuroimage.2016.01.052>
- Bastos, A.M., Schoffelen, J.-M., 2016. A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls. *Front. Syst. Neurosci.* 9, 1–23. <https://doi.org/10.3389/fnsys.2015.00175>
- Bolognini, N., Senna, I., Maravita, A., Pascual-Leone, A., Merabet, L.B., 2010. Auditory enhancement of visual phosphene perception: the effect of temporal and spatial factors and of stimulus intensity. *Neurosci. Lett.* 477, 109–14. <https://doi.org/10.1016/j.neulet.2010.04.044>
- Brunet, D., Murray, M.M., Michel, C.M., 2011. Spatiotemporal analysis of multichannel EEG: CARTOOL. *Comput. Intell. Neurosci.* 2011, 813870. <https://doi.org/10.1155/2011/813870>
- Campus, C., Sandini, G., Concetta Morrone, M., Gori, M., 2017. Spatial localization of sound elicits early responses from occipital visual cortex in humans. *Sci. Rep.* 7, 1–12. <https://doi.org/10.1038/s41598-017-09142-z>
- Cappagli, G., Finocchietti, S., Baud-Bovy, G., Cocchi, E., Gori, M., 2017. Multisensory Rehabilitation Training Improves Spatial Perception in Totally but Not Partially Visually Deprived Children. *Front. Integr. Neurosci.* 11, 1–11. <https://doi.org/10.3389/fnint.2017.00029>
- Csibra, G., Davis, G., Spratling, M.W., Johnson, M.H., 2000. Gamma oscillations and object processing in the infant brain. *Science* 290, 1582–5.
- De Meo, R., Murray, M.M., Clarke, S., Matusz, P.J., 2015. Top-down control and early multisensory processes: chicken vs. egg. *Front. Integr. Neurosci.* 9, 1–6. <https://doi.org/10.3389/fnint.2015.00017>
- Doniger, G.M., Foxe, J.J., Murray, M.M., Higgins, B. a, Snodgrass, J.G., Schroeder, C.E., Javitt, D.C., 2000. Activation timecourse of ventral visual stream object-recognition areas: high density electrical mapping of perceptual closure processes. *J. Cogn. Neurosci.* 12, 615–21.
- Dura-Bernal, S., Wennekers, T., Denham, S.L., 2011. The role of feedback in a hierarchical model of object perception. *Adv. Exp. Med. Biol.* 718, 165–79. https://doi.org/10.1007/978-1-4614-0164-3_14
- Fiebelkorn, I.C., Foxe, J.J., Schwartz, T.H., Molholm, S., 2010. Staying within the lines: the formation of visuospatial boundaries influences multisensory feature integration. *Eur. J. Neurosci.* 31, 1737–43. <https://doi.org/10.1111/j.1460-9568.2010.07196.x>
- Ghazanfar, A. a, Schroeder, C.E., 2006. Is neocortex essentially multisensory? *Trends Cogn. Sci.* 10, 278–85. <https://doi.org/10.1016/j.tics.2006.04.008>
- Grave de Peralta Menendez, R., Gonzalez Andino, S., Lantz, G., Michel, C.M., Landis, T., 2001. Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. *Brain Topogr.* 14, 131–7.
- Grave de Peralta Menendez, R., Murray, M.M., Michel, C.M., Martuzzi, R., Gonzalez Andino, S.L., 2004. Electrical neuroimaging based on biophysical constraints. *Neuroimage* 21, 527–39. <https://doi.org/10.1016/j.neuroimage.2003.09.051>
- Grosf, D.H., Shapley, R.M., Hawken, M.J., 1993. Macaque V1 neurons can signal “illusory” contours. *Nature* 365, 550–552. <https://doi.org/10.1038/365550a0>

- Guthrie, D., Buchwald, J.S., 1991. Guthrie & Buchwald Waveform Comparison Psychophysiology 1991.pdf. *Psychophysiology* 28, 240–4.
- Hadad, B.S., Maurer, D., Lewis, T.L., 2017. The role of early visual input in the development of contour interpolation: the case of subjective contours. *Dev. Sci.* 20. <https://doi.org/10.1111/desc.12379>
- Hirsch, J., DeLaPaz, R.L., Relkin, N.R., Victor, J., Kim, K., Li, T., Borden, P., Rubin, N., Shapley, R., 1995. Illusory contours activate specific regions in human visual cortex: evidence from functional magnetic resonance imaging. *Proc. Natl. Acad. Sci. U. S. A.* 92, 6469–73.
- Holm, S., 1979. A Simple Sequentially Rejective Multiple Test Procedure. *Scand J Stat.* 6, 65–70. <https://doi.org/10.2307/4615733>
- Kanizsa, G., 1976. Subjective Contours. *Sci. Am.* 234, 48–52.
- Knebel, J.-F., Murray, M.M., 2012. Towards a resolution of conflicting models of illusory contour processing in humans. *Neuroimage* 59, 2808–17. <https://doi.org/10.1016/j.neuroimage.2011.09.031>
- Koenig, T., Stein, M., Grieder, M., Kottlow, M., 2014. A tutorial on data-driven methods for statistically assessing ERP topographies. *Brain Topogr.* 27, 72–83. <https://doi.org/10.1007/s10548-013-0310-1>
- Kok, P., Bains, L.J., Van Mourik, T., Norris, D.G., De Lange, F.P., 2016. Selective activation of the deep layers of the human primary visual cortex by top-down feedback. *Curr. Biol.* 26, 371–376. <https://doi.org/10.1016/j.cub.2015.12.038>
- Lakatos, P., Karmos, G., Mehta, A.D., Ulbert, I., Schroeder, C.E., 2008. Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–3. <https://doi.org/10.1126/science.1154735>
- Larsson, J., Amunts, K., Gulyás, B., Malikovic, a, Zilles, K., Roland, P.E., 1999. Neuronal correlates of real and illusory contour perception: functional anatomy with PET. *Eur. J. Neurosci.* 11, 4024–36.
- Lee, T.S., Nguyen, M., 2001. Dynamics of subjective contour formation in the early visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 98, 1907–1911.
- Lehmann, D., 1987. Principles of spatial analysis, in: Gevins, A.S., A, R. (Eds.), *Handbook of Electroencephalography and Clinical Neurophysiology, Vol. 1: Methods of Analysis of Brain Electrical and Magnetic Signals*. Elsevier, Amsterdam, pp. 309–354.
- Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalogr. Clin. Neurophysiol.* 48, 609–621.
- Leshner, G.W., 1995. Illusory contours: Toward a neurally based perceptual theory. *Psychon. Bull. Rev.* 2, 279–321. <https://doi.org/10.3758/BF03210970>
- Maertens, M., Pollmann, S., Hanke, M., Mildner, T., Möller, H., 2008. Retinotopic activation in response to subjective contours in primary visual cortex. *Front. Hum. Neurosci.* 2, 2. <https://doi.org/10.3389/neuro.09.002.2008>
- Marks, L.E., 1993. Contextual processing of multidimensional and unidimensional auditory stimuli. *J. Exp. Psychol. Hum. Percept. Perform.* 19, 227–49.
- Marr, D.C., 1982. *Vision*. MIT Press, Cambridge, MA.
- Martuzzi, R., Murray, M.M., Meuli, R. a, Thiran, J.-P., Maeder, P.P., Michel, C.M., Grave de Peralta Menendez, R., Gonzalez Andino, S.L., 2009. Methods for determining frequency- and region-dependent relationships between estimated LFPs and BOLD responses in humans. *J. Neurophysiol.* 101, 491–502. <https://doi.org/10.1152/jn.90335.2008>

