Special Issue "Multisensory integration": Research Report

Trimodal processing of complex stimuli in inferior parietal cortex is modality-independent

Danja K. Porada a,*, Christina Regenbogen a,b,c, Jessica Freiherr d, Janina Seubert a,1 and Johan N. Lundström a,e,f,g,1,**

a Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
b Department of Psychiatry, Psychotherapy and Psychosomatics, Medical School, RWTH Aachen University, Aachen, Germany
c JARA Institute Brain Structure Function Relationship, RWTH Aachen University, Aachen, Germany
d Department of Psychiatry and Psychotherapy, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
e Monell Chemical Senses Center, Philadelphia, USA
f Department of Psychology, University of Pennsylvania, Philadelphia, USA
g Stockholm University Brain Imaging Centre, Stockholm University, Stockholm, Sweden

A R T I C L E   I N F O

Article history:
Received 13 June 2020
Reviewed 27 October 2020
Revised 29 November 2020
Accepted 9 March 2021
Published online 19 March 2021

Keywords:
Multisensory integration
Object information
Trimodal
Auditory-visual-olfactory
Intraparietal sulcus
Inferior parietal cortex

A B S T R A C T

In humans, multisensory mechanisms facilitate object processing through integration of sensory signals that match in their temporal and spatial occurrence as well as their meaning. The generalizability of such integration processes across different sensory modalities is, however, to date not well understood. As such, it remains unknown whether there are cerebral areas that process object-related signals independently of the specific senses from which they arise, and whether these areas show different response profiles depending on the number of sensory channels that carry information.

To address these questions, we presented participants with dynamic stimuli that simultaneously emitted object-related sensory information via one, two, or three channels (sight, sound, smell) in the MR scanner. By comparing neural activation patterns between various integration processes differing in type and number of stimulated senses, we showed that the left inferior frontal gyrus and areas within the left inferior parietal cortex were engaged independently of the number and type of sensory input streams. Activation in these areas was enhanced during bimodal stimulation, compared to the sum of unimodal activations, and increased even further during trimodal stimulation.

Taken together, our findings demonstrate that activation of the inferior parietal cortex during processing and integration of meaningful multisensory stimuli is both modality-independent and modulated by the number of available sensory modalities. This suggests that the processing demand placed on the parietal cortex increases with the number...
1. Introduction

While queueing in our favorite coffee shop to get our morning coffee, several of our senses are simultaneously stimulated with overlapping information. The appetizing smell of freshly brewed coffee is combined with the sight and sound of the coffee machine, and these dynamic sensory impressions are swiftly integrated. Relating the different sensory channels to each other allows us to selectively attend to those inputs that form a familiar whole and to separate them from other simultaneously occurring events, such as the chatter of the barista in the background. The combination of redundant information originating from the same object but from different sensory sources increases our ability to separate relevant object-related signals from object-irrelevant noise, and thereby generates a more reliable object percept than any of the contributing sources could produce alone (Ernst & Banks, 2002). Multisensory integration thus “increases the collective impact of biologically significant signals” (Stein, Stanford, & Rowland, 2014), which serves an evolutionarily adaptive function that is ubiquitous across species (Grefkes & Fink, 2005).

In human experimental research, the mechanisms of multisensory processing have predominantly been examined by studying participants’ responses to simplified bimodal combinations of visual and auditory stimuli, such as flashes and beeps, that occur in spatial or temporal proximity (Slutsky & Recanzone, 2001). Results demonstrate that integration is achieved both by direct enhancement of activity of, and functional connectivity between, sensory-specific areas (Foote & Schroeder, 2005; Lewis & Noppeney, 2010), as well as the recruitment of other, modality-independent brain regions, such as the superior temporal sulcus (STS) or intraparietal sulcus (IPS) (Calvert, 2001). While such paradigms provide important insights into the most essential crossmodal binding mechanisms across species, they tend to overlook the ecological relevance of the stimuli. Most importantly, beeps and flashes do not allow us to understand how humans routinely use their knowledge about objects to relate information from different channels to one another, as illustrated in the coffee example above. To achieve this, our brains need to have the capacity for not only temporal and spatial information, but also object knowledge and past experiences to modulate signal integration across sensory channels (Hein et al., 2007; Noppeney, Josephs, Hocking, Price, & Friston, 2008). This provides an important basis for sensory integration to be adaptable to the individual circumstances of the observer, and also, to facilitate changes in sensory processing over the life course.

A number of recent studies have shown substantial cortical activation differences between the perceptual processing of simplified laboratory stimuli, and that of more naturalistic stimuli that carry meaning for the perceiver. The visual system, for example, has been shown to be particularly well adapted to processing properties of the natural environment, exhibiting qualitatively different response profiles to simplified laboratory stimuli and more complex naturalistic objects (Felsen & Dan, 2005; Hein et al., 2007; Kaysker, Kording, & Konig, 2004). Artificial and everyday stimuli do not only differ in their basic sensory processing, but also in their corresponding integration processes. For example, earlier and stronger integration effects have been shown for dynamic auditory-visual object stimuli than for abstract stimuli, and the effects were attributed to a shift in activation towards a more specialized fronto-temporal network for semantic processing (Senkowski, Saint-Amour, Kelly, & Foxe, 2007). The engagement of inferior parietal areas, which have been shown to be substantially involved in multisensory processing (Bremmer et al., 2001; Calvert, Campbell, & Brammer, 2000; Driver & Noesselt, 2008; Grefkes & Fink, 2005; Macaluso & Driver, 2001; Molholm et al., 2006), was, however, lower for naturalistic stimuli than for abstract stimuli coinciding in space and time (Senkowski et al., 2007). The authors speculated that this reflected a lower need for engagement of intraparietal areas during integration of naturally related meaningful compared to unrelated meaningless multisensory stimuli, indicating an ongoing need for studies exploring the integration of such meaningful multisensory object stimuli.

Another important difference between laboratory studies of multisensory processing and our everyday environment lies in the number of sensory modalities present. When object-like stimulus material was presented in multisensory studies, it most commonly consisted of combined dynamic audio-visual recordings of stimuli such as natural speech (Calvert et al., 2000), dynamic naturalistic objects (Beauchamp, Lee, Argall, & Martin, 2004), and manual tools or real-life events (Beauchamp, Argall, Bodurka, Duyn, & Martin, 2004; Senkowski et al., 2007; Stevenson, Geoghegan, & James, 2007). Whether the central multisensory processes identified for relating and integrating meaningful auditory and visual stimuli are generalizable to other sensory modalities, and to what extent processing changes when more than two modalities are stimulated at once, remains to be explored.

To our knowledge, the only neuroimaging study that has studied simultaneous processing of naturally meaningful stimuli in a larger number of sensory modalities (Kassuba et al., 2011) revealed that both bi- and trimodal combinations of images, sounds, and haptic information of real-world objects activated the inferior frontal gyrus (IFG), insular cortex, superior temporal gyrus/sulcus (STG/STS), and left fusiform gyrus (FG). These areas are thought to serve key functions for integration of object information, including linkage of sensory information with semantic memory.
(Amedi, von Kriegstein, van Atteveldt, Beauchamp, & Naumer, 2005), mediation of communication and information exchange between unisensory cortices (Amedi et al., 2005), integration of object identity and complex featural information (Beauchamp, Lee, et al., 2004; Calvert, 2001), and activation of unified multisensory object representations (Kassuba et al., 2011; Lundström, Regenbogen, Ohla, & Seubert, 2018). Regions that were more active during trimodal than bimodal processing, such as the inferior parietal (IPC) and dorsolateral prefrontal cortex (DLPFC), were attributed to attention- and working memory-related processes.

