Behavioral/Cognitive

Distinct Neural Mechanisms of Spatial Attention and Expectation Guide Perceptual Inference in a Multisensory World

Arianna Zuanazzi and Uta Noppeney

Computational Cognitive Neuroimaging Laboratory, Computational Neuroscience and Cognitive Robotics Centre, University of Birmingham, B15 2TT Birmingham, United Kingdom

Spatial attention (i.e., task-relevance) and expectation (i.e., signal probability) are two critical top-down mechanisms guiding perceptual inference. Spatial attention prioritizes processing of information at task-relevant locations. Spatial expectations encode the statistical structure of the environment. An unresolved question is how the brain allocates attention and forms expectations in a multisensory environment, where task-relevance and signal probability over space can differ across sensory modalities. We used functional magnetic resonance imaging in human participants (female and male) to investigate whether the brain encodes task-relevance and signal probability over space separately or interactively across sensory modalities. In a novel multisensory paradigm, we manipulated spatial attention and expectation selectively in audition and assessed their effects on behavioral and neural responses to auditory and visual stimuli. Our results show that both auditory and visual stimuli increased activations in a right-lateralized frontoparietal system, when they were presented at locations that were task-irrelevant in audition. Yet, only auditory stimuli increased activations in the medial prefrontal cortex when presented at expected locations and in audiovisual and frontoparietal cortices signaling a prediction error when presented at unexpected locations. This dissociation in multisensory generalization for attention and expectation effects shows that the brain controls attentional resources interactively across the senses but encodes the statistical structure of the environment as spatial expectations independently for each sensory system. Our results demonstrate that spatial attention and expectation engage partly overlapping neural systems via distinct mechanisms to guide perceptual inference in a multisensory world.

Key words: attention; expectation; fMRI; multisensory; perceptual decisions; space

---

Significance Statement

In our natural environment the brain is exposed to a constant influx of signals through all our senses. How does the brain allocate attention and form spatial expectations in this multisensory environment? Because observers need to respond to stimuli regardless of their sensory modality, they may allocate attentional resources and encode the probability of events jointly across the senses. This psychophysics and neuroimaging study shows that the brain controls attentional resources interactively across the senses via a frontoparietal system but encodes the statistical structure of the environment independently for each sense in sensory and frontoparietal areas. Thus, spatial attention and expectation engage partly overlapping neural systems via distinct mechanisms to guide perceptual inference in a multisensory world.

Introduction

Spatial attention (i.e., task-relevance) and expectation (i.e., signal probability) are two critical top-down mechanisms that guide

perceptual inference. Spatial attention prioritizes signal processing at locations that are relevant for the observer's goals. Spatial expectations encode the event probability over space, i.e., the statistical structure of the environment (Summerfield and Egner, 2009).

Behaviorally, both spatial attention and expectation typically facilitate perception leading to faster and more accurate responses for stimuli presented at attended and/or expected locations (Posner et al., 1980; Downing, 1988; Geng and Behrmann, 2002, 2005; Doherty et al., 2005; Carrasco, 2011). At the neural level, spatial attention is thought to increase stimulus-evoked

responses at task-relevant locations (Tootell et al., 1998; Brefczynski and DeYoe, 1999; Bressler et al., 2013), whereas expectations often reduce stimulus-evoked responses (Summerfield et al., 2008; Alink et al., 2010; Kok et al., 2012a; but see Kok et al., 2012b). Importantly, spatial attention and expectations are intimately related (Zuanazzi and Noppeney, 2018). In many situations observers will allocate attentional resources to locations where events are likely to occur (Summerfield and Egner, 2009). Likewise, the majority of previous paradigms, most prominently the classical Posner paradigm (Posner, 1980), manipulated observer's endogenous spatial attention via probabilistic cues that indicate where a task-relevant target is likely to occur. Only recently has unisensory research attempted to dissociate attention and expectation (Shulman et al., 2009; Doricchi et al., 2010; Kok et al., 2012b; Auksztulewicz and Friston, 2015). Recent accounts of predictive coding suggest that attention may increase the precision of prediction errors that are elicited when expectations are violated (Feldman and Friston, 2010; Auksztulewicz and Friston, 2015).

Crucially, in our natural environment the brain is exposed to a constant influx of signals furnished by all our senses. This raises the critical question of how the brain allocates spatial attention and forms spatial expectations in a multisensory environment. Because observers need to respond to stimuli regardless of the sense by which they are perceived, they may allocate attentional resources interactively across the senses and form an "amodal map" that encodes the probability of events. In line with this conjecture, parietal cortices have previously been shown to integrate audiovisual signals weighted by their bottom-up sensory reliabilities and top-down task-relevance into audiovisual spatial priority maps (Rohe and Noppeney, 2015, 2016). Likewise, attentional resources were shown to be allocated interactively across the senses. Shifts in spatial attention that were endogenously or exogenously induced in one sensory modality affected stimulus processing in other sensory systems (Spence and Driver, 1996, 1997; Eimer and Schröger, 1998; Eimer, 1999; McDonald et al., 2000; Spence et al., 2000; Ward et al., 2000). Regardless of stimulus modality, reorienting of spatial attention was associated with activations in ventral and to some extent dorsal frontoparietal cortices (Corbetta and Shulman, 2002; Wu et al., 2007; Corbetta et al., 2008; Krumbholz et al., 2009; Santangelo et al., 2009; Macaluso, 2010; Santangelo and Macaluso, 2012).

Less is known about how the brain forms spatial expectations across sensory modalities (Stekelenburg and Vroomen, 2012). Because information is initially gathered by distinct sensory organs and enters the brain via parallel pathways, each sensory system may initially encode the probability of signals selectively for its preferred sensory modality. These modality-specific spatial expectations may be reinforced particularly in environments where auditory and visual signals arise from separate sources such as in experiments that present auditory or visual signals independently (Spence and Driver, 1996).

The current study investigated how the brain allocates spatial attention and forms spatial expectations across the senses. Further, we assessed whether spatial attention and expectation rely on distinct or common neural systems and guide perceptual inference via additive or interactive mechanisms. Combining fMRI and a novel multisensory paradigm we orthogonally manipulated spatial attention (i.e., task-relevance) and expectation (i.e., spatial signal probability) selectively in audition and assessed their effects on observers' behavioral and neural responses in audition and vision. We expected attentional resources to be interactively allocated across sensory modalities (Eimer and Schröger, 1998;

Macaluso et al., 2002). By contrast, given the hierarchical organization of multisensory integration, spatial expectations, and prediction errors for unexpected stimuli may be modality-specific in early sensory cortices but shared across the senses in parietal cortices (Rohe and Noppeney, 2015, 2016, 2018).

Materials and Methods

Participants
Thirty-one healthy volunteers (8 males; mean age: 21.4 years; range: 18–27 years) participated in the psychophysics experiment. All participants had normal or corrected to normal vision, reported normal hearing and had no history of neurological or psychiatric illness. All participants were right-handed, according to the Edinburgh Handedness Inventory (Oldfield, 1971; mean laterality index: 84; range: 60–100). A subgroup of 22 participants (5 males; mean age: 21.2 years; range: 18–27 years) was selected to take part in the fMRI experiment (see Inclusion criteria). Data collection was terminated when 22 participants had undergone the fMRI study. This sample size was determined based on Thirion et al. (2007). All participants provided written informed consent, as approved by the local ethics committee of the University of Birmingham (Science, Technology, Mathematics and Engineering Ethical Review Committee) and the experiment was conducted in accordance with these guidelines and regulations.

Inclusion criteria
A subgroup of 22 participants who had taken part in the psychophysics experiment was selected to take part in the fMRI experiment. Inclusion criteria were participants' accuracy and fixation performance in the psychophysics experiment. Only participants who in the psychophysics experiment produced <20 saccades averaged across blocks and showed overall accuracy >95% (calculated as the percentage of hits + correct rejections, pooling over auditory and visual stimuli) were selected for the fMRI experiment.