- Matusz, P.J., Retsa, C., Murray, M.M., 2016. The context-contingent nature of cross-modal activations of the visual cortex. *Neuroimage* 125. <https://doi.org/10.1016/j.neuroimage.2015.11.016>
- Matusz, P.J., Thelen, A., Amrein, S., Geiser, E., Anken, J., Murray, M.M., 2015. The role of auditory cortices in the retrieval of single-trial auditory-visual object memories. *Eur. J. Neurosci.* 41, 699–708. <https://doi.org/10.1111/ejn.12804>
- Matusz, P.J., Wallace, M.T., Murray, M.M., 2017. A multisensory perspective on object memory. *Neuropsychologia* 105, 243–252. <https://doi.org/10.1016/j.neuropsychologia.2017.04.008>
- Matusz, P.J., Dikker, S., Huth, A., Perrodin, C., 2018. Are we ready for real-world neuroscience? *Journal of Cognitive Neuroscience*, in press.
- McKyton, A., Ben-Zion, I., Doron, R., Zohary, E., 2015. The Limits of Shape Recognition following Late Emergence from Blindness. *Curr. Biol.* 25, 2373–2378. <https://doi.org/10.1016/j.cub.2015.06.040>
- Meijer, G.T., Montijn, J.S., Pennartz, C.M.A., Lansink, C.S., 2017. Audiovisual Modulation in Mouse Primary Visual Cortex Depends on Cross-Modal Stimulus Configuration and Congruency. *J. Neurosci.* 37, 8783–8796. <https://doi.org/10.1523/JNEUROSCI.0468-17.2017>
- Mendola, J.D., Dale, A.M., Fischl, B., Liu, A.K., Tootell, R.B., 1999. The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *J. Neurosci.* 19, 8560–72.
- Mercier, M.R., Foxe, J.J., Fiebelkorn, I.C., Butler, J.S., Schwartz, T.H., Molholm, S., 2013. Auditory-driven phase reset in visual cortex: human electrocorticography reveals mechanisms of early multisensory integration. *Neuroimage* 79, 19–29. <https://doi.org/10.1016/j.neuroimage.2013.04.060>
- Michel, C.M., Murray, M.M., 2012. Towards the utilization of EEG as a brain imaging tool. *Neuroimage* 61, 371–85. <https://doi.org/10.1016/j.neuroimage.2011.12.039>
- Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R., 2004. EEG source imaging. *Clin. Neurophysiol.* 115, 2195–222. <https://doi.org/10.1016/j.clinph.2004.06.001>
- Murray, M.M., Brunet, D., Michel, C.M., 2008. Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr.* 20, 249–64. <https://doi.org/10.1007/s10548-008-0054-5>
- Murray, M.M., Foxe, D.M., Javitt, D.C., Foxe, J.J., 2004. Setting boundaries: brain dynamics of modal and amodal illusory shape completion in humans. *J. Neurosci.* 24, 6898–903. <https://doi.org/10.1523/JNEUROSCI.1996-04.2004>
- Murray, M.M., Herrmann, C.S., 2013. Illusory contours: a window onto the neurophysiology of constructing perception. *Trends Cogn. Sci.* 17, 471–481. <https://doi.org/10.1016/j.tics.2013.07.004>
- Murray, M.M., Imber, M.L., Javitt, D.C., Foxe, J.J., 2006. Boundary completion is automatic and dissociable from shape discrimination. *J. Neurosci.* 26, 12043–54. <https://doi.org/10.1523/JNEUROSCI.3225-06.2006>
- Murray, M.M., Lewkowicz, D.J., Amedi, A., Wallace, M.T., 2016a. Multisensory Processes: A Balancing Act across the Lifespan. *Trends Neurosci.* 39, 567–579. <https://doi.org/10.1016/j.tins.2016.05.003>
- Murray, M.M., Matusz, P.J., Amedi, A., 2015. Neuroplasticity: Unexpected Consequences of Early Blindness. *Curr. Biol.* 25, R998–R1001. <https://doi.org/10.1016/j.cub.2015.08.054>
- Murray, M.M., Thelen, A., Thut, G., Romei, V., Martuzzi, R., Matusz, P.J., 2016b. The