The study thus yielded important insights into the similarities and differences between bi- and trimodal processing of meaningful stimuli, and revealed candidate regions for their multisensory processing. However, little is known to date about the specificity of the observed multisensory effects to the sensory modalities involved: can multisensory processing sites be identified that relate sensory object information independently of the sensory modalities through which the object was perceived? If such modality-independent multisensory processing sites exist, do they respond differently when more than two sensory modalities are to be related to each other?

The present functional magnetic resonance imaging (fMRI) study sought to determine the cortical activation patterns representing the processing of object-related multisensory information by presenting participants with videos, sounds, and smells of everyday objects. Our first goal was to define the cortical areas that process bimodal object information, independently of the specific sensory modalities involved. As a next step, we assessed multisensory effects for trimodal object information to identify possible regional differences depending on the number of stimulated senses. Finally, we computed the overlap between bimodal and trimodal multisensory effects to identify brain regions relating object information to each other, independently of the type and number of stimulated sensory modalities.

2. Material and methods

2.1. Participants

Sixteen volunteers (7 women, mean age 26.9 years, SD 3.2 years) participated in the study. All participants were right-handed, did not take any medication, had no history of functional sensory impairments (i.e., anosmia, hearing, or vision deficits), reported normal hearing, and had confirmed normal smell [MONEX-40 (Freiherr et al., 2012)] and vision (Snellen’s visual acuity evaluation, www.mdsupport.org/snellen.html). All aspects of the study were approved by the local Institutional Review Board of the University of Pennsylvania and participants provided written informed consent prior to study inclusion.

2.2. Stimuli

The stimulus set and experimental design is described in detail in a prior article on this dataset (Porada, Regenbogen, Seubert, Freiherr, & Lundström, 2019). In brief, the stimulus set consisted of six familiar objects and one baseline stimulus, presented through sight, sound, and smell. Objects were selected based on their clear and distinct features for all three sensory modalities: ‘Coffee brewer’, ‘Wood fire’, ‘Lawn mower’, ‘Popcorn’, ‘Flopping fish’, and ‘Flushing urinal’. Examples of the visual and auditory stimuli are available in the online supplementary material.

Visual stimuli (V) consisted of three different 3 sec long digital video sequences for each of the six objects and one blank white screen video as visual baseline stimulus. The object video sequences were individually cut from a longer video to create three unique, but nearly identical alternative versions of the videos. All stimuli contained a centered fixation cross. Video clips were cropped to a duration of 3 sec, resized to 640*480 pixels resolution, equalized in luminance levels and overlaid with a fixation cross using Virtual Dub (www.virtualdub.org).

Auditory stimuli (A) consisted of three unique, but nearly identical alternative versions of 3 sec long sequences for each of the six objects and one silent audio track as an auditory baseline stimulus. The object audio sequences were created from the video clips’ audio tracks. Like the video sequences, the audio tracks were individually created from a longer sound file that was synchronized with the video. All sounds were volume-matched using MP3Gain (www.mp3gain.sourceforge.net) and presented at a volume that allowed clear audibility of the objects inside the scanner.

Olfactory stimuli (O) consisted of six different iso-intense odors, one for each object, and a neat odorless diluent (1,2-propanediol, Sigma Aldrich) as olfactory baseline stimulus. The object-related odors comprised of: freshly brewed coffee (object ‘Coffee brewer’, 50% v/v of Brazilian coffee oil, International Flavor and Fragrances [IFF], dissolved in diethyl phthalate, Sigma Aldrich), smoky fire (object ‘Wood fire’, 20% v/v of Guaiacol, Sigma Aldrich, dissolved in 1,2-propanediol, Sigma Aldrich), freshly cut grass (object ‘Lawn mower’, 7.5% v/v of triple-A-aldehyde, IFF, dissolved in 1,2-propanediol, Sigma Aldrich), butty popcorn (object ‘Popcorn’, 2.5% v/v of 2,3-butanedione, Sigma Aldrich, dissolved in 1,2-propanediol, Sigma Aldrich), fresh fish (object ‘Flopping fish’, 3% v/v of fish aroma, Givaudan Inc, dissolved in 1,2-propanediol, Sigma Aldrich), and urine (object ‘Flushing urinal’, 100% v/v Urine malodor, Givaudan Inc.). Odors were presented binocularly with an MRI-compatible computer controlled olfactometer (Lundström, Gordon, Alden, Boesveldt, & Albrecht, 2010) using an airflow of 3.0 L per minute (l/m). This odor stream was embedded in a constant binorial flow of .5 L/m to mask tactile or thermal sensations caused by a sudden flow onset, resulting in a total airflow of 3.5 l/m. Between odor presentations, an identical flow of clean air was presented to remove residual odor and reduce potential odor adaptation/habitation effects. The time lag of odor onset was measured using a photoionization detector and odor triggering was adjusted accordingly to assure synchronous onset of the odor with the other sensory stimuli (Lundström et al., 2010).
2.3. **Experimental design**

The olfactory, visual, and auditory stimuli were presented in combinations that formed unimodal (object-related information in one modality, baseline stimulation in the others), bimodal (object-related information in two modalities, baseline stimulation in the other), trimodal (object-related information in all three modalities), or baseline (object-related information completely absent, baseline stimulation in all three modalities) stimuli. This resulted in the following eight combinations: 3x unimodal stimulation (O, A, V), 3x bimodal stimulation (OA, OV, AV), 1x trimodal stimulation (OAV), 1x baseline stimulation (see Fig. 1B). To prevent participants from being able to guess the content of one modality based on focusing their attention on another, bimodal and trimodal combinations could further be either congruent (modalities receive object information associated with the same object), or incongruent (modalities receive object information associated with different objects). For the present study, only the unimodal, congruent bi- and trimodal conditions as well as the baseline condition were of interest and will in the following be referred to as: O, A, V, OA, OV, AV, OAV, BL.

Each condition was repeated 18 times (three times per object) resulting in a total of 216 stimulations, of which 144 were trials of interest for this analysis. Trials were pseudo-randomized, so that the same type of stimulation was not repeated more than once every 36 trials.

2.4. **Procedure and task**

The experiment consisted of two sessions, separated by two days. In Session 1, participants were familiarized with the scanning environment, tasks and stimuli. Specifically, participants underwent an extensive stimulus association training inside a mock scanner, where the congruent trimodal stimuli were repeatedly presented to assure that the information from the three modalities would be strongly associated with each other and to assure comparable expertise between participants. After training, participants rated the odors for several perceptual attributes. They then completed one experimental run identical to the task in Session 2 as training.