Stimuli
Auditory spatialized stimuli (100 ms duration) were created by convolving a burst of white noise (with 5 ms onset and offset ramps) with spatially specific head-related transfer functions based on the KEMAR dummy head of the MIT Media Lab (http://sound.media.mit.edu/resources/KEMAR.html; Gardner and Martin, 1995).

Visual stimuli (i.e., the so-called "flashes") were white disks (100 ms duration; radius: 0.88° visual angle, luminance: 165 cd/m^2) presented on a gray background (luminance: 78 cd/m^2). Both auditory and visual stimuli were presented at ±10° visual angle along the azimuth (0° visual angle for elevation). A fixation cross was presented in the center of the screen throughout the entire experiment.

Experimental design
In both the psychophysics and the fMRI experiment, we orthogonally manipulated spatial attention (i.e., task-relevance or response requirement) and expectation (i.e., stimulus probability) across the two hemifields selectively in audition and evaluated their effects on observers' neural and behavioral responses to auditory and visual signals. Thus, the 2 × 2 × 2 × 2 design manipulated auditory spatial attention (left vs right hemifield), auditory spatial expectation (left vs right hemifield), stimulus location (left vs right hemifield) and stimulus modality (auditory vs visual; Fig. 1A). For the behavioral and fMRI data analysis we pooled over stimulus locations (left/right) leading to a 2 (attended vs unattended) × 2 (expected vs unexpected) × 2 (auditory vs visual stimulus modality) factorial design. Across days, auditory spatial expectation was manipulated as spatial signal probability, i.e., the probability for auditory stimuli to be presented in the left or right hemifield. Both the psychophysics and fMRI experiments were preceded by training runs, in which the spatial probability ratio of auditory targets was set to 9:1 for the expected/unexpected hemifields to boost the implicit learning of auditory spatial signal probability. In the psychophysics and fMRI experiments the auditory stimuli were presented with a ratio of 4:1 in the expected/unexpected hemifields. Observers were not explicitly informed about those probabilities. Auditory spatial attention was manipulated as "task-relevance", i.e.,
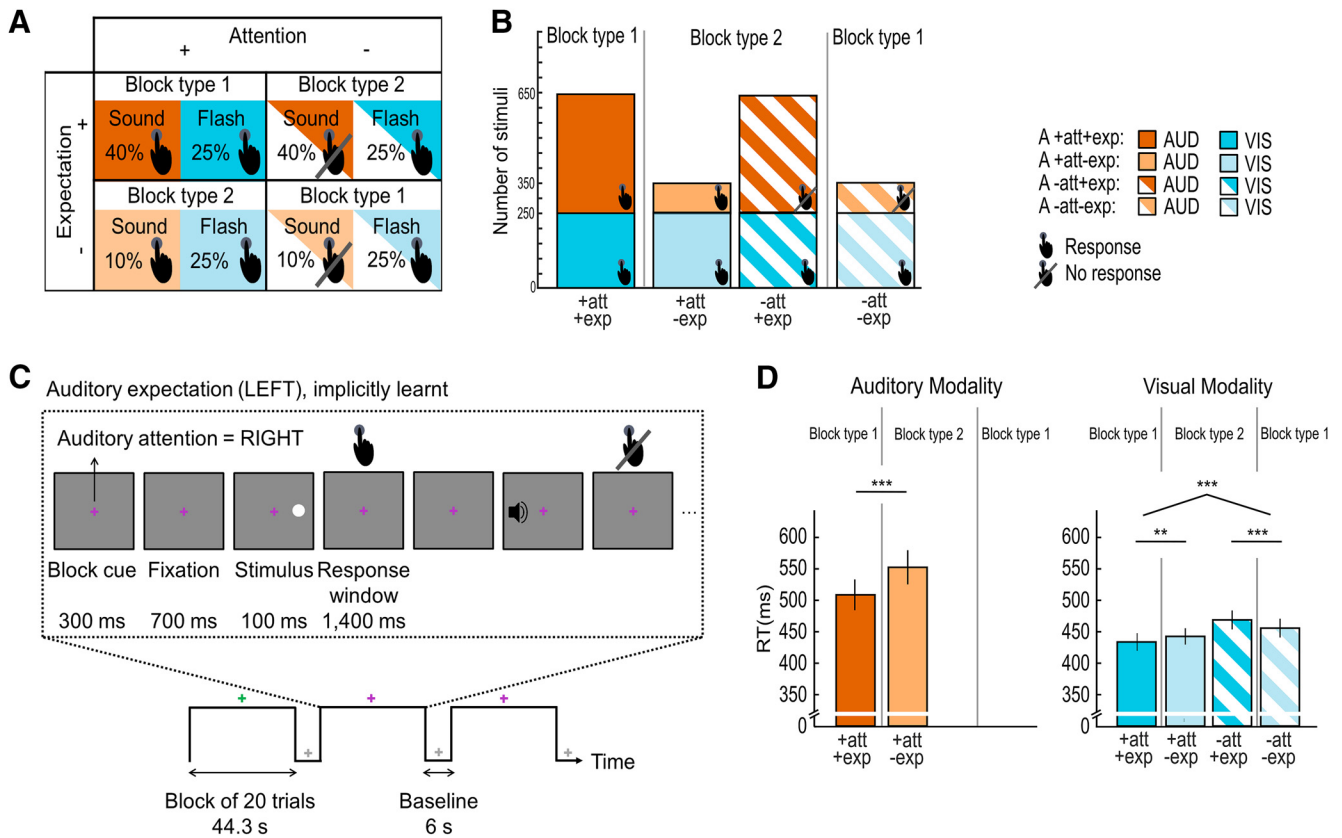
Figure 1. Experimental design, example stimuli of the psychophysics and fMRI experiment and behavioral results of the fMRI experiment. ***A***, The factorial design manipulated: auditory (A) spatial attention (attended hemifield — full pattern, vs unattended hemifield — striped pattern), A spatial expectation (expected hemifield — dark shade, vs unexpected hemifield — light shade) and stimulus modality (auditory modality — orange, vs visual modality — blue). For illustration purposes and analysis, we pooled over stimulus locations (left/right). Presence versus absence of response requirement is indicated by the hand symbol. ***B***, Number of auditory (orange) and visual (blue) trials in the 2 (A attended vs unattended) × 2 (A expected vs unexpected) design. Presence versus absence of response requirement is indicated by the hand symbol. The fraction of the area indicated by the "Response" hand symbol pooled over the two bars of one particular block type (e.g., block type 1) represents the "general response probability" (i.e., the overall probability that a response is required on a particular trial); the general response probability is greater for block type 1 (90%), where attention and expectation are congruent, than block type 2 (60%), where they are incongruent. The fraction of the area indicated by the Response hand symbol for each bar represents the "spatially selective response probability", i.e., the probability that the observer needs to make a response conditioned on the signal being presented in a particular hemifield; the spatially selective response probability is greater when unattended signals are presented in the unexpected (71.4%) than expected (38.4%) hemifield. ***C***, fMRI runs included 10 blocks of 20 trials alternating with fixation periods. A fixation cross was presented throughout the entire run. The colors indicate as follows: white, fixation period; green or pink, activation period with auditory attention directed to the left (or right) hemifield. On each trial participants were presented with an auditory or visual stimulus (100 ms duration) either in their left or right hemifield. They were instructed to respond as fast and accurately as possible with their right index finger within a response window of 1400 ms. ***D***, Bar plots show response times (across subjects' mean ± SEM) for each of the six conditions with response requirements in the fMRI experiment. The brackets and stars indicate significance of main effects and interactions. **$p < 0.01$, ***$p < 0.001$. Audition, orange; vision, blue; attended, full pattern; unattended, striped pattern; expected, dark shade; unexpected, light shade.

the requirement to respond to an auditory target in the left versus right hemifield. Critically, spatial attention and expectation were manipulated only in audition but not in vision. Participants needed to respond to all visual targets that were presented in either spatial hemifield with equal probability (i.e., 1:1 in the expected/unexpected hemifields; Fig. 1 *A*, *B*). Throughout the entire experiment a central fixation cross coded in color whether participants should attend and respond to sounds in either their left or right hemifield. The mapping between color and task-relevant hemifield was counterbalanced across participants.