- multisensory function of the human primary visual cortex. *Neuropsychologia* 83, 161–169. <https://doi.org/10.1016/j.neuropsychologia.2015.08.011>
- Murray, M.M., Wallace, M.T., 2012. *The Neural Bases of Multisensory Processes*. CRC Press, Boca Raton, FL.
- Murray, M.M., Wylie, G.R., Higgins, B. a, Javitt, D.C., Schroeder, C.E., Foxe, J.J., 2002. The spatiotemporal dynamics of illusory contour processing: combined high-density electrical mapping, source analysis, and functional magnetic resonance imaging. *J. Neurosci.* 22, 5055–73.
- Nieder, a, Wagner, H., 1999. Perception and neuronal coding of subjective contours in the owl. *Nat. Neurosci.* 2, 660–3. <https://doi.org/10.1038/10217>
- Noesselt, T., Tyll, S., Boehler, C.N., Budinger, E., Heinze, H.-J., Driver, J., 2010. Sound-induced enhancement of low-intensity vision: multisensory influences on human sensory-specific cortices and thalamic bodies relate to perceptual enhancement of visual detection sensitivity. *J. Neurosci.* 30, 13609–23. <https://doi.org/10.1523/JNEUROSCI.4524-09.2010>
- Ohshiro, T., Angelaki, D.E., DeAngelis, G.C., 2017. A Neural Signature of Divisive Normalization at the Level of Multisensory Integration in Primate Cortex. *Neuron* 95, 399–411.e8. <https://doi.org/10.1016/j.neuron.2017.06.043>
- Pegna, A.J., Khateb, A., Murray, M.M., Landis, T., Michel, C.M., 2002. Neural processing of illusory and real contours revealed by high-density ERP mapping. *Neuroreport* 13, 965–8.
- Perrin, F., Pernier, J., Bertrand, O., Giard, M.H., Echallier, J.F., 1987. Mapping of scalp potentials by surface spline interpolation. *Electroencephalogr. Clin. Neurophysiol.* 66, 75–81.
- Poscoliero, T., Girelli, M., 2017. Electrophysiological Modulation in an Effort to Complete Illusory Figures: Configuration, Illusory Contour and Closure Effects. *Brain Topogr.* 0, 0. <https://doi.org/10.1007/s10548-017-0582-y>
- Ramsden, B.M., Hung, C.P., Roe, a W., 2001. Real and illusory contour processing in area V1 of the primate: a cortical balancing act. *Cereb. Cortex* 11, 648–65.
- Redies, C., Crook, J.M., Creutzfeldt, O.D., 1986. Neuronal responses to borders with and without luminance gradients in cat visual cortex and dorsal lateral geniculate nucleus. *Exp. Brain Res.* 61, 469–81.
- Romei, V., Gross, J., Thut, G., 2012. Sounds reset rhythms of visual cortex and corresponding human visual perception. *Curr. Biol.* 22, 807–13. <https://doi.org/10.1016/j.cub.2012.03.025>
- Romei, V., Murray, M.M., Cappe, C., Thut, G., 2009. Preperceptual and Stimulus-Selective Enhancement of Low-Level Human Visual Cortex Excitability by Sounds. *Curr. Biol.* 19, 1799–1805. <https://doi.org/10.1016/j.cub.2009.09.027>
- Romei, V., Murray, M.M., Merabet, L.B., Thut, G., 2007. Occipital transcranial magnetic stimulation has opposing effects on visual and auditory stimulus detection: implications for multisensory interactions. *J. Neurosci.* 27, 11465–72. <https://doi.org/10.1523/JNEUROSCI.2827-07.2007>
- Romei, V., Murray, M.M.M.M., Cappe, C., Thut, G., 2013. The contributions of sensory dominance and attentional bias to cross-modal enhancement of visual cortex excitability. *J. Cogn. Neurosci.* 25, 1122–35. https://doi.org/10.1162/jocn_a_00367
- Sarmiento, B.R., Matusz, P.J., Sanabria, D., Murray, M.M., 2016. Contextual factors multiplex to control multisensory processes. *Hum. Brain Mapp.* 37.

<https://doi.org/10.1002/hbm.23030>

- Schroeder, C.E., Foxe, J.J., 2002. The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res. Cogn. Brain Res.* 14, 187–98.
- Seghier, M., Dojat, M., Delon-Martin, C., Rubin, C., Warnking, J., Segebarth, C., Bullier, J., 2000. Moving illusory contours activate primary visual cortex: an fMRI study. *Cereb. Cortex* 10, 663–70.
- Seghier, M.L., Vuilleumier, P., 2006. Functional neuroimaging findings on the human perception of illusory contours. *Neurosci. Biobehav. Rev.* 30, 595–612.
<https://doi.org/10.1016/j.neubiorev.2005.11.002>
- Shams, L., Kamitani, Y., Shimojo, S., 2000. Illusions. What you see is what you hear. *Nature* 408, 788. <https://doi.org/10.1038/35048669>
- Sheth, B.R., Sharma, J., Rao, S.C., Sur, M., 1996. Orientation maps of subjective contours in visual cortex. *Science* 274, 2110–5.
- Spieler, L., Manuel, A.L., Buetti, D., Murray, M.M., 2013. Contributions of pitch and bandwidth to sound-induced enhancement of visual cortex excitability in humans. *Cortex* 49, 2728–34. <https://doi.org/10.1016/j.cortex.2013.01.001>
- Spinelli, L., Andino, S.G., Lantz, G., Seeck, M., Michel, C.M., 2000. Electromagnetic inverse solutions in anatomically constrained spherical head models. *Brain Topogr.* 13, 115–25.
- Stein, B.E., London, N., Wilkinson, L.K., Price, D.D., 1996. Enhancement of perceived visual intensity by auditory stimuli: a psychophysical analysis. *J. Cogn. Neurosci.* 8, 497–506.
<https://doi.org/10.1162/jocn.1996.8.6.497>
- Stein, B.E., Meredith, M.A., 1993. *The Merging of the Senses*. MIT Press.
- Talairach, J., Tournoux, P., 1988. *Co-Planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System: An Approach to Cerebral Imaging*. Thieme.
- ten Oever, S., Romei, V., van Atteveldt, N., Soto-Faraco, S., Murray, M.M., Matusz, P.J., 2016. The COGs (context, object, and goals) in multisensory processing. *Exp. Brain Res.* 234. <https://doi.org/10.1007/s00221-016-4590-z>
- Tinga, A.M., Visser-Meily, J.M.A., van der Smagt, M.J., Van der Stigchel, S., van Ee, R., Nijboer, T.C.W., 2016. Multisensory Stimulation to Improve Low- and Higher-Level Sensory Deficits after Stroke: A Systematic Review. *Neuropsychol. Rev.* 26, 73–91.
<https://doi.org/10.1007/s11065-015-9301-1>
- Toepel, U., Knebel, J.-F., Hudry, J., le Coutre, J., Murray, M.M., 2009. The brain tracks the energetic value in food images. *Neuroimage* 44, 967–74.
<https://doi.org/10.1016/j.neuroimage.2008.10.005>
- van Atteveldt, N., Murray, M.M.M., Thut, G., Schroeder, C.E.C.E., 2014. Multisensory integration: Flexible use of general operations. *Neuron* 81, 1240–1253.
<https://doi.org/10.1016/j.neuron.2014.02.044>