Session 2 was conducted inside an MRI scanner (3T Siemens TIM, Erlangen, Germany). Each trial started with a ‘get ready’ message, displayed for 4.33 sec on average (jittered between 2.2 sec and 8.8 sec). This message was followed by a fixation cross, which indicated trial onset and cued the

---

**Fig. 1** – Overview of experimental paradigm and stimulus combinations. A) During each trial, participants were first asked to prepare to sniff and then presented with a uni-, bi- or trimodal stimulus presentation or a blank baseline trial. Afterwards, participants indicated via button-press how many stimuli they had perceived. B) Experimental conditions comprised baseline trials (BL, blue crosshair); unimodal conditions: visual (V), auditory (A), or olfactory (O); bimodal conditions: olfactory-auditory (OA, odor + sound), olfactory-visual (OV, odor + video), auditory-visual (AV, sound + video); trimodal condition: olfactory-auditory-visual (OAV, odor + sound + video). C) Illustration of the visual stimuli.
participant to sniff. Concurrently with the fixation cross, participants were exposed to a baseline, uni-, bi-, or trimodal stimulus combination for 3 sec (Fig. 1A). At the end of each trial, participants indicated the number of stimulated sensory modalities during a 3 sec long response window, a task administered to maintain alertness and to ensure equally shared attention to all sensory modalities. Stimulus presentation and response collection were controlled by E-Prime (RRID:SCR_009567).

2.5. Functional imaging measurements

Functional blood oxygen level dependent (BOLD) images were acquired with a T2*-weighted, gradient-echo, echoplanar imaging (EPI) sequence using a standard eight-channel head coil (TE = 22 msec, TR = 2500 msec, field of view = 240 × 240 mm, matrix size = 96 × 96, pixel size = 2.5 × 2.5 mm, flip angle = 70°, 40 oblique slices with a 2.5 mm slice thickness, no gap). The scanning session consisted of two runs with 449 volumes each. To maximize the signal-to-noise-ratio in areas of the OFC and the frontal–temporal junction, image acquisition was tilted by 30° from the horizontal axis (Deichmann, Gottfried, Hutton, & Turner, 2003).

A structural image was obtained using a whole-brain, high-resolution, three-dimensional, T1-weighted, gradient-echo (MPRAGE) sequence (TI = 1100 msec, TE = 3.87 msec, TR = 1630 msec, field of view = 240 × 180 mm, pixel size = 0.9 × 0.9 mm, flip angle = 15°, 160 slices, 1 mm slice thickness).

2.6. Functional imaging data analysis

2.6.1. Preprocessing

Data were preprocessed and analyzed using the SPM8 software (Welcome Trust Centre for Neuroimaging, London, UK, RRID:SCR_007037). The origin of the structural image was manually adjusted to the anterior commissure and all functional volumes were spatially realigned correspondingly. Within each run, functional images were slice-time corrected, and realigned to the mean functional image. After coregistering the structural image to each run’s mean functional image, the structural image delivered the priors for a unified segmentation process (Ashburner & Friston, 2005). The structural image was non-linearly segmented into gray matter, white matter, and cerebrospinal fluid (CSF). This yielded the normalization parameters that were applied to the structural as well as all functional images. After normalization, the voxel size was 3 × 3 × 3 mm for the functional image volumes and 1 × 1 × 1 mm for the structural image volume. Functional images were finally spatially smoothed using an 8 mm full-width-at-half-maximum (FWHM) isotropic Gaussian kernel.

2.6.2. Single-subject analysis

A general linear model (GLM) was set up with eight regressors of interest corresponding to the onset times of the eight experimental conditions. Incongruent conditions and realignment parameters were modeled as regressors of no interest. The regressors were convolved with the canonical hemodynamic response function (HRF) with time and dispersion derivatives. A high-pass filter with a cutoff of 128 sec removed low frequency drifts in the signal. Serial autocorrelations were accounted for by including a first order autoregressive model (AR-1).

2.6.3. Group-level analysis

Simple main contrasts for each subject were submitted to a flexible factorial design modelling subjects as random effects and the eight experimental conditions as fixed effects. Departures from sphericity were corrected for by variance components assuming a compound symmetry structure for within-subjects (correlated) measures and heteroscedasticity between subjects and conditions.

Separate sets of analyses were conducted to 1. ensure the basic effectiveness of our stimulation paradigm, 2. identify brain regions processing bimodal sensory information and regions processing trimodal information, and 3. establish which brain regions demonstrate multisensory processing independent of the number and type of stimulated channels.

To test the effectiveness of our stimulation paradigm, we assessed the effect of sensory relative to baseline stimulation for each sensory modality separately (O > BL, A > BL, V > BL). p-values of all contrasts were corrected for multiple comparisons across whole-brain volume using family wise error rate (FWE) (p < .05 corrected for multiple comparisons using Gaussian random field theory). As these analyses revealed typical activation patterns associated with olfactory, visual and auditory processing (Appendix D Figure D1, Table D1), we proceeded with testing multisensory effects.

To infer multisensory processing from BOLD responses, several statistical criteria of different stringency have been suggested (Beauchamp, 2005). For the purpose of the present study, we decided to characterize multisensory processing as superadditive activation. This means that activation measured during multisensory stimulation was considered to be specific to the process of relating sensory components if it was statistically higher than the summed activation measured during separate unimodal stimulation. To determine modality-independent bimodal effects, we first analyzed all possible bimodal combinations of modalities (olfactory-visual, olfactory-auditory, visual-auditory) separately. This was achieved by contrasting each bimodal condition to the sum of the respective unimodal conditions. Because participants were instructed to sniff during any kind of stimulation, even if there was no odor present, we adjusted for multiple subtraction of activation attributed to sniffing when computing the superadditive contrasts. In other words, we contrasted the sum of bimodal and baseline condition to the sum of the respective unimodal conditions (e.g., OA + BL > O + A, see Table 1). Each contrast revealed brain regions relating specific bimodal information to each other. These contrasts were subsequently entered in a conjunction analysis (conjunction null hypothesis; Nichols, Brett, Andersson, Wager, & Poline, 2005) to determine the overlap between all three superadditive bimodal contrasts (see Table 1). This allowed us to exclude brain regions that only processed bimodal information for specific modalities, and isolate the brain regions showing superadditive responses across all three possible bimodal
Table 1 – Contrast descriptions of all group-level contrasts used in the study.

| Contrasts                                                                 |
|---------------------------------------------------------------------------|
| **Unimodal stimulation**                                                 |
| O > BL                                                                    |
| A > BL                                                                    |
| V > BL                                                                    |
| **Bimodal processing**                                                    |
| OA + BL > O + A                                                           |
| OV + BL > O + V                                                           |
| AV + BL > A + V                                                           |
| (OA + BL > O + A) ∩ (OV + BL > O + V) ∩ (AV + BL > A + V)                 |
| **Trimodal processing**                                                   |
| (OA + BL > O + A) ∩ (OV + BL > O + V) ∩ (AV + BL > A + V)                 |
| OAV + BL > OA + V                                                         |
| OAV + BL > OV + A                                                        |
| OAV + BL > AV + O                                                        |
| (OAV + BL > OA + V) ∩ (OAV + BL > OV + A) ∩ (OAV + BL > AV + O)           |
| **Superadditive processing independent of the number and type of stimulated channels** |
| (OAV + BL > OA + V) ∩ (OAV + BL > OV + A) ∩ (OAV + BL > AV + O)           |

combinations. Hence, the resulting isolated regions could be labeled as modality-independent bimodal processing sites.