Spatial signal, general response, and spatially selective response probability
Our experiment orthogonally manipulated spatial attention as task-relevance and expectation as spatial signal probability selectively in audition. The attentional manipulation is therefore operationally linked with response requirement over space. Further, attention as response requirement and expectation as signal probability are intimately linked by codetermining general (i.e., the probability that the observer needs to make a response regardless of the hemifield in which the signal is presented) and spatially selective (i.e., the probability that the observer needs to make a response conditioned on that the signal is presented in a particular hemifield) response probabilities.

As shown in Figures 1, *A* and *B*, the general response probability is greater in block type 1, where attention and expectation are directed to the same hemifields, than in block type 2, where attention and expectation are directed to different hemifields. Put differently, greater demands are placed on response inhibition in block type 2 where the hemifield with the more frequent auditory stimuli is task-irrelevant (i.e., a response needs to be inhibited).

Likewise, the spatially selective response probability is codetermined by both attention and expectation. Observers need to respond to both auditory and visual stimuli in the attended hemifield, so that the response probability in the attended hemifield is always equal to one. By contrast, in the unattended hemifield observers need to respond only to the visual stimuli. Hence, in the unattended hemifield the response probability also depends on the frequency of the auditory stimuli and hence on expectation. In the unattended hemifield the response probability is thus smaller and hence response inhibition greater when the task-irrelevant auditory stimuli are more frequent.

Importantly, in our paradigm general and spatially selective response probabilities would predict an interaction between attention and expectation that is common to auditory and visual stimuli. Conversely, main effects of attention and expectation cannot be explained by differences in response probability.

Experimental procedures

The current study included two experiments: (1) a psychophysics experiment conducted across 2 d (i.e., auditory spatial expectation was manipulated between the 2 d), and (2) an fMRI experiment conducted across other 2 d (i.e., auditory spatial expectation was manipulated between the 2 d). The psychophysics experiment was conducted before the fMRI experiment. On each day, the psychophysics and the fMRI experimental runs were preceded by two training runs (see Experimental design).

Each experimental run (duration: ~8 min/run) included 10 attention blocks with 20 trials each, interleaved with 6 s fixation baseline periods. As a result of our balanced factorial design, blocks were of two types: in block type 1, spatial attention and expectation were congruent (i.e., spatial attention was directed to the hemifield with higher auditory target frequency); in block type 2, spatial attention and expectation were incongruent (i.e., attention was directed to the hemifield with lower auditory target frequency; Fig. 1B). Thus, both psychophysics and fMRI experiments included 2000 trials = 20 trials × 10 blocks (attention manipulation: 5 blocks of type 1 and 5 blocks of type 2) × 5 experimental runs × 2 d (expectation manipulation) in total. Therefore, each block type included 400 auditory stimuli for the expected hemifield (pooled over left and right) and 100 auditory stimuli for the unexpected hemifield (pooled over left and right). Each block type also included 250 visual stimuli for the expected hemifield and 250 visual stimuli for the unexpected hemifield (pooled over left and right). For further details, Figure 1B shows the absolute number of trials for each condition and block type and their response requirement for the psychophysics and the fMRI experiment.

The order of "expectation" days (i.e., left vs right) and the color (i.e., pink or green) of the fixation cross (i.e., attention instruction) were counterbalanced across participants, the order of attention blocks was counterbalanced within and across participants and the order of stimulus location and stimulus modality were pseudorandomized within each participant. Brief breaks were included after every run to provide feedback to participants about their performance accuracy (averaged across all conditions) in the target detection task. In the psychophysics experiment participants' fixation performance was monitored via eye tracking, and participants were provided with feedback about their eye movements (i.e., fixation maintenance) during the breaks. For the psychophysics experiment, mean group number of saccades was 22.9 ± 5.2 (across subjects mean ± SEM) and mean group accuracy was 97% ± 0.2% (across subjects mean ± SEM) for the psychophysics experiment and 97% ± 0.5% (across subjects mean ± SEM) for the fMRI experiment.

Each trial (SOA: 2200 ms) included three periods (Fig. 1C): (1) the fixation cross alone (700 ms duration), (2) the brief flash or sound (stimulus duration: 100 ms), and (3) the fixation cross alone, i.e., response window (1400 ms). Participants responded to the auditory targets in the attended hemifield and to all visual targets via key press with their right index finger (i.e., the same response for all auditory and visual targets) as fast and accurately as possible. They fixated the cross in the center of the screen which was presented throughout the entire experiment.

On each day, participants were first familiarized with the stimuli in brief practice runs (with equal spatial signal probability) to train them on target detection performance and, only in the psychophysics experiment, also on fixation (i.e., a warning signal was shown when the disparity between the central fixation cross and the eye-data samples exceeded 2.5°).

After the final fMRI day, participants indicated in a questionnaire whether they thought the sound or the flash was presented more frequently in one of the two spatial hemifields. Eighteen of the total 22 participants correctly reported that the auditory stimuli were more frequent in one hemifield and 20 of 22 participants reported the visual stimuli to be equally frequent across the two hemifields, suggesting that most participants were aware of the manipulation of signal probability.

Experimental setup

Psychophysics experiment. The psychophysics experiment (training and experimental runs) was conducted in a darkened room. Participants rested their chin on a chinrest with the height held constant across all participants. Auditory stimuli were presented at ~72 dB SPL, via HD 280

PRO headphones (Sennheiser). To mimic the scanner environment, the scanner noise was reproduced for the whole duration of the experiment at ~80 dB SPL via external loudspeakers. Visual stimuli were displayed on a gamma-corrected LCD monitor (2560 × 1600 resolution, 60 Hz refresh rate, 30 inch Dell UltraSharp U3014), at a viewing distance of ~50 cm from the participant's eyes. Stimuli were presented using Psychtoolbox v3 (Brainard, 1997; http://www.psychtoolbox.org; RRID: SCR_002881), running under MATLAB R2014a (MathWorks; RRID: SCR_001622) on a Windows machine. Participants responded to all targets with their right index finger and responses were recorded via one key of a small keypad (Targus). Throughout the study, participants' eye movements and fixations were monitored using Tobii Eyex eyetracking system.

fMRI experiment. During the training runs, participants lay in a mock scanner, which mimicked all features of the MRI scanner. The scanner noise was reproduced at ~80 dB SPL via external loudspeakers. During the experimental runs, participants lay in the MRI scanner. Auditory stimuli were presented at ~72 dB SPL using MR-compatible headphones (MR Confon). Visual stimuli were back-projected onto a Plexiglas screen using a BARCO projector (F35). Participants viewed the screen through a mirror mounted on the MR head coil at a viewing distance of ~68 cm. Stimuli were presented using Psychtoolbox v3 (Brainard, 1997; http://www.psychtoolbox.org; RRID:SCR_002881), running under MATLAB R2014a (MathWorks; RRID:SCR_001622) on a MacBook Pro machine. Participants responded to all targets with their right index finger and responses were recorded via an MR-compatible keypad (NATA).

fMRI data acquisition

A 3T Philips MRI scanner with 32 channel head coil was used to acquire both T1-weighted anatomical images (TR = 8.4 ms, TE = 3.8 ms, flip angle = 8°, FOV = 288 mm × 232 mm, image matrix = 288 × 232, 175 sagittal slices acquired in ascending direction, voxel size = 1 × 1 × 1 mm) and T2*-weighted axial echoplanar images with bold oxygenation level-dependent (BOLD) contrast (TR = 2600 ms, TE = 40 ms, flip angle = 85°, FOV = 240 mm × 240 mm, image matrix 80 × 80, 38 transversal slices acquired in ascending direction, voxel size = 3 × 3 × 3 mm). For each participant, an overall of 196 volumes × 5 experimental runs × 2 d = 1960 volumes were acquired. The anatomical image volume was acquired at the end of the experiment.