Figure Legends

Figure 1. Examples of stimuli. The upper left panel depicts an array inducers resulting in a modally-completed illusory contour (MIC). The upper right panel depicts an array of modal inducers that do not form an illusory contour (MNC). The lower left panel depicts an array inducers resulting in an amodally-completed illusory contour (AIC). The lower right panel depicts an array of amodal inducers that do not form an illusory contour (ANC).

Figure 2. Group-averaged unsubtracted ERPs. Data are displayed from an exemplar occipital midline scalp location (Oz) and show the responses to the MIC_SND, MNC_SND, MIC_NOSND, and MNC_NOSND conditions (see inset). Enhanced responses to the MIC_SND condition are apparent, peaking at ~150ms post-stimulus onset at this scalp site.

Figure 3. Group-averaged IC_{effect} ERPs and results. **a.** The $IC_{\text{effect_MSND}}$ and $IC_{\text{effect_MNOSND}}$ waveforms at an occipital midline scalp site (Oz) are displayed (mean \pm SEM indicated). **b.** The percentage of the 128-channel electrode montage exhibiting a significant difference (non-parametric randomization test) is displayed. Note that this graph does illustrate results after applying the temporal threshold criterion of significant effects lasting for at least 15ms contiguously. **c.** The spatial correlation between the $IC_{\text{effect_MSND}}$ and $IC_{\text{effect_MNOSND}}$ conditions is plotted as a function of time (blue trace) alongside the timing of significant differences in the ERP topography (yellow area plot).

Figure 4. Topographic clustering and single-subject fitting results. **a.** Two template maps were observed in the group-averaged IC_{effect} ERPs over the 70-170ms period and are displayed from a top view as well as back view. **b.** Single-subject fitting was based on the spatial correlation of each template map in *a* with the data from each participant from each condition. This yields a total percentage of the 70-170ms time period better characterized topographically by each template map (mean \pm SEM shown). Significant differences (Wilcoxon signed rank test) are indicated by asterisks.

Figure 5. Source estimations and correlation results. **a.** The statistical contrast of distributed source estimations from the 70-170ms period resulted in 3 clusters of robust differential

responses ($p < 0.05$ at each solution point and a minimal cluster size of 15 points). These were located within the inferior parietal lobe (IPL), lateral occipital cortex (LOC), and primary visual cortex (V1). **b.** Non-directed interaction was assessed via linear correlation. Source estimations across all three clusters were significantly correlated for the IC_{effect_MSND} condition, but not for the IC_{effect_MNOSND} condition.

Supplementary Figure S1. Schematic overview of the topographic clustering and single-subject fitting procedure.

Supplementary Figure S2. Behavioral data. **a.** Group-averaged values of d' are displayed. **b.** Group-averaged values of c are displayed. **c.** Group-averaged reaction times (in milliseconds) are displayed. Error bars indicate the standard error of the mean.

Supplementary Figure S3. Mean Global Field Power waveforms are displayed for the IC_{effect_MSND} and IC_{effect_MNOSND} conditions over the 300ms post-stimulus period. There was no evidence of statistically reliable differences.

Supplementary Figure S4. Topographic clustering results. **a.** The template maps identified via the hierarchical topographic clustering are displayed for the 300ms post-stimulus period. Note that 9 template maps were identified over the full 500ms post-stimulus period and accounted for 89% of the global explained variance. **b.** The sequence of template maps observed in the group-averaged data from each condition (IC_{effect_MSND} and IC_{effect_MNOSND}). Each template map in **a** is represented by a color, and the time intervals when the maps were observed in the group-averaged data are indicated.