All conjunction analyses used a statistical threshold of $p < .001$ uncorrected for multiple comparisons, for each contrast entering the conjunction analysis (see Seubert, Ohla, Yokomukai, Kellermann, & Lundström, 2014 for a similar approach). This approach was chosen because significance testing against a conjunction null-hypothesis relies on the simultaneous statistical significance of three contrasts, where a threshold of FWE $p < .05$ applied to three such tests would yield an extremely conservative overall statistical significance threshold of $p < .05^3 = .000125$ (Fisher’s method of estimating the conjoint significance of independent tests Fisher, 1950).

Finally, we identified brain regions where the response to trinomodal stimulation exceeded the sum of bimodal and unimodal responses. Thus, we tested for superadditive effects for trinomodal relative to bimodal stimulation. For this purpose, we first computed three separate contrasts, each contrasting the sum of trinomodal and baseline condition to the sum of one bimodal condition and the thus missing unimodal condition (e.g., OAV + BL > OA + V, see Table 1).

We then entered these contrasts into a conjunction analysis to isolate brain regions exhibiting trinomodal effects relative to any combination of bimodal and unimodal sensory input. The identified regions could be therefore labeled as modality-independent trinomodal processing sites. Finally, we computed the overlap between modality-independent bimodal and modality-independent trinomodal processing sites by means of a conjunction analysis over all bimodal and trinomodal contrasts (see Table 1). We thereby isolated brain regions demonstrating multisensory processing independent of the number and type of integrated sensory modalities.

2.7 Data, code, and stimulus availability

Due to the conditions of our ethics approval, public archiving of data from this study is not permitted. However, any individual that seeks access to the data with an explicit research purpose by contacting the corresponding authors will be granted access to the anonymized data without restrictions. Auditory and visual stimulus material, as well as analysis and experiment presentation code are freely and publicly available at https://osf.io/MQXHY/ (DOI 10.17605/OSF.IO/MQXHY). We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. No part of this study was pre-registered.

3. Results

3.1 Unimodal analyses

Unimodal contrasts for each of the studied modalities showed robust activations relative to baseline in their respective primary and secondary perceptual processing areas. See Appendix D Figure D.1 and Table D.1 for detailed activations.

3.2 Bimodal processing sites

To identify brain regions exhibiting a superadditive profile for all types of bimodal stimulation, we first computed separate contrasts assessing superadditive activation for each bimodal condition relative to its unimodal counterparts and subsequently assessed the overlap between these contrasts. The results for modality-specific bimodal effects that are reported below are all corrected for multiple comparisons across whole-brain volume using FWE ($p < .05$ corrected for multiple comparisons using Gaussian random field theory). For olfactory-auditory stimulation, superadditive activation was found in the right superior medial prefrontal gyrus (sPFC), and the anterior part of the cingulate cortex (ACC) (Appendix A Table A.1).

Olfactory-visual stimulation resulted in superadditive activation of the left intraparietal sulcus (IPS) (Appendix A Table A.1).

Auditory-visual stimulation activated the bilateral superior frontal gyrus (SFG) and middle temporal gyrus (MTG) in a superadditive fashion. In the left hemisphere, the IPS, MFG and IFG, middle orbital gyrus, and middle cingulate cortex (MCC) exhibited superadditive activation. Additionally, the right SFG, MTG and IPC/IPS (Appendix A Table A.1) were activated in a superadditive fashion.

To identify brain regions that consistently exhibited multisensory processing in terms of superadditive activation...
Across all bimodal conditions, we assessed the overlap of all three modality-specific bimodal contrasts. This conjunction analysis showed modality-independent superadditive activation in the left IPC/IPS, MFG, IFG, as well as right smPFC (Fig. 2A, Table 2) during bimodal stimulation.

### 3.3. Trimodal processing sites

Trimodal processing sites, that means brain regions responding in a superadditive manner to trimodal stimulation relative to their added bimodal and unimodal responses, were identified in a two-stage procedure. First, we separately contrasted the trimodal condition to the sum of activation elicited by each bimodal condition and its respective unimodal complement.

Bilateral IPC/IPS, IFG and MTG, as well as the right SFG, and the left MFG and middle orbital gyrus showed superadditive activation during trimodal stimulation relative to the sum of olfactory-auditory and visual stimulation (Appendix B Table B1).

Superadditive effects of trimodal stimulation over the sum of olfactory-visual and auditory stimulation were found in the right precentral gyrus, the left MFG, IFG, ACC, MCC, precuneus, and posterior-medial frontal gyrus, as well as bilateral IPC/IPS, MTG, and SFG (Appendix B Table B1).

Activation elicited during trimodal stimulation surpassed the sum of auditory-visual and olfactory stimulation in the left IFG and IPC/IPS (Appendix B Table B1).

To reveal brain regions exhibiting a superadditive response profile for all three trimodal contrasts, we computed a conjunction analysis across all three trimodal contrasts outlined above. Overlapping activation was present in bilateral IPC/IPS and left IFG (Fig. 2B, Table 2).

### 3.4. Core multisensory processing sites

To isolate brain regions demonstrating multisensory processing independent of the number and type of stimulated sensory modalities, we computed the overlap between modality-independent bimodal and modality-independent trimodal processing sites by means of a conjunction analysis over all bimodal and trimodal contrasts (Fig. 2C, Table 2). This revealed overlapping activation in the left IPC/IPS and IFG.

### 3.5. Behavioral data

The analysis of participants’ responses during the fMRI experiment revealed a main effect of number of stimulated

---

**Table 2** — Conjunction analysis of modality-independent bimodal and trimodal processing. Contrasts resulted from a random-effects GLM (ts > 3.16, uncorrected, ps < .001, k > 10). Stereotaxic coordinates of local maxima of activation are expressed as x; y; z values in MNI space. Numbers in parenthesis indicate Brodmann areas (BA).

| Brain region | Hemisphere | Size | t     | x     | y     | z     |
|--------------|------------|------|-------|-------|-------|-------|
| **Bimodal processing** |            |      |       |       |       |       |
| (OA + BL > O + A) ∩ (OV + BL > O + V) ∩ (AV + BL > A + V) | L          | 28   | 3.75  | −54   | −40   | 46    |
| Inferior parietal cortex/intraparietal sulcus (BA 40) | L          | 22   | 4.2   | −33   | 17    | 37    |
| Superior medial frontal gyrus (BA 9) | R          | 43   | 3.88  | 60    | −31   | 49    |
| Middle frontal gyrus | L          | 31   | 4.92  | −48   | 17    | −2    |
| Inferior frontal gyrus (BA 44, 45) | L          | 31   | 4.92  | −48   | 17    | −2    |
| **Trimodal processing** |            |      |       |       |       |       |
| (OAV + BL > OA + V) ∩ (OAV + BL > A + V) ∩ (OAV + BL > AV + O) | L          | 32   | 3.75  | −54   | −40   | 46    |
| Inferior parietal cortex/intraparietal sulcus (BA 40) | R          | 33   | 3.16  | 10    | 1     |
| Inferior frontal gyrus (BA 44, 45) | L          | 24   | 3.75  | −54   | −40   | 46    |
| **Superadditive processing independent of the number and type of stimulated channels** |            |      |       |       |       |       |
| (OA + BL > O + A) ∩ (OV + BL > O + V) ∩ (AV + BL > A + V) ∩ (OAV + BL > OA + V) ∩ (OAV + BL > AV + O) | R          | 33   | 3.16  | 10    | 1     |
sensory modalities on reaction times (RTs) [Wald $\chi^2 (2) = 18.41, p < .001$]. Pairwise comparisons showed significantly faster RTs for trimodal conditions compared to both unimodal [$t(15) = -4.15, p = .001$] and bimodal [$t(15) = -3.21, p = .006$] conditions. A main effect of modality was also found on the accuracy of subjects’ performances [Wald $\chi^2 (2) = 7.69, p = .021$], but pairwise comparisons with Bonferroni correction did not reach significance.