Statistical analysis

Behavioral data analysis: psychophysics and fMRI experiments. For the behavioral analysis of the psychophysics experiment, we excluded trials where participants did not successfully fixate the central cross based on a dispersion criterion [i.e., distance of fixation from subject's median of fixation (as defined in calibration trials) >1.3° for 3 subsequent samples; Blignaut, 2009]. Percentage (across subjects mean ± SEM) of trials excluded for auditory stimuli: 1.4 ± 0.4%; for visual stimuli: 1.3 ± 0.4%. The response time analysis was limited to correct trials and response times within the range of participant- and condition-specific mean ± 2 SD and <1400 ms (i.e., within the response window).

For auditory targets in the attended hemifield, median response times for each subject were entered into a two-sided paired *t* test with auditory spatial expectation (expected vs unexpected stimulus) as factor.

For visual targets, median response times for each subject were entered into a 2 (auditory spatial attention: attended vs unattended stimulus) × 2 (auditory spatial expectation: expected vs unexpected stimulus) repeated-measures ANOVA.

Unless otherwise indicated, we only report effects that are significant at $p < 0.05$.

fMRI data analysis. The functional MRI data were analyzed with statistical parametric mapping (SPM12, Wellcome Department of Imaging Neuroscience, London; http://www.fil.ion.ucl.ac.uk/spm; Friston et al., 1995). Scans from each subject were realigned using the first as a reference, unwarped, slice-time-corrected, and spatially normalized into MNI standard space using parameters from segmentation of the T1 structural image (Ashburner and Friston, 2005), resampled to a spatial resolution of 2 × 2 × 2 mm³ and spatially smoothed with a Gaussian kernel of 8 mm full-width at half-maximum. The time series of all voxels were high-pass filtered to 1/128 Hz.

The fMRI experiment was modeled in a mixed block/event-related fashion with regressors entered into the design matrix after convolving each event-related unit impulse with a canonical hemodynamic response function and its first temporal derivative. In addition to modeling the 16 conditions in our 2 (stimulus modality: auditory vs visual) × 2 (auditory spatial attention: left vs right hemifield) × 2 (auditory spatial expectation: left vs right hemifield) × 2 (stimulus location: left vs right hemifield) factorial design, the statistical model included the onsets of the attention cue (i.e., auditory attention to the left hemifield, auditory attention to the right hemifield) as a separate regressor. Nuisance covariates included the realignment parameters to account for residual motion artifacts.

Condition-specific effects for each subject were estimated according to the general linear model and passed to a second-level analysis as contrasts. This involved creating 16 contrast images (i.e., each of the 16 conditions relative to fixation, summed over the 10 runs) for each subject and entering them into a second-level ANOVA. Inferences were made at the second level to allow a random-effects analysis and inferences at the population level (Friston et al., 1995).

At the random effects or group level, we pooled over stimulus locations (left/right) and, separately for each sensory modality, we tested for (1) the main effect of spatial attention (i.e., attended > unattended auditory stimuli and vice versa, attended > unattended visual stimuli and vice versa), and (2) the main effect of spatial expectation (i.e., expected > unexpected auditory stimuli and vice versa, expected > unexpected visual stimuli and vice versa).

To assess whether these effects of spatial attention and expectation rely on amodal or modality-specific systems, we investigated (1) whether the effects of attention and expectation are common for audition and vision (i.e., a logical "AND" conjunction over stimulus modalities), or (2) whether the effects differ between audition and vision (i.e., the interaction between attention and stimulus modality and the interaction between expectation and stimulus modality).

Finally, we investigated whether spatial attention and expectation effects were additive or interactive. Separately for each stimulus modality, we tested for (1) the effects that are common for attention and expectation (i.e., a logical "AND" conjunction over each attention and expectation main effects, i.e., additive effects), and (2) the interaction between attention and expectation.

Unless otherwise stated, we report activations at $p < 0.05$ at the cluster level corrected for multiple comparisons within the entire brain using an auxiliary (uncorrected) voxel threshold of $p < 0.001$.

ROI analysis
Based on our *a priori* hypothesis that spatial attention and expectation influence activations in primary sensory cortices, we tested for the effects of auditory spatial attention and expectation selectively within the primary auditory cortex and primary visual cortex. These areas of interests were defined using bilateral ROI maps from SPM Anatomy Toolbox v2.2b (Eickhoff et al., 2005). The anatomical mask for the primary auditory cortex encompassed 890 voxels in the bilateral cytoarchitectonic maps TE 1.0, TE 1.1, and TE 1.2. The anatomical mask for the primary visual cortex encompassed 2936 voxels in the bilateral cytoarchitectonic maps hOC1. We extracted parameter estimates from each ROI, for each of the 16 conditions relative to fixation and for each subject and entered them into a 2 (auditory spatial attention: attended vs unattended stimulus) × 2 (auditory spatial expectation: expected vs unexpected stimulus) repeated-measures ANOVA, separately for each stimulus modality (pooling over stimulus locations).

To sensitize our analysis to attentional modulation of evoked responses, we re-performed this repeated-measures ANOVA separately for (1) the "unilateral ROIs" ipsilateral to the stimulus location and (2) the "unilateral ROIs" contralateral to the stimulus location. Practically, this involved normalization to a symmetric MNI standard template (created by averaging the standard MNI template with its flipped version; Didelot et al., 2010) and (1) pooling over activations in the left ROI (for stimuli in the left hemifield) and the right ROI (for stimuli in the right hemifield; i.e., ipsilateral ROIs), and (2) pooling over activations in the left ROI (for stimuli in the right hemifield) and the right ROI (for stimuli in the left

Table 1. Behavioral results

Experiment	Auditory modality		Visual modality			
	+att +exp	+att −exp	+att +exp	+att −exp	−att +exp	−att −exp
Psychophysics						
RT, ms (SEM)	530.7 (17.1)	566.8 (15.3)	446.9 (10.2)	458.3 (9.4)	487 (11.4)	472.6 (11.8)
fMRI						
RT, ms (SEM)	508.4 (24.5)	552.9 (27.4)	432.3 (13.9)	441.2 (12.9)	467.1 (14.9)	454.2 (14.7)

Group mean reaction times (RTs) for each stimulus modality in each condition for the psychophysics and fMRI experiments. SEM is given in parentheses. att: attention, exp: expectation.

hemifield; i.e., contralateral ROIs), for the corresponding conditions in our 2 (attention) × 2 (expectation) × 2 (stimulus modality) design (for similar analyses, see Lipschutz et al., 2002; Macaluso and Patria, 2007). Because the results of these two "flipped" analyses that separately tested for the effects of attention and expectation on ipsilateral and contralateral stimuli were comparable (with small deviations in *p* values) to our main ROI analysis (because these effects were anyhow bilateral), we do not report these results.

Results

In the following, we report (1) the behavioral results of the psychophysics and the fMRI experiment and (2) the imaging results of the fMRI experiment.

Behavioral results: psychophysics and fMRI experiments
In a target detection task, participants responded to auditory targets presented in their attended hemifield (i.e., auditory attention manipulation) and to all visual targets (Fig. 1A–C).

For both psychophysics and fMRI experiments, the two-sided paired-sample *t* tests on response times for auditory stimuli in the attended hemifield showed significantly faster responses when this hemifield was expected than unexpected [psychophysics: $t_{(30)} = -4.56$, $p < 0.001$, Cohen's d_{av} (95% CI) $= -0.40$ (-0.59, -0.19); fMRI: $t_{(21)} = -5.06$, $p < 0.001$, Cohen's d_{av} (95% CI) $= -0.36$ (-0.54, -0.18); Table 1; Fig. 1D, left].