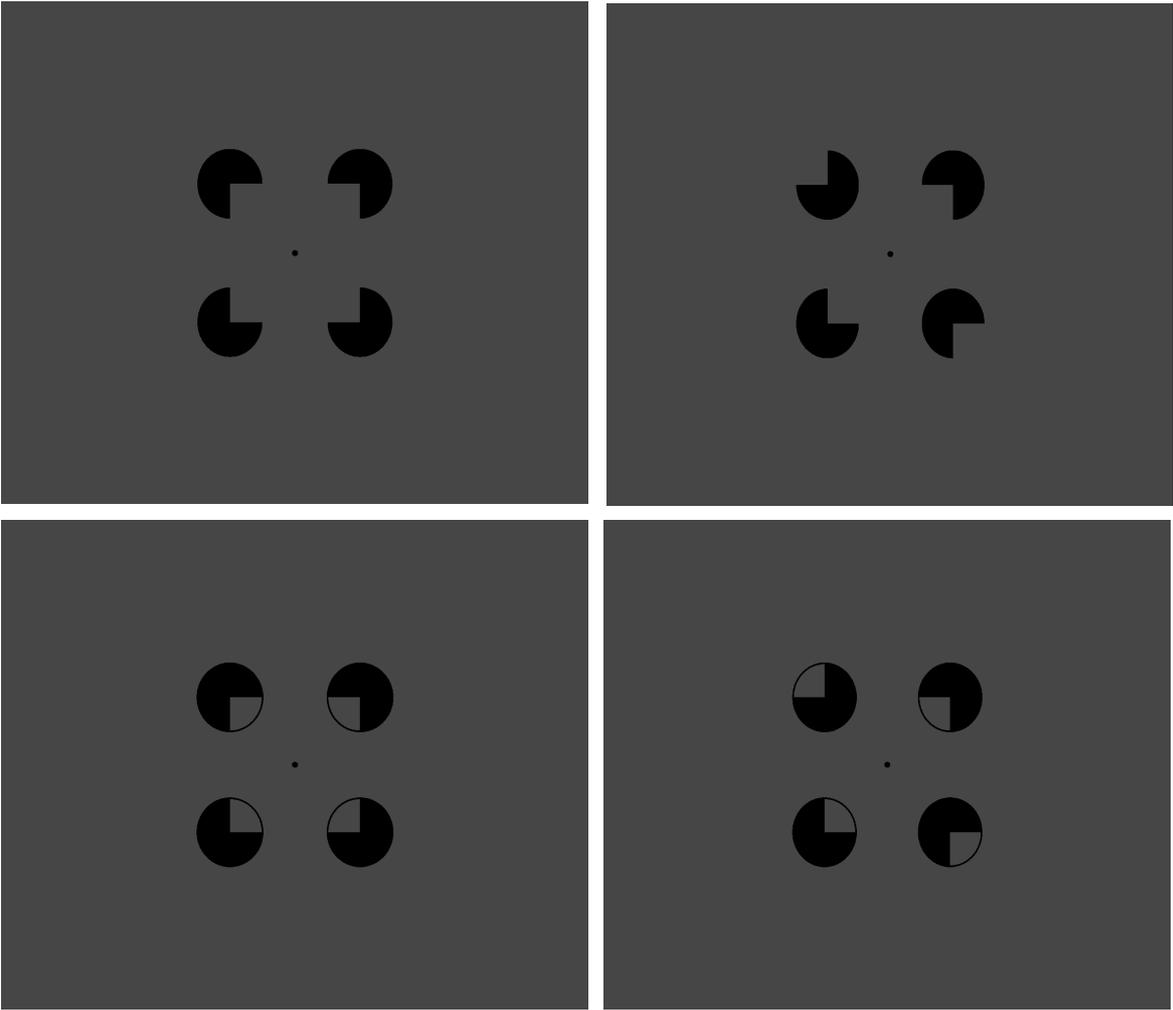


Figure 1

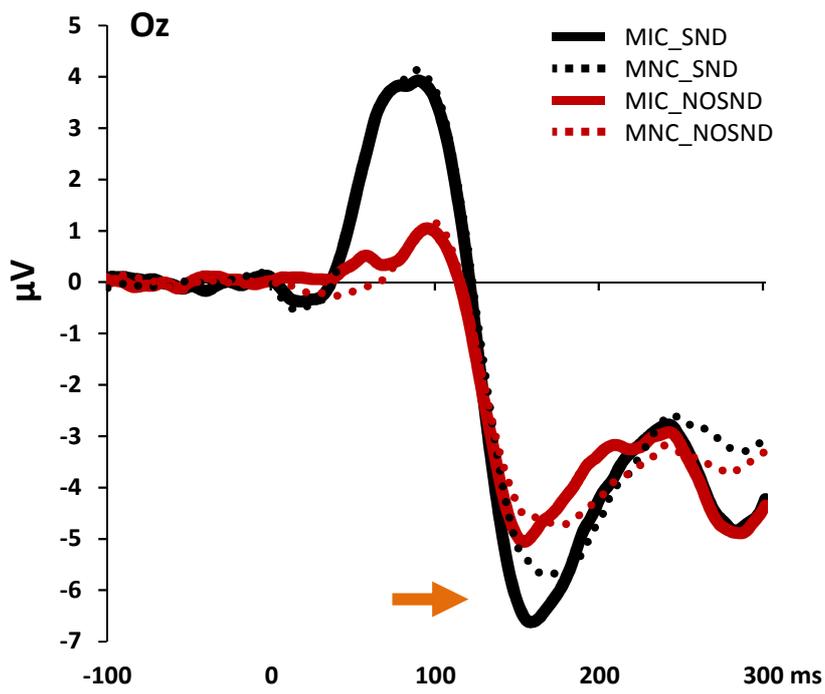


Figure 2

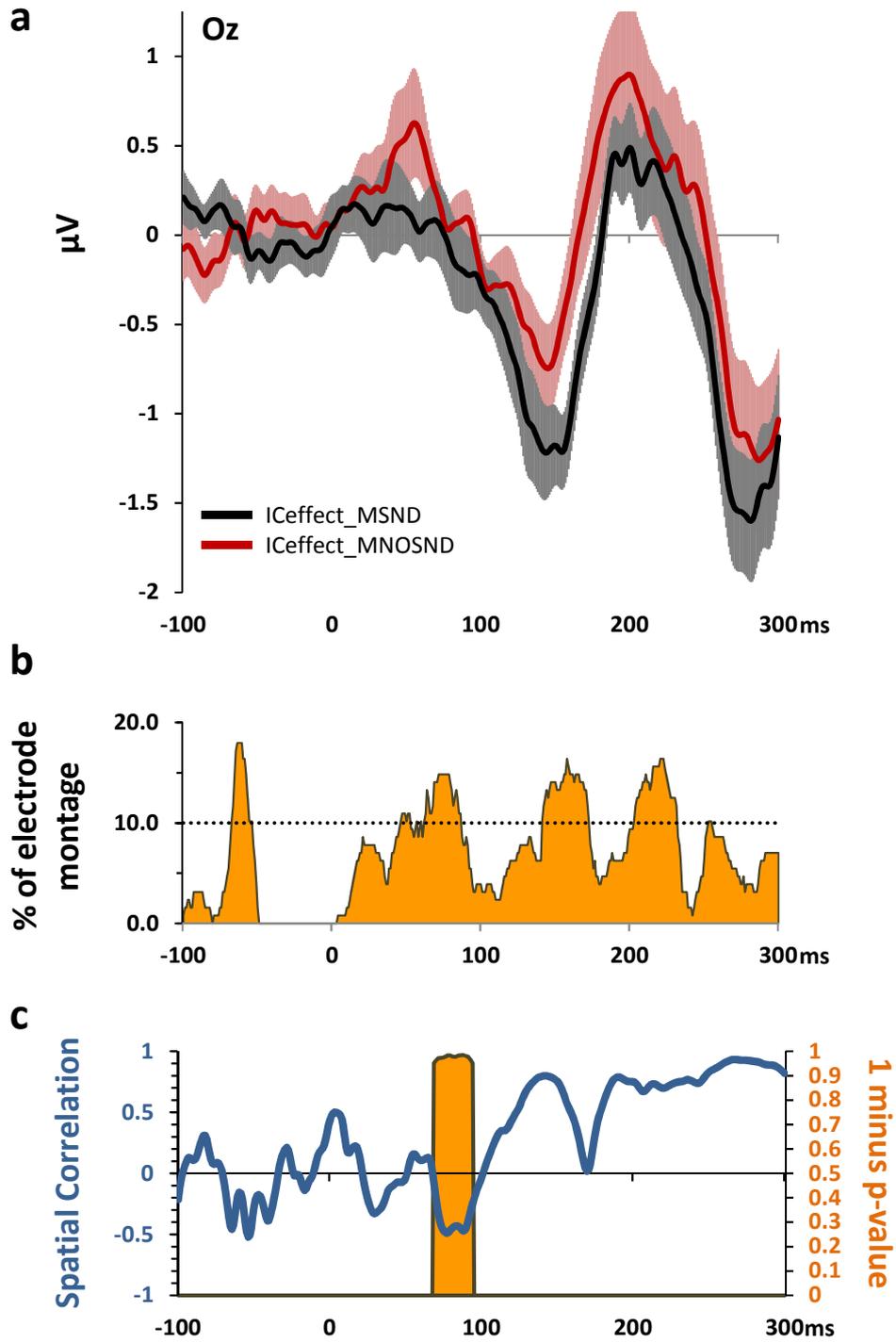


Figure 3

a Topographic Cluster Analysis
and Single-Subject Fitting
(70-170ms)