4. Discussion

By presenting participants with the sight, sound and smell of everyday stimuli in various combinations, the present study aimed to identify a modality-independent neural correlate for processing of meaningful multisensory object information, and describe possible changes in its activation pattern depending on the number of activated senses. In line with previous work, overlapping activations across all bimodal conditions were hereby observed in the inferior parietal and inferior frontal cortices. Given that sensory-specific effects were removed during analyses, we conclude that these activations constitute a modality-independent neural correlate of multisensory processing that is shared across all stimulated senses. While IFG and IPC/IPS were also active when three sensory modalities were simultaneously activated, a number of important differences between bi- and trimodal stimulation emerged independently of the specific modalities involved. Specifically, processing of trimodal sensory information was linked to superadditive bilateral activation of the IPC/IPS, while engagement of the left MFG and right smPFC, which was evident during bimodal stimulation, was absent during trimodal stimulation. Together, these findings highlight the important modality-independent role of IPC/IPS in relating object information from several sensory channels to each other, and indicate that their recruitment increases with the number of channels that are simultaneously present.

A large number of studies to date have converged on the IPC and IPS as central areas for multisensory integration (Bremmer et al., 2001; Calvert et al., 2000; Driver & Noesselt, 2008; Grefkes & Fink, 2005; Macaluso & Driver, 2001; Molholm et al., 2006). It has hereby been proposed that this activation is not a result of mere sensory convergence (Bremmer et al., 2001; Macaluso & Driver, 2001), but can be attributed to processing of integration in terms of a superadditive response. The exact function of intraparietal areas in that process, however, has remained poorly understood. In its role as part of the dorsal attention network (Tang, Wu, & Shen, 2016), the IPS has been discussed to allocate additional attentional resources to the process of integrating multisensory stimuli. For example, Regenbogen et al. (2018) showed that the IPS is recruited for the integration of degraded bimodal naturalistic stimuli (high attentional load) but not for clear ones (low attentional load). Along the same line of reasoning, Kassuba and colleagues (Kassuba et al., 2011) attributed the observed higher activation in the IPC for trimodal relative to bimodal naturalistic stimulation to the increased recruitment of attentional resources for processing of three opposed to two sensory input streams. The present study replicates this result of increased inferior parietal engagement in tri-relative to bimodal processing. It is important, however, to note that our study did not explicitly test for multisensory integration in terms of a unification of multisensory signals into one meaningful whole. We hypothesize that the observed superadditive effect instead reflects IPC/IPS engagement whenever several channels need to be matched against each other for overlapping meaning, regardless of the actual outcome. Whether the here observed superadditive activation can be attributed to unifying or matching sensory inputs can only be conclusively distinguished by demonstrating privileged processing of congruent compared to incongruent inputs. In an exploratory analysis (see Appendix C), we have assessed whether the superadditivity effect is specific to the processing of sensory stimuli matching in meaning, and found no evidence for specificity. While one might interpret this finding in the sense of the “relating inputs” interpretation, these results should be interpreted with caution because they partially depict only statistical trends. Replication of this null result in a larger sample would be advisable.

Our findings further indicate that in addition to the amount of noise in each to-be-integrated stimulus, also the number of to-be-integrated sensory channels may determine the extent to which attentional resources are recruited to aid with multisensory integration. This reasoning is in line with Senkowski and colleagues’ speculation that IPC engagement during multisensory processing could reflect an increased demand in relating multisensory signals (Senkowski et al., 2007). They, as well as others (Hein et al., 2007), observed IPC activation for processing of artificial audio-visual stimuli but not for naturalistic ones and speculated that it reflected an extended search for relatedness between the visual and auditory signal in the artificial stimuli. Senkowski and colleagues reasoned that an artificial multisensory stimulus does not entail an intrinsically meaningful relation between the individual sensory components, rendering their processing more demanding. A summary interpretation of these findings might thus propose that with increasing dimensionality of a stimulus more information needs to be related to each other, which requires more attentional resources. In turn, allocation of these resources would then be controlled by enhanced recruitment of inferior parietal cortex.

In line with this idea, the crucial role of the parietal cortex in relaying complex sensory input has further been demonstrated outside the multisensory literature; for example, through tasks studying the grouping of individual components into a global holistic percept (Zaretskaya & Bartels, 2011). Activation of the IPS was found when a larger number of moving dots were perceived as two moving squares instead of ungrouped, individually moving dots. Moreover, temporal disruption of the IPS through transcranial magnetic stimulation inhibited the holistic percept of moving squares while leaving the perception of moving dots unaffected. These findings suggest that a role of IPS activation might be to place individual perceptual components, e.g., moving dots, in relation to each other, which allows for integration into a holistic whole. Disruption of IPS activation, in return, might prevent linkage of individual parts, rendering a holistic percept impossible. Indeed, it has previously been proposed that the IPS relates and integrates multisensory stimuli by inferring whether the individual signals arise from one common source.
or independent sources. It does so by taking the reliability of the individual signals as well as the uncertainty about the world's causal structure into account (Rohe & Noppeney, 2015).

In addition to intraparietal areas, also the left IFG exhibited superadditive activation during integration of both bimodal and trimodal stimuli in our study, which might reflect a linking process between multisensory object representations and concepts in semantic memory (Amedi et al., 2005). Apart from these commonalities, we observed differences in the activation of areas within the prefrontal cortex during bimodal and trimodal integration. Most importantly, bimodal, but not trimodal, sensory integration activated the left MFG and right smPFC. While the reasons for this additional activation remain speculative, it might be related to the performed task. The task of identifying the number of stimulated senses was more difficult to accomplish in the bimodal condition than in the trimodal condition because it was more difficult to evaluate if one of three sensory modalities was not stimulated than to identify that all three senses received stimulation. It is possible that this increased task demand might have required additional engagement of prefrontal areas.

As a limitation of the present study, it should be noted that although we found strong activations for all uni-, bi-, and trimodal stimuli using a conservative integration criterion of required superadditivity, replication in a larger sample would be desirable to more reliably quantify the extent of the observed effects and to assess individual differences.

It is notable that, in contrast to previous work (e.g., Kassuba et al., 2011; Regenbogen et al., 2018), our results did not support multisensory integration at the level of unisensory cortices or in commonly reported integration sites, such as superior temporal or cingulate regions.