For both psychophysics and fMRI experiments, the 2 (attended vs unattended) × 2 (expected vs unexpected) repeated-measures ANOVA on response times for visual stimuli revealed a significant main effect of attention [psychophysics: $F_{(1,30)} = 109.88$, $p < 0.001$, η_p^2 (90% CI) $= 0.79$ (0.64, 0.84); fMRI: $F_{(1,21)} = 78.69$, $p < 0.001$, η_p^2 (90% CI) $= 0.79$ (0.61, 0.85)]. Participants responded faster to visual stimuli in their attended than unattended hemifield. Moreover, a significant crossover interaction between attention and expectation was observed [psychophysics: $F_{(1,30)} = 41.59$, $p < 0.001$, η_p^2 (90% CI) $= 0.58$ (0.36, 0.69); fMRI: $F_{(1,21)} = 49.29$, $p < 0.001$, η_p^2 (90% CI) $= 0.70$ (0.47, 0.79)]. The simple main effects showed that participants responded significantly faster to visual targets in the attended hemifield when this hemifield was expected than unexpected (psychophysics: $t_{(30)} = -5.46$, $p < 0.001$, Cohen's d_{av} (95% CI) $= -0.20$ (-0.30, -0.11); fMRI: $t_{(21)} = -3.94$, $p = 0.001$, Cohen's d_{av} (95% CI) $= -0.14$ (-0.22, -0.06); Table 1; Fig. 1D, right]. By contrast, they responded significantly more slowly to visual targets in the unattended hemifield when this hemifield was expected than unexpected (psychophysics: $t_{(30)} = 5.44$, $p < 0.001$, Cohen's d_{av} (95% CI) $= 0.22$ (0.12, 0.32); fMRI: $t_{(21)} = 5.79$, $p < 0.001$, Cohen's d_{av} (95% CI) $= 0.18$ (0.09, 0.26); Table 1; Fig. 1D, right]. Importantly, we observed equivalent response time results in the psychophysics and the fMRI experiment. As we discussed in the Materials and Methods section, this crossover interaction between attention and expectation can be explained by the profile of general and spatially selective response probabilities across conditions. Most prominently, when attention and expectation are
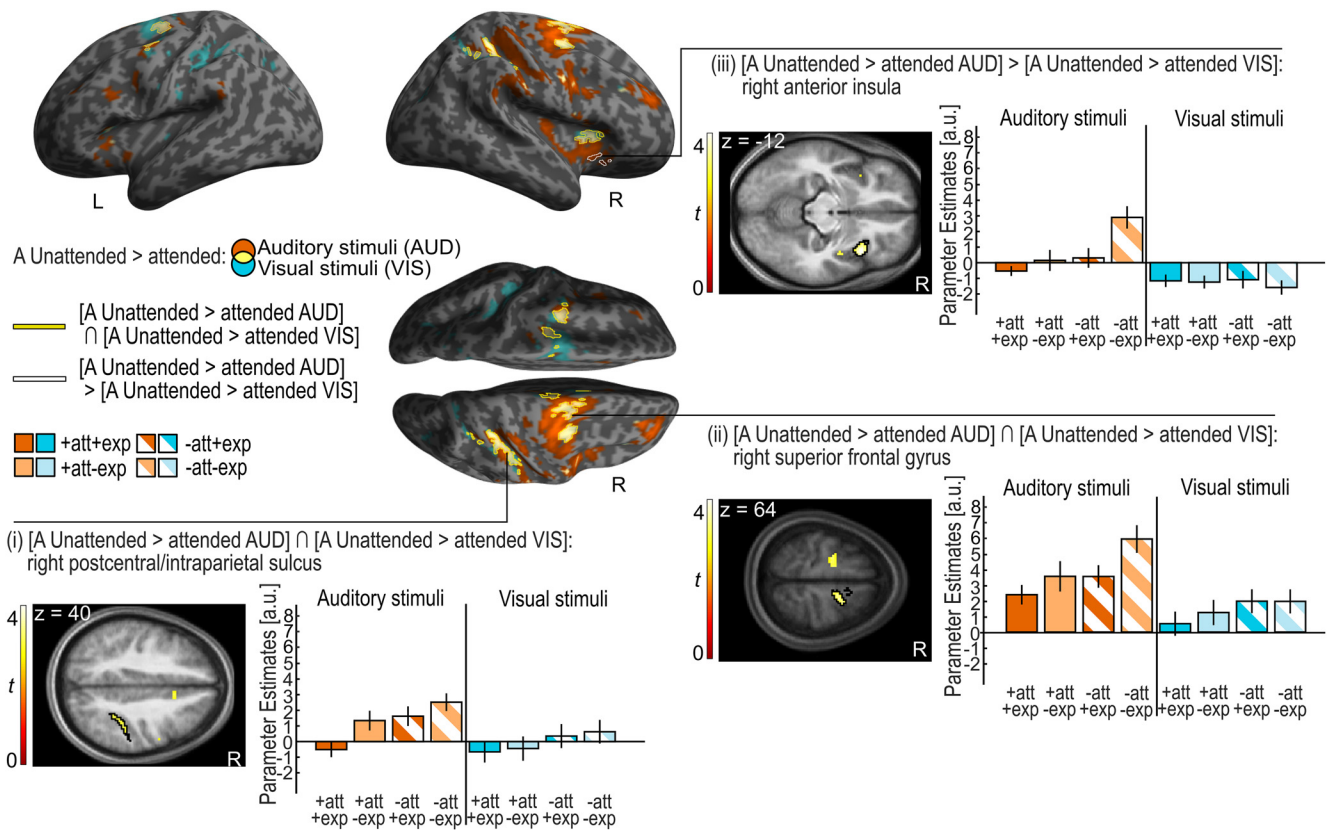
Figure 2. Auditory (A) unattended > attended for auditory and visual stimuli. Activation increases for A unattended > attended stimuli for auditory (AUD; orange, height threshold: $p < 0.001$, uncorr., extent threshold $k > 0$ voxels) and visual (VIS; blue, height threshold: $p < 0.001$, uncorr., extent threshold $k > 0$ voxels) stimuli (overlap, yellow) are rendered on an inflated canonical brain. The conjunction of A unattended > attended for auditory and visual stimuli is encircled in yellow (height threshold: $p < 0.001$, uncorrected, extent threshold $k > 0$ voxels). Activation increases for A unattended > attended that are greater for auditory than visual stimuli (i.e., interaction) are encircled in white (height threshold: $p < 0.001$, uncorrected, extent threshold $k > 0$ voxels). Bar plots show the parameter estimates (across participants mean ± SEM, averaged across all voxels in the black encircled cluster) in the (**i**) right postcentral/intraparietal sulcus, (**ii**) right superior frontal gyrus, and (**iii**) right anterior insula that are displayed on axial slices of a mean image created by averaging the subjects' normalized structural images. The bar graphs represent the size of the effect pertaining to BOLD magnitude in non-dimensional units (corresponding to percentage whole-brain mean). Audition, orange; vision, blue; attended, full pattern; unattended, striped pattern; expected, dark shade; unexpected, light shade.

directed to different hemifields as in block type 2, observers need to inhibit responses on a greater proportion of trials, leading to slower response times.

fMRI results

Effects of auditory spatial attention separately for auditory and visual stimuli

We first evaluated the main effect of spatial attention, separately for each stimulus modality. For auditory stimuli, auditory spatial attention (i.e., A attended vs unattended auditory stimuli) increased activations in bilateral thalami, caudates, hippocampi, left frontoparietal operculum, left putamen, and in a motor network encompassing the left central sulcus and the right cerebellum. The increased activations for auditory stimuli in motor areas can be explained by the motor responses that were given to auditory stimuli only in the attended hemifield. Conversely, because visual stimuli required a motor response in both hemifields, no attentional effects were observed in the motor network for visual stimuli.

Reorienting spatial attention to an auditory stimulus presented in the auditory unattended hemifield (i.e., A unattended vs attended auditory stimuli) induced activations in a predominantly right lateralized frontoparietal system encompassing the bilateral superior frontal gyri (SFG)/sulci and the right postcentral sulcus extending into the intraparietal sulcus (IPS) and the inferior parietal lobule (IPL). Significant activation increases

were also observed in the right anterior cingulate cortex/SFG, right middle frontal gyrus (MFG), and bilateral insulae (Fig. 2, orange and yellow on the inflated brain).