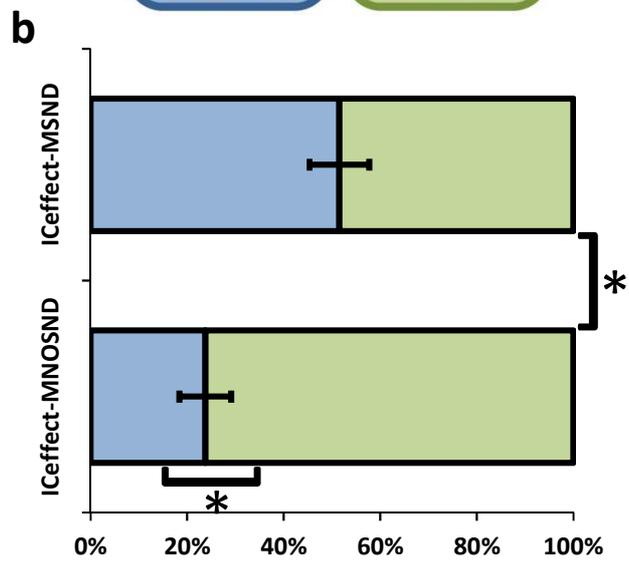
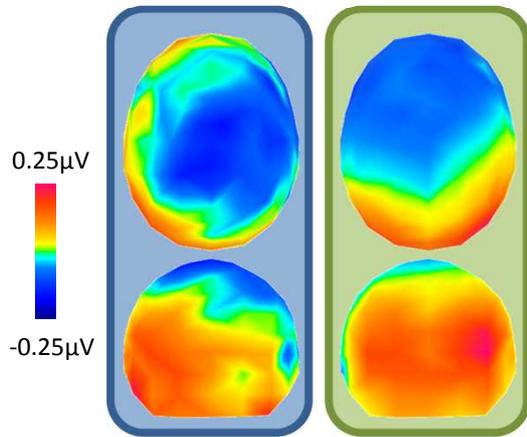


Figure 4

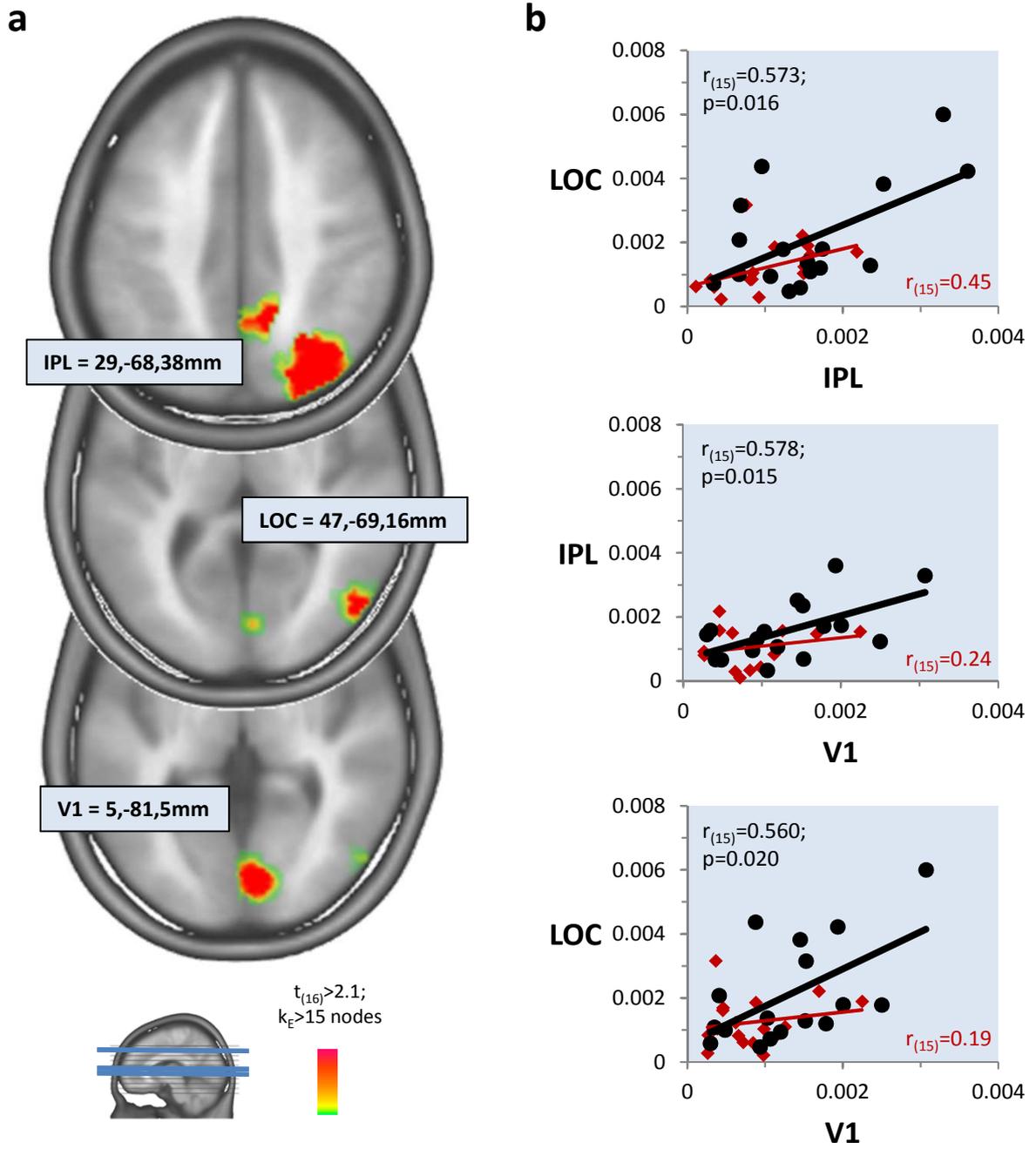
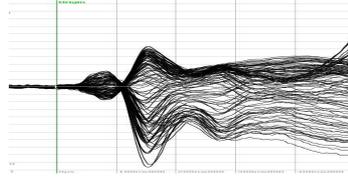
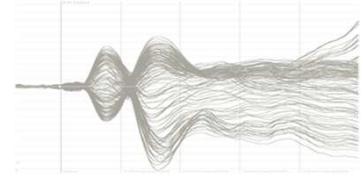