The reasons for this absence remain speculative. One possibility is that activation in these regions does not generalize across modalities, but instead is specific to the sensory modalities involved in the integration process. This view is supported by the observation that anterior and middle cingulate cortices, as well as middle temporal regions, exhibited significant superadditive activation for some modality-specific bimodal contrasts, but did not survive the modality-independent conjunction analysis.

Another possibility is that the use of a superadditivity criterion in the present study provides a conservative estimate of modality-specific and modality-independent multisensory processes. Our rather strict statistical criterion, combined with a relatively small sample size may have allowed for identification of brain regions exhibiting strong multisensory effects, but resulted in false negative findings, or in other words, a failure to detect brain regions that were actually engaged in multisensory processes but exhibited effects below our threshold. This point is especially relevant considering that our stimuli were presented in a clearly perceivable manner, which usually evokes weaker multisensory effects (Meredith, Nemitz, & Stein, 1987; Regenbogen et al., 2018; Stein & Meredith, 1993). Future replication of the observed effects in a larger sample would be desirable to assess individual differences and to more reliably quantify the extent of the observed effects.

In summary, our results demonstrate that processing of real-world multisensory stimuli engages left inferior frontal and intraparietal areas independently of the specific senses that are stimulated. This indicates that they likely represent modality-independent hubs for multisensory processing, where activation differences stem from the increased processing demand for relating and integrating the components of multisensory stimuli, with processing of more sensory input streams being more demanding. The specific function of such modality-independent processing in relation to more sensory-specific processing and to direct communication between primary sensory cortices remains to be explored and promises to yield valuable insights into the formation of perceptual objects in real-world settings.

### Authors' contributions

Danja K. Porada: Data curation, Formal analysis, Visualization, Writing - original draft, Writing - review & editing.

Christina Regenbogen: Formal analysis, Supervision, Validation, Visualization, Writing - review & editing.

Jessica Freiherr: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing - review & editing.

Janina Seubert: Conceptualization, Supervision, Writing - review & editing.

Johan N. Lundström: Conceptualization, Funding acquisition, Investigation, Methodology, Resources, Supervision, Writing - review & editing.

### Funding

This work was supported by the Knut and Alice Wallenberg Foundation (KAW 2018.0152), project grants from the Swedish Research Council awarded to JNL (2014-1346) and JS (2018-01305), and DAAD postdoctoral fellowships awarded to JF (D/08/40252) and CR (91518477).

### Open practices

The study in this article earned an Open Materials badge for transparent practices. Auditory and visual stimulus material, as well as analysis and experiment presentation code are freely and publicly available at [https://osf.io/MQXHY/](https://osf.io/MQXHY/) (DOI 10.17605/OSF.IO/MQXHY).

### Declaration of competing interest

None.

### Acknowledgements

We would like to thank Eva C. Alden, Amy M. Gordon, Valentin A. Schriever, and Sanne Boesveldt for assistance during data collection.
Appendix A

Sensory-specific bimodal processing

Table A.1 – Brain activation evoked by odor-sound, odor-video, and sound-video stimulation. Contrasts resulted from a random-effects GLM (ts > 4.79, ps < .05, whole-brain-corrected for multiple comparisons (FWE), k > 10 voxels). Stereotaxic coordinates of local maxima of activation are expressed as x; y;z values in MNI space. Numbers within parenthesis indicate Brodmann areas (BA).

| Contrast      | Brain region                                      | Hemisphere | Size | t   | x   | y   | Z   |
|---------------|---------------------------------------------------|------------|------|-----|-----|-----|-----|
| OA + BL > O + A | Superior medial gyrus/Anterior cingulate cortex | R          | 10   | 5.7 | 12  | 56  | 28  |
|               | Anterior cingulate cortex                         | R          | 11   | 5.6 | 15  | 47  | 13  |
| OV + BL > O + V | Intraparietal sulcus                              | L          | 22   | 5.5 | −51 | −40 | 43  |
| AV + BL > A + V | Superior frontal gyrus (BAs 6, 44, 45)           | L          | 800  | 7.7 | 18  | 53  | 28  |
|                | Middle temporal gyrus                             | L          | 158  | 7.8 | 57  | −43 | −2  |
|                | Intraparietal sulcus                              | L          | 405  | 7.6 | −51 | −58 | 43  |
|                | Middle temporal gyrus                             | R          | 117  | 7.3 | −57 | −37 | −8  |
|                | Middle cingulate cortex                           | L          | 12   | 5.5 | −3  | −19 | 37  |
|                | Middle frontal gyrus                              | L          | 37   | 6.8 | −24 | 47  | 31  |
|                | Superior frontal gyrus                            | L          | 58   | 6.7 | −18 | 23  | 55  |
|                | Supramarginal gyrus                               | R          | 91   | 6.6 | 57  | −46 | 34  |
|                | Middle frontal gyrus                              | L          | 137  | 6.6 | −36 | 20  | 40  |
|                | Middle frontal gyrus                              | L          | 71   | 6.0 | −27 | 56  | 10  |
|                | Middle orbital gyrus                              | L          | 12   | 6.0 | −24 | 44  | −17 |
|                | Middle cingulate cortex                           | L          | 12   | 5.5 | −3  | −19 | 37  |
|                | Superior frontal gyrus                            | L          | 5.4  | −51 | 20  | 1   |

Appendix B

Sensory-specific trimodal processing

Table B.1 – Brain activation evoked by trimodal stimulation. Contrasts resulted from a random-effects GLM (ts > 4.79, ps < .05, whole-brain-corrected for multiple comparisons (FWE), k > 10 voxels). Stereotaxic coordinates of local maxima of activation are expressed as x; y;z values in MNI space. Numbers within parenthesis indicate Brodmann areas (BA).