Likewise, shifting attention to a visual stimulus in the auditory unattended hemifield (i.e., A unattended vs attended visual stimuli) increased activations in a more bilateral frontoparietal network including bilateral SFG, superior frontal, precentral and postcentral sulci extending into IPS. We also observed activation increases for unattended visual stimuli in the bilateral anterior cingulate cortices and right anterior insula (Fig. 2, blue and yellow on the inflated brain). Thus, even though spatial attention was manipulated selectively in the auditory modality, we observed similar effects for visual and auditory stimuli when they were presented in the hemifield that was task-irrelevant in audition.

For completeness, no significant main effects of auditory attention for auditory or visual stimuli were found in primary auditory or visual cortices in our selective ROI analysis. However, we observed a main effect of auditory attention on visual stimuli at threshold significance ($p = 0.05$) in auditory cortices (see Table 4).

Effects of auditory spatial attention: commonalities and differences between auditory and visual stimuli

Next, we investigated the extent to which the neural systems engaged by attention shifts are common (i.e., amodal) or distinct (i.e., modality-specific) for auditory (A) and visual stimuli. The conjunc-

Table 2. Amodal and modality dependent mechanisms of auditory spatial attention

Brain regions	MNI coordinates, mm			z-score, peak	Cluster size, voxels	p_{FWE} value, cluster
	x	y	z			
[A unattended > attended auditory stimuli] ∩ [A unattended > attended visual stimuli]						
R SFG	18	−4	64	4.09	731	0.000
R superior frontal sulcus	28	−6	46	4.38		
R anterior cingulate gyrus	10	18	36	3.67		
L SFG	−14	−10	64	3.69	268	0.005
L superior frontal sulcus	−30	−8	48	4.07		
R postcentral sulcus/R IPS	42	−32	40	3.74	304	0.003
R anterior insula	30	20	6	4.41	185	0.027
[A unattended > attended auditory stimuli] > [A unattended > attended visual stimuli]						
R anterior insula	38	16	−12	4.32	209	0.016

p values are FWE-corrected at the cluster level for multiple comparisons within the entire brain. Auxiliary uncorrected voxel threshold of p < 0.001. L, Left; R, right. A, auditory.

Table 3. Main effects of auditory spatial expectation for auditory stimuli

Brain regions	MNI coordinates, mm			z-score, peak	Cluster size, voxels	p_{FWE} value, cluster
	x	y	z			
A Expected > unexpected auditory stimuli						
R SFG	8	54	18	4.91	1458	0.000
L SFG	−6	54	36	5.80		
R precuneus	−4	−52	26	3.90	260	0.006
L precuneus	6	−56	26	3.28		
A Unexpected > expected auditory stimuli						
R STG	60	−44	16	7.47	18,305	0.000
L STG	−62	−34	14	5.44		
R postcentral sulcus/R IPS	34	−58	46	5.93		
L postcentral sulcus/L IPS	−38	−46	42	5.79		
R precuneus	4	−54	54	6.49		
L precuneus	−8	−54	54	6.48		
R anterior insula	38	16	2	7.48		
L anterior insula	−32	16	2	6.94		
R posterior cingulate gyrus/ L posterior cingulate gyrus	4	−28	26	5.18	339	0.001
R anterior cingulate gyrus	8	22	32	5.40	4222	0.000
R SFG	18	2	66	4.14		
L SFG	−26	−8	70	4.23		
L precentral sulcus	−38	0	38	5.08		
R precentral sulcus	40	6	30	4.97	2186	0.000
R middle frontal gyrus	40	34	36	4.42		
L middle frontal gyrus	−34	46	24	4.56	810	0.000
R calcarine cortex	12	−84	8	3.75	680	0.000
L calcarine cortex	−12	−84	6	3.71		

p values are FWE-corrected at the cluster level for multiple comparisons within the entire brain. Auxiliary uncorrected voxel threshold of p < 0.001. L, Left; R, right. A, auditory.

tion analysis over sensory modalities showed increased activations for attention shifts (i.e., [A unattended > attended auditory stimuli] ∩ [A unattended > attended visual stimuli]) in the bilateral SFG and sulci, right anterior cingulate gyrus, right postcentral sulcus extending into IPS, and right anterior insula (Table 2; Fig. 2).

Only the right insula, which was also part of the attentional system that was commonly engaged by unattended auditory and visual stimuli, showed a stronger attentional effect for auditory than visual stimuli (i.e., interaction: [A unattended > attended auditory stimuli] > [A unattended > attended visual stimuli]; Table 2; Fig. 2).

Table 2 and Figure 2 show the significant clusters of the conjunction analysis: (1) right postcentral sulcus/IPS (encircled in yellow on the inflated SPM template and encircled in black in the axial slice), and (2) right SFG (encircled in yellow on the inflated SPM template and encircled in black in the axial slice). (3) The interaction between attention and stimulus modality in the right insula is shown in Table 2 and Figure 2 (encircled in white on the inflated SPM template and encircled in black in the axial slice).

In summary, in line with our behavioral results, our fMRI analysis suggests that the effect of auditory spatial attention generalizes to visual stimuli. Spatial reorienting to both auditory and visual stimuli in the hemifield that was task-irrelevant selectively in audition increased activations in a widespread right lateralized frontoparietal system (Shomstein and Yantis, 2006; Indovina and Macaluso, 2007; Santangelo et al., 2009; Shulman et al., 2009; Doricchi et al., 2010). Although the right insula exhibited significantly stronger attentional effects for auditory than visual stimuli, we did not observe attentional effects that were truly selective for stimuli from either the visual or auditory modality. Collectively, these results suggest that spatial attention and reorienting rely predominantly on neural systems that are interactively shared across sensory modalities, even though they may be more strongly engaged by stimuli of the sensory modality where spatial attention is directly manipulated.

Effects of auditory spatial expectation separately for auditory and visual stimuli
Auditory stimuli in the expected relative to unexpected hemifield elicited significantly greater activation in the bilateral medial prefrontal cortices (i.e., anterior portions of the SFG) and the bilateral precunei/posterior cingulate gyri (Table 3; Fig. 3; Summerfield et al., 2006).

By contrast, auditory stimuli in the unexpected relative to expected hemifield increased activations in a widespread frontoparietal system encompassing bilateral SFG/MFG and sulci and the postcentral/parietal sulci extending into IPL. We also ob-

served activation increases for unexpected auditory stimuli in the bilateral precunei, anterior insulae, anterior and posterior cingulate gyri, and in the bilateral plana temporalia and superior temporal gyri (STG) previously implicated in spatial processing (Griffiths and Warren, 2002; Brunetti et al., 2005; Ahveninen et al., 2006; Table 3; Fig. 4). Critically, the effects of expectation in regions of the auditory cortices were not observed for unattended relative to attended stimuli (p < 0.05 uncorrected) suggesting that they were selective for auditory expectation. Surprisingly, unexpected relative to expected auditory stimuli increased activations also in the bilateral calcarine cortices (Table 3; Fig. 4).

Our selective ROI analysis also revealed higher activations for unexpected relative to expected auditory stimuli in primary auditory and visual cortices (Table 4, asterisks).

Surprisingly, neither whole-brain nor ROI analyses revealed any significant effects of spatial expectation for visual stimuli.

Effects of auditory spatial expectation: commonalities and differences between auditory and visual stimuli
Our results suggest that a widespread neural system forms spatial expectations selectively for stimuli from the auditory modality where signal probability was manipulated. Indeed, this was confirmed by the significant interaction between expectation and stimulus modality ([A unexpected > expected auditory stimuli] > [A unexpected > expected visual stimuli]) that was observed in large parts of the neural system showing auditory expectation effects for auditory stimuli (Fig. 4, areas on the inflated brain with white outline). By contrast, the conjunction analyses over stimulus modality did not reveal any significant effects of auditory expectation that were common to auditory and visual stimuli.