Figure 5

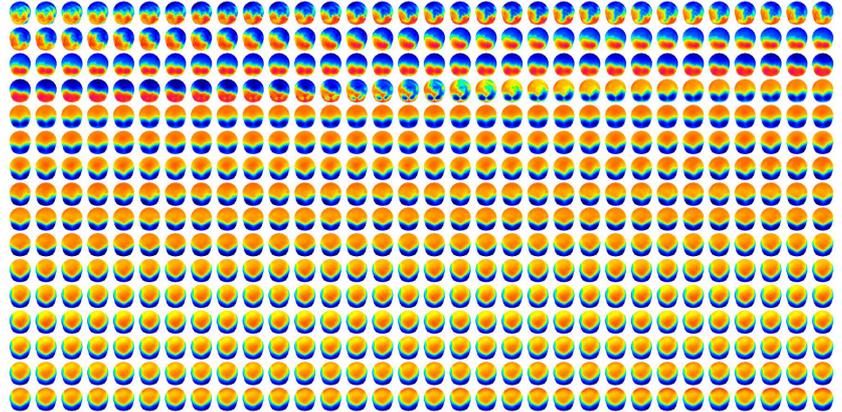
Condition A



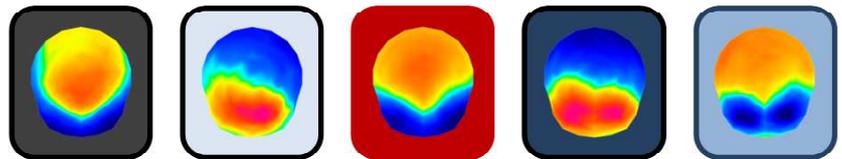
Condition B



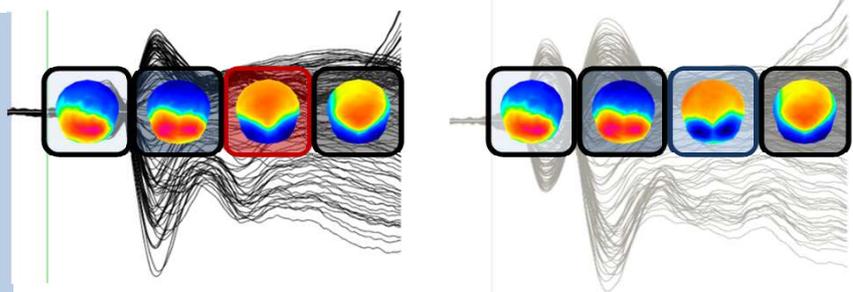
- The concatenated dataset is treated as a series of strength-normalized maps (i.e. N-dimensional vectors, where N equals the number of electrodes)
- Clustering is performed on these data (e.g. hierarchical, k-means, etc.)



- As output, the clustering yields a set of "template" maps (independent of condition and time labels).
- These characterize the group-averaged ERPs (i.e. the fewest maps that provide the highest global explained variance).



- Template maps are then shown as a function of time according to their characterization of the group-averaged data.
- Topographic clustering thus generates hypotheses regarding whether and when different template maps characterize ERPs to different conditions (e.g. red and blue maps each better characterize each condition over the 200-300ms time window).



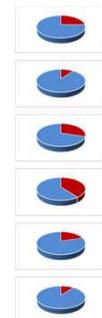
- Such hypotheses are tested by calculating the spatial correlation (C) between specific template maps (e.g. red and blue) and a circumscribed time window of the single-subject data from each condition.
- Each time point is labelled according to the template map with which there is a higher spatial correlation in a winner-takes-all manner.
- The pie charts show the amount of time over the 200-300ms time window when the red or blue map better correlated spatially.