| Contrast      | Brain region                                      | Hemisphere | Size | t   | x   | y   | Z   |
|---------------|---------------------------------------------------|------------|------|-----|-----|-----|-----|
| OAV + BL > OA + V | Superior frontal gyrus (BA 44)                         | R          | 39   | 6.9 | 18  | 53  | 31  |
|                | Inferior parietal cortex/Intraparietal sulcus        | L          | 456  | 8.2 | −45 | −58 | 46  |
|                | Inferior parietal cortex/Intraparietal sulcus        | R          | 338  | 7.3 | 51  | −37 | 43  |
|                | Middle frontal gyrus                                | L          | 18   | 6.2 | −33 | 38  | 19  |
|                | Middle temporal gyrus                               | R          | 155  | 6.7 | 60  | −40 | −5  |
|                | Inferior frontal gyrus                              | R          | 16   | 5.7 | 48  | 41  | −14 |
|                | Middle temporal gyrus                               | L          | 53   | 6.5 | −60 | −34 | −11 |
|                | Middle orbital gyrus                                | L          | 99   | 6.2 | −33 | 50  | −5  |
|                | Superior frontal gyrus                              | R          | 11   | 5.6 | 18  | 23  | 55  |
|                | Superior frontal gyrus                              | L          | 16   | 6.0 | −51 | 17  | 1   |
|                | Inferior parietal cortex/Intraparietal sulcus        | R          | 216  | 7.5 | 21  | 53  | 31  |
|                | Inferior parietal cortex/Intraparietal sulcus        | R          | 155  | 7.2 | 57  | −43 | 31  |
|                | Inferior parietal cortex/Intraparietal sulcus        | L          | 201  | 5.9 | −48 | −52 | 43  |
|                | Middle temporal gyrus                               | R          | 175  | 6.7 | 54  | −34 | −8  |
|                | Inferior frontal gyrus                              | L          | 18   | 5.7 | −51 | 17  | −2  |
|                | Precuneus                                           | L          | 11   | 5.6 | −6  | 73  | 37  |
|                | Superior frontal gyrus                              | R          | 128  | 6.4 | 18  | 23  | 55  |
|                | Anterior cingulate cortex                           | L          | 87   | 6.4 | −6  | 47  | 7   |
|                | Middle temporal gyrus                               | L          | 41   | 6.3 | −57 | −34 | −14 |
|                | Middle cingulate cortex                             | L          | 57   | 6.2 | −3  | −28 | 37  |
|                | Superior frontal gyrus                              | L          | 16   | 5.9 | −21 | 50  | 31  |
|                | Precentral gyrus                                    | R          | 29   | 5.8 | 45  | 5   | 46  |
|                | Middle frontal gyrus                                | L          | 10   | 5.5 | −33 | 41  | 22  |
|                | Posterior-medial frontal gyrus                      | L          | 19   | 5.3 | −3  | 2   | 49  |
|                | Middle frontal gyrus                                | L          | 21   | 5.3 | −30 | 56  | 10  |
| OAV + BL > AV + O | Inferior parietal cortex/Intraparietal sulcus     | L          | 1   | 4.8 | −54 | −37 | 46  |
|                | Inferior frontal gyrus (BA 44, 45)                  | L          | 4    | 5.3 | −48 | 17  | 1   |

* Note that these activations represent statistical trends with clusters smaller than k = 10.
Appendix C

**Congruency-sensitivity of multisensory processing**

As an additional exploratory analysis, we aimed to determine whether the observed modality-independent superadditive effects would be sensitive to the congruency of the sensory signals. We therefore assessed whether superadditive activation could also be observed for incongruent bimodal and trimodal stimulation. To this end, we first computed separate contrasts assessing superadditive activation for each incongruent bimodal condition relative to its unimodal counterparts (e.g., $O_A + BL > O + A$) and subsequently assessed the overlap between these contrasts. For incongruent trimodal stimulation, we followed the same analysis procedure as for the bimodal conditions. Meaning, we first separately contrasted the incongruent trimodal condition to the sum of activation elicited by each congruent bimodal condition and its respective unimodal complement (e.g., $OAV + BL > O + V$), and afterwards computed a conjunction analysis across all three contrasts. To specifically test for an effect in the previously identified core multisensory processing sites, we applied an inclusive mask consisting of a conjunction across the superadditive contrasts for congruent bimodal and trimodal stimulation ($p < .05$, uncorrected). We observed significant superadditive activation in the left IPC/IPS (statistical threshold of $p < .001$ uncorrected for multiple comparisons, for each contrast entering the conjunction analysis) for incongruent trimodal stimulation, and a trend towards superadditive activation for incongruent bimodal stimulation. These results suggest that the superadditive effects we observed in the left inferior parietal cortex are not sensitive to the congruency of the incoming sensory signals. Although these results might be surprising, they are in line with the interpretation that the superadditive activation in IPC/IPS reflects a process of relating multisensory signals to each other to determine whether they arise from a common source. As this process would be independent of the congruency between the signals, congruent and incongruent inputs should require an equal amount of cognitive resources. Given that the stimuli in our study entailed an intrinsically meaningful relation, merely the number of sensory input streams should have modulated the demand placed on the IPC/IPS. Hence, superadditive responses should be equally observable for congruent as well as incongruent multisensory signals as long as they match in the number of stimulated senses.

Appendix D

**Neural correlates of unimodal sensory processing**

The main effects of unimodal stimulation (odor, sound, and video) demonstrate that our experimental paradigm induced

---

Table C.1 – Contrast descriptions of group-level contrasts used to assess the sensitivity of the superadditivity effect to the congruency of the sensory signals. Incongruent stimulus combinations are labeled with $i$ (e.g., $OAV_i$), while congruent stimulus combination are labeled with $c$ (e.g., $OAV_c$).

| Contrasts | Descriptions |
|-----------|--------------|
| Congruency-sensitivity of bimodal processing | $(OAI + BL > O + A) \cap (OVI + BL > O + V) \cap (AVI + BL > A + V)$ masked with $(OAVc + BL > OAc + V) \cap (OAVc + BL > OVc + A) \cap (OAVc + BL > AVc + O) \cap (OAc + BL > O + A) \cap (OVc + BL > O + V) \cap (AVc + BL > A + V)$ |
| Congruency-sensitivity of trimodal processing | $(OAVi + BL > OAc + V) \cap (OAVi + BL > OVc + A) \cap (OAVi + BL > AVc + O)$ masked with $(OAVc + BL > OAc + V) \cap (OAVc + BL > OVc + A) \cap (OAVc + BL > AVc + O) \cap (OAc + BL > O + A) \cap (OVc + BL > O + V) \cap (AVc + BL > A + V)$ |

Fig. D.1 – Unimodal sensory processing. Brain activation evoked by unimodal olfactory (red), auditory (blue), and visual (green) stimulation, compared to baseline, respectively ($t_s > 4.79, p_s < .05$, whole-brain-corrected for multiple comparisons (FWE), $k > 10$ voxels).
activation in all three modalities and thereby assured that the experimental design was sufficient to study multisensory processes. Unimodal odor stimulation showed significant activation in typical olfactory cortex areas, namely bilateral piriform cortex, bilateral amygdala, bilateral orbitofrontal cortex, right insula and areas associated with higher olfactory processes, namely left middle and superior frontal gyrus, left posterior cingulate cortex, right precuneus, and left postcentral gyrus. Unimodal sound stimulation showed activation in bilateral superior temporal gyrus. Unimodal video stimulation showed activation in primary visual cortex (bilateral middle occipital gyrus, right calcarine gyrus) and bilateral hippocampi (all ps < 0.05, corrected for multiple comparisons, k > 10, Figure D.1; Table D.1).

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2021.03.008.