Additive and interactive effects of spatial attention and expectation: separately for auditory and visual stimuli
Finally, we investigated separately for auditory and visual stimuli whether attention and expectation effects engage common neural
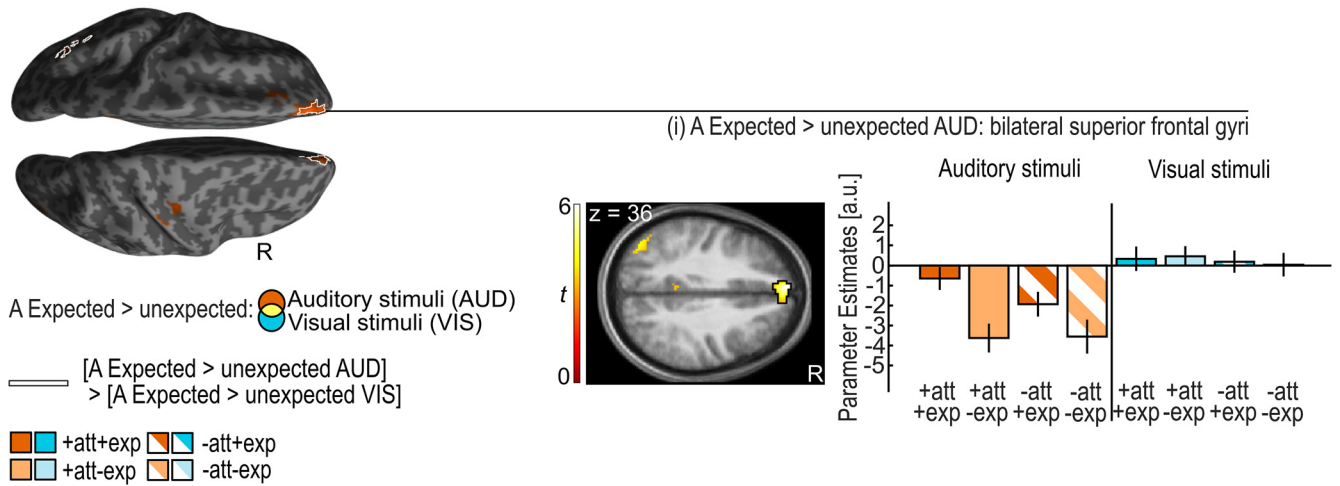
Figure 3. Auditory (A) expected > unexpected for auditory and visual stimuli. Activation increases for A expected > unexpected auditory stimuli (orange) are rendered on an inflated canonical brain; they are encircled in white if they are significantly greater for auditory than visual stimuli (i.e., interaction). Height threshold of $p < 0.001$, uncorrected; extent threshold $k > 0$ voxels. Bar plots show the parameter estimates (across participants mean $\pm$ SEM, averaged across all voxels in the black encircled cluster) in the medial prefrontal cortices (i.e., anterior portions of the superior frontal gyri) that are displayed on axial slices of a mean image created by averaging the subjects' normalized structural images; the bar graphs represent the size of the effect in non-dimensional units (corresponding to percentage whole-brain mean). Audition: orange; vision: blue; attended: full pattern; unattended: striped pattern; expected: dark shade; unexpected: light shade.
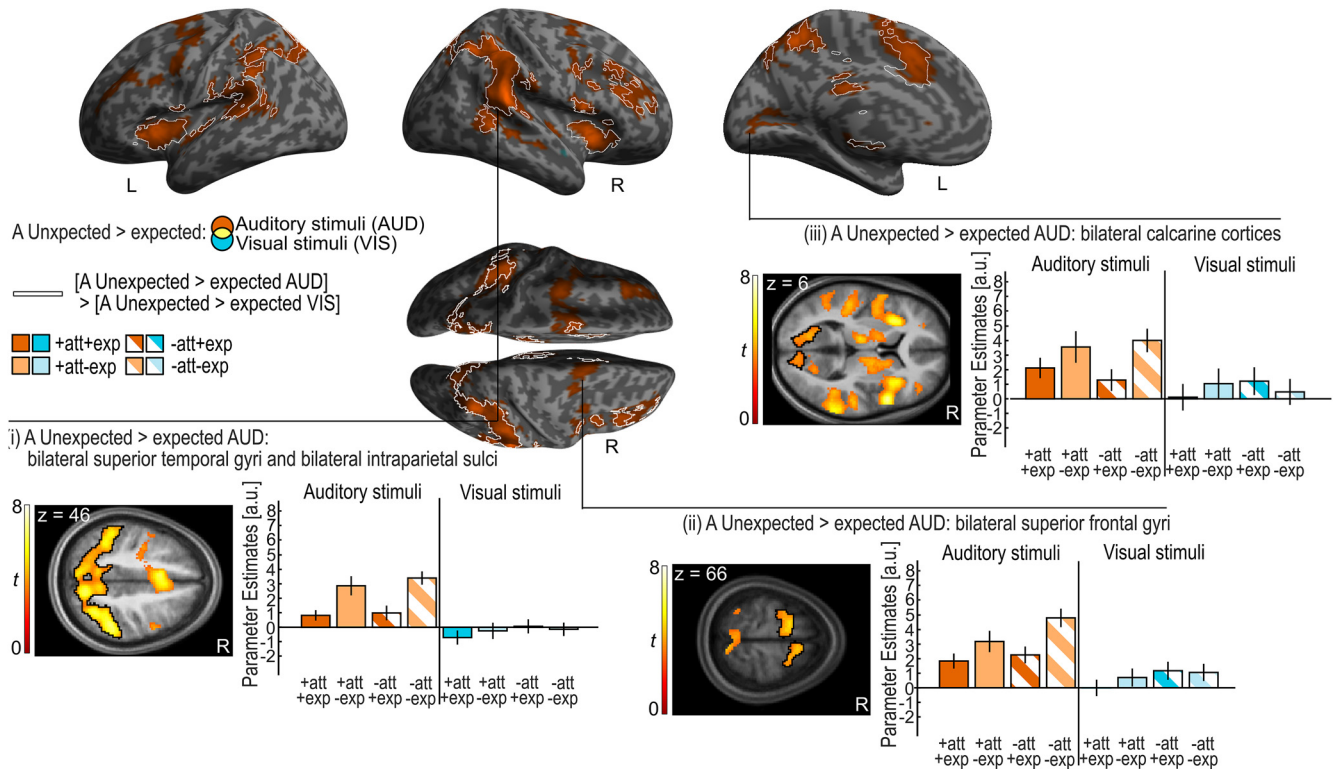

Figure 4. Auditory (A) unexpected > expected for auditory and visual stimuli. Activation increases for A unexpected > expected stimuli for auditory stimuli (orange) are rendered on an inflated canonical brain; they are encircled in white if they are significantly greater for auditory than visual stimuli (i.e., interaction). Height threshold of $p < 0.001$, uncorrected; extent threshold $k > 0$ voxels. Bar plots show the parameter estimates (across participants mean $\pm$ SEM, averaged across all voxels in the black encircled cluster) in (**i**) bilateral superior temporal gyri and bilateral intraparietal sulci, (**ii**) bilateral superior frontal gyri, and (**iii**) bilateral calcarine cortices that are displayed on axial slices of a mean image created by averaging the subjects' normalized structural images. The bar graphs represent the size of the effect in non-dimensional units (corresponding to percentage whole-brain mean). Audition, orange; vision, blue; attended, full pattern; unattended, striped pattern; expected, dark shade; unexpected, light shade.

systems in an additive (i.e., conjunction over attention and expectation) or interactive fashion (i.e., interaction between attention and expectation).