subject 1
subject 2
subject 3
subject 4
subject 5
...
subject n

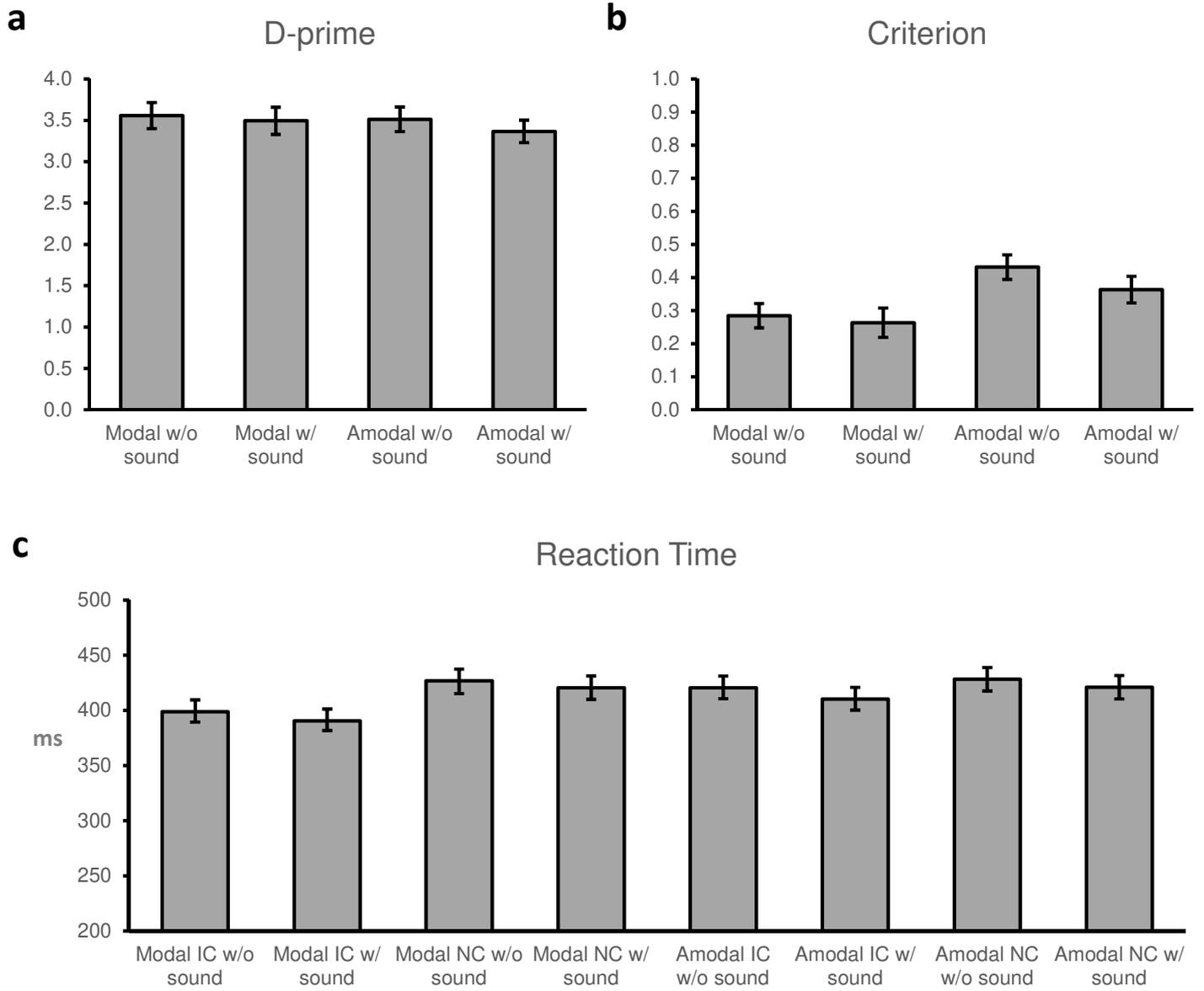


$$C = \frac{\sum_{i=1}^n (u_i \cdot v_i)}{\sqrt{\sum_{i=1}^n u_i^2} \cdot \sqrt{\sum_{i=1}^n v_i^2}}$$

Where
 n = #electrodes
 u_i = measured voltage at electrode i of single-subject data from a given condition
 v_i = measured voltage at electrode i of template map

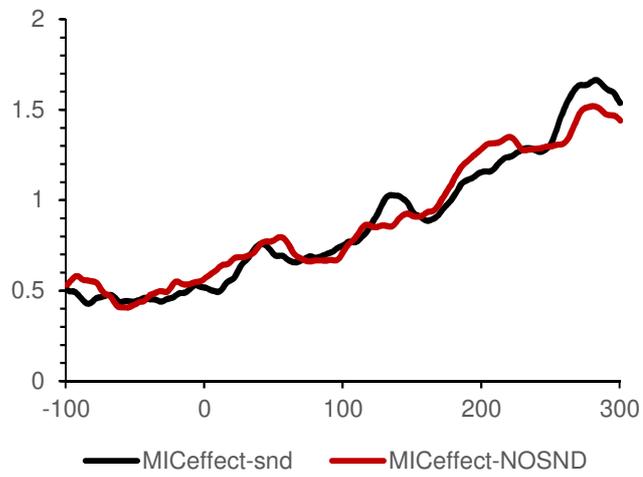


Supplementary Figure S1



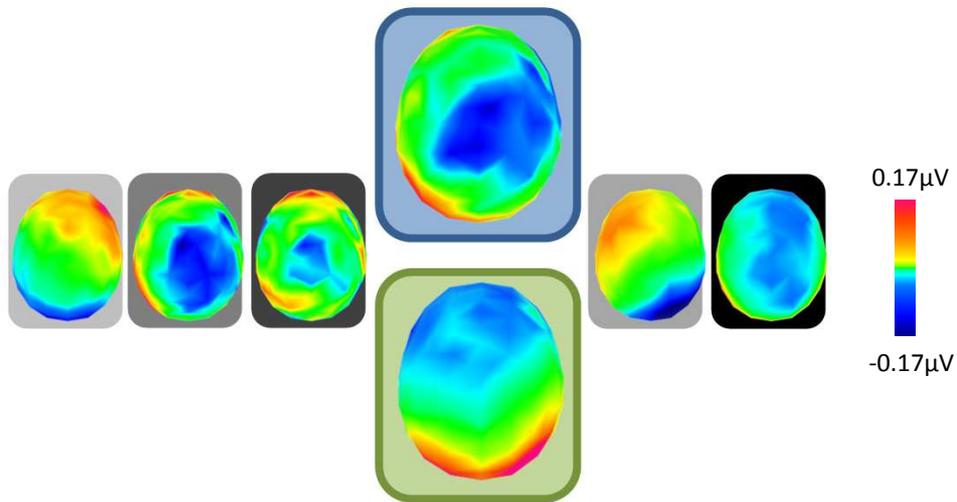
Supplementary Figure S2

Global Field Power



Supplementary Figure S3

a Template maps identified via hierarchical clustering over the 300ms post-stimulus interval



b The pattern of template maps observed in the group-averaged data from each condition