REFERENCES

Amedi, A., von Kriegstein, K., van Attervelt, N. M., Beauchamp, M. S., & Naumer, M. J. (2005). Functional imaging of human crossmodal identification and object recognition. Experimental Brain Research, 166(3–4), 559–571. https://doi.org/10.1007/s00221-005-2396-5

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. Neuroimage, 26(3), 839–851. https://doi.org/10.1016/j.neuroimage.2005.02.018

Beauchamp, M. S. (2005). Statistical criteria in fMRI studies of multisensory integration. Neuroinformatics, 3(2), 93–113. https://doi.org/10.1385/NI:3:2:093

Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyun, J. H., & Martin, A. (2004). Unraveling multisensory integration: Patchy organization within human STS multisensory cortex. Nature Neuroscience, 7(11), 1190–1192. https://doi.org/10.1038/nn1333

Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. Neuron, 41(5), 809–823. https://doi.org/10.1016/s0896-6273(04)00070-4

Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K., et al. (2001). Polymodal motion processing in posterior parietal and premotor cortex: A human fMRI study strongly implies equivalences between humans and monkeys. Neuro, 29(1), 287–296. https://doi.org/10.1016/s0896-6273(01)00198-2

Calvert, G. A. (2001). Crossmodal processing in the human brain: Insights from functional neuroimaging studies. Cerebral Cortex, 11(12), 1110–1123. https://doi.org/10.1093/cercor/11.12.1110

Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Current Biology, 10(11), 649–657. https://doi.org/10.1016/S0960-9822(00)00513-3

Deichmann, R., Gottfried, J. A., Hutton, C., & Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. Neuroimage, 19(2 Pt 1), 430–441. https://doi.org/10.1016/s1053-8119(03)00073-9

Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on "sensory-specific" brain regions, neural responses, and judgments. Neuron, 57(1), 11–23. https://doi.org/10.1016/j.neuron.2007.12.013

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. Nature, 415(6870), 429–433. https://doi.org/10.1038/415429a

Felsen, G., & Dan, Y. (2005). A natural approach to studying vision. Nature Neuroscience, 8(12), 1643–1646. https://doi.org/10.1038/nn1608

Fisher, R. A. (1950). Statistical methods for research workers. Statistical Methods for Research Workers.

Foxe, J. J., & Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. Neuroreport, 16(5), 419–423. https://doi.org/10.1097/00001756-200504040-00001

Freiherr, J., Gordon, A. R., Alden, E. C., Ponting, A. L., Hernandez, M. F., Boesveldt, S., et al. (2012). The 40-item monell extended sniffin’ sticks identification test (MONEX-40). Journal of Neuroscience Methods, 205(1), 10–16. https://doi.org/10.1016/j.jneumeth.2011.12.004

Grefkes, C., & Fink, G. R. (2005). The functional organization of the intraparietal sulcus in humans and monkeys. Journal of Anatomy, 207(1), 3–17. https://doi.org/10.1111/j.1469-7580.2005.00426.x

Hein, G., Doehrmann, O., Müller, N. G., Kaiser, J., Muckli, L., & Naumer, M. J. (2007). Object familiarity and semantic congruency modulate responses in cortical audiovisual integration areas. The Journal of Neuroscience, 27(30), 7881–7887. https://doi.org/10.1523/JNEUROSCI.1740-07.2007
Kassuba, T., Klinge, C., Holig, C., Menz, M. M., Ptito, M., Röder, B., & et al. (2011). The left fusiform gyrus hosts trisensory representations of manipulable objects. *Neuroimage*, 56(3), 1566–1577. https://doi.org/10.1016/j.neuroimage.2011.02.032

Kayser, C., Körding, K. F., & König, P. (2004). Processing of complex stimuli and natural scenes in the visual cortex. *Current Opinion in Neurobiology*, 14(4), 468–473. https://doi.org/10.1016/j.conb.2004.06.002

Lewis, R., & Noppeney, U. (2010). Audiovisual synchrony improves motion discrimination via enhanced connectivity between early visual and auditory areas. *The Journal of Neuroscience*, 30(37), 12329–12339. https://doi.org/10.1523/JNEUROSCI.5745-09.2010

Lundström, J. N., Gordon, A. R., Alden, E. C., Boesveldt, S., & Albrecht, J. (2010). Methods for building an inexpensive computer-controlled olfactometer for temporally-precise experiments. *International Journal of Psychophysiology*, 78(2), 179–189. https://doi.org/10.1016/j.ijpsycho.2010.07.007

Lundström, J. N., Regenbogen, C., Ohla, K., & Seubert, J. (2018). Prefrontal control over occipital responses to crossmodal overlap varies across the congruency spectrum. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhy168

Macaluso, E., & Driver, J. (2001). Spatial attention and crossmodal interactions between vision and touch. *Neuropsychologia*, 39(12), 1304–1316.

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *The Journal of Neuroscience*, 7(10), 3215–3229. https://doi.org/10.1523/JNEUROSCI.07-10-03215.1987

Molholm, S., Sehatpour, P., Mehta, A. D., Shpaner, M., Gomez-Ramirez, M., Ortigue, S., et al. (2006). Audio-visual multisensory integration in superior parietal lobule revealed by human intracranial recordings. *Journal of Neurophysiology*, 96(2), 721–729. https://doi.org/10.1152/jn.00285.2006

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage*, 25(2), 653–660. https://doi.org/10.1016/j.neuroimage.2004.12.005

Noppeney, U., Josephy, O., Hocking, J., Price, C. J., & Friston, K. J. (2008). The effect of prior visual information on recognition of speech and sounds. *Cerebral Cortex*, 18(3), 598–609. https://doi.org/10.1093/cercor/bhn091

Porada, D. K., Regenbogen, C., Seubert, J., Freiherr, J., & Lundström, J. N. (2019). Multisensory enhancement of odor object processing in primary olfactory cortex. *Neuroscience*, 418, 254–265. https://doi.org/10.1016/j.neuroscience.2019.08.040

Regenbogen, C., Seubert, J., Johansson, E., Finkelmeyer, A., Andersson, P., & Lundström, J. N. (2018). The intraparietal sulcus governs multisensory integration of audiovisual information based on task difficulty. *Human Brain Mapping*, 39(3), 1313–1326. https://doi.org/10.1002/hbm.23918

Roe, T., & Noppeney, U. (2015). Cortical hierarchies perform Bayesian causal inference in multisensory perception. *Plos Biology*, 13(2), Article e1002073. https://doi.org/10.1371/journal.pbio.1002073

Senkowski, D., Saint-Amour, D., Kelly, S. P., & Foxe, J. J. (2007). Multisensory processing of naturalistic objects in motion: A high-density electrical mapping and source estimation study. *Neuroimage*, 36(3), 877–888. https://doi.org/10.1016/j.neuroimage.2007.01.053

Seubert, J., Ohla, K., Yokomukai, Y., Kellermann, T., & Lundström, J. N. (2014). Supraadditive opercular activation to food flavor is mediated by enhanced temporal and limbic coupling. *Human Brain Mapping*, 36(5), 1662–1676. https://doi.org/10.1002/hbm.22728

Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *NeuroReport*, 12(1), 7–10. https://doi.org/10.1097/00001756-200101220-00009

Stein, B. E., & Meredith, M. A. (1993). *The Merging of the Senses*. MIT Press.

Stein, B. E., Stanford, T. R., & Rowland, B. A. (2014). Development of multisensory integration from the perspective of the individual neuron. *Nature Reviews. Neuroscience*, 15(8), 520–535.

Stevenson, R. A., Geoghegan, M. L., & James, T. W. (2007). Superadditive BOLD activation in superior temporal sulcus with threshold non-speech objects. *Experimental Brain Research*, 179(1), 85–95. https://doi.org/10.1007/s00221-006-0770-6

Tang, X., Wu, J., & Shen, Y. (2016). The interactions of multisensory integration with endogenous and exogenous attention. *Neuroscience and Biobehavioral Reviews*, 61, 208–224. https://doi.org/10.1016/j.neubiorev.2015.11.002

Zaretskaya, N., & Bartels, A. (2011). Parietal cortex mediates perceptual grouping across space: Evidence from fMRI and TMS.