For auditory stimuli, neither the whole-brain nor the selective ROI analyses (Table 4) revealed any significant interaction between attention and expectation. By contrast, the conjunc-

tion analysis over attention and expectation revealed activation increases jointly for unattended > attended and unexpected > expected (i.e., in an additive fashion) in a predominantly right-lateralized frontoparietal system including the bilateral superior/MFG/sulci and the right postcentral/IPS extending into right IPL. Further, we observed additive effects

Table 4. Results of the ROI analysis for each stimulus modality

ROIs	Main effect of A attention	Main effect of A expectation	Interaction A attention × expectation
Primary auditory cortex			
Auditory stimuli			
$F_{(1,21)}$	0.148	12.671	0.846
p	0.704	0.002**	0.368
η_p^2 (90% CI)	0.007 (0, 0.140)	0.376 (0.106, 0.558)	0.039 (0, 0.223)
Visual stimuli			
$F_{(1,21)}$	4.310	0.213	0.117
p	0.050	0.649	0.736
η_p^2 (90% CI)	0.170 (0, 0.383)	0.010 (0, 0.154)	0.006 (0, 0.131)
Primary visual cortex			
Auditory stimuli			
$F_{(1,21)}$	0.165	7.213	2.096
p	0.689	0.014*	0.162
η_p^2 (90% CI)	0.008 (0, 0.144)	0.256 (0.032, 0.461)	0.091 (0, 0.296)
Visual stimuli			
$F_{(1,21)}$	0.995	0.054	5.062
p	0.330	0.819	0.035*
η_p^2 (90% CI)	0.045 (0, 0.233)	0.003 (0, 0.091)	0.194 (0.008, 0.406)

ROIs: primary auditory and primary visual cortex. 90% CI of η_p^2 is given in parentheses. **$p < 0.01$, *$p < 0.05$. A, auditory.

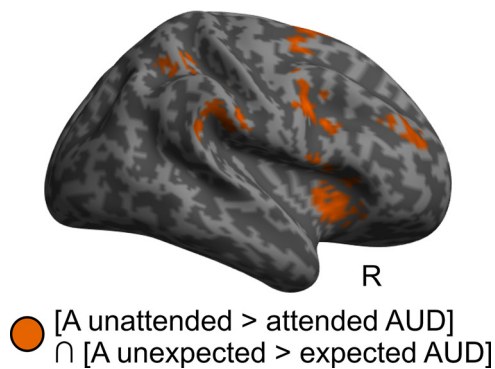○ [A unattended > attended AUD]
∩ [A unexpected > expected AUD]

Figure 5. Additive effects of auditory (A) attention and expectation in audition (AUD). Activation increases common (i.e., conjunction) for A attention and expectation main effects in the auditory modality ([A unattended > attended AUD] ∩ [A unexpected > expected AUD]) are rendered in orange on an inflated canonical brain; height threshold of $p < 0.001$, uncorrected; extent threshold $k > 0$ voxels.

in the right anterior cingulate gyrus and in the bilateral insulae (Fig. 5).

For visual stimuli, the whole-brain analysis did not reveal any significant additive or interactive effects for attention and expectation. The ROI analysis revealed a significant interaction between attention and expectation in the primary visual cortex, with greater activations for unexpected than expected visual stimuli in the attended hemifield, but lower activation for unexpected than expected visual stimuli in the unattended hemifield (Table 4, asterisk). As we discussed in Materials and Methods, this interaction between attention and expectation may potentially be caused by differences in response probabilities, response times, and associated processes of response inhibition that are increased for block type 2 relative to block type 1.

Effect of "awareness of auditory expectation manipulation"
Only 4 of the 22 participants were not aware of the spatial expectation manipulation in audition. For completeness, we therefore investigated whether the expectation effects for auditory stimuli depended on observers' explicit knowledge about auditory signal probability. Hence, at the second between participants level we

compared the auditory expectation effects between these 4 "unaware" and 18 "aware" participants (i.e., interaction between unexpected > expected for auditory stimuli for aware vs unaware). However, this interaction did not reveal significant clusters (whole-brain corrected). By contrast, a conjunction-null conjunction analysis over both groups replicated the effects for unexpected relative to expected auditory stimuli in planum temporale, anterior insula, and parietal cortex. These results suggest that explicit knowledge may not be required for the brain to express activation increases signaling a prediction error for unexpected auditory stimuli.

Discussion
The current study was designed to investigate whether the brain allocates attentional resources and forms expectations over space separately or interactively across the senses. To dissociate the effects of spatial attention and expectation we orthogonally manipulated spatial attention as response requirement and expectation as stimulus probability over space selectively in audition and assessed their effects on neural and behavioral responses in audition and vision.

Consistent with previous research, our behavioral results show that participants responded significantly faster to visual stimuli that were presented in the hemifield where auditory stimuli were task-relevant (Spence and Driver, 1996, 1997). In other words, directing observers' spatial attention to one hemifield selectively in audition impacted participants' response speed to auditory and visual stimuli, suggesting that attentional resources are at least partly shared across sensory modalities.

Likewise, the neural responses to both auditory and visual stimuli depended on auditory spatial attention. Regardless of their sensory modality, unattended relative to attended stimuli increased activations in a widespread right-lateralized dorsal and ventral frontoparietal system that has previously been implicated in sustained spatial attention (Leitão et al., 2015, 2017) and spatial (re)orienting and contextual updating based on attentional cuing paradigms that conflated attention and expectation (Nobre et at., 2000; Macaluso et al., 2002; Kincade et al., 2005; Bressler et al., 2008; Santangelo et al., 2009). By orthogonally manipulating task-relevance (i.e., response requirement) and expectation (i.e., signal probability), the current study allowed us to attribute these frontoparietal activations to attentional mechanisms. Our results corroborate that the brain has only limited abilities to define spatial locations as task-relevant or irrelevant independently for audition and vision (Eimer and Schröger, 1998; Eimer, 1999; Macaluso, 2010). As a result, visual stimuli engaged spatial reorienting even though they should have been attended equally in both hemifields. Conversely, auditory stimuli induced attentional reorienting, even though they should have been ignored as task-irrelevant. The, at least to some extent, amodal definition of spatial task-relevance may also explain the extensive activations that we observed for unattended stimuli not only in the ventral, but also the dorsal attentional network that is typically more associated with sustained attention. Greater sustained attention may be required for stimuli in the "auditory unattended" hemifield, because the brain needs to decide whether to respond (i.e., visual stimuli) or not to respond (i.e., auditory stimuli; Indovina and Macaluso, 2007; Santangelo et al., 2009). In summary, our behavioral and neuroimaging findings suggest that spatial attention, when defined as task-relevance, operates interactively across the senses.

Next, we asked whether the neural systems encode spatial signal probability independently across audition and vision. Behaviorally, we observed faster responses to expected than unexpected

stimuli regardless of sensory modality in the task-relevant hemifield. Yet, surprisingly we observed faster responses for unexpected than expected visual stimuli in the task-irrelevant hemifield (note that auditory stimuli did not require a response in the task-irrelevant hemifield). Hence, we observed a significant interaction between attention and expectation for visual response times. As discussed in Materials and Methods and in greater detail by Zuanazzi and Noppeney (2018), this interaction for visual response times most likely results from the differences in the general response probability across conditions. The response probability is greater when attention and expectation are congruent and directed to the same hemifield (90% of the trials in blocks of type 1) than when they are directed to different hemifields (60% of the trials in blocks of type 2; Fig. 1A,B). Put differently, observers need to inhibit their responses on a greater proportion of trials in block type 2, when the frequency of task-irrelevant auditory stimuli in the auditory unattended hemifield is high.

Critically, fMRI allows us to move beyond response times and track the neural processes across the entire visual and auditory processing hierarchy regardless of whether (e.g., auditory) stimuli are responded to. This provides us with the opportunity to investigate whether the brain forms expectations or spatial event probability maps separately for different sensory modalities. Based on the notion of predictive coding, we would then expect activation increases signaling a prediction error for stimuli that are presented at unexpected spatial locations (Rao and Ballard, 1999; Friston, 2005). Indeed, spatially unexpected relative to expected auditory stimuli increased activations as a prediction error signal in the plana temporalia that are critical for auditory spatial encoding as well as higher order frontoparietal areas. These results suggest that the planum temporale forms spatial prediction error signals for spatial unexpected auditory stimuli that then propagate up the hierarchy into frontoparietal areas (Friston, 2005). Alternatively, prediction error signals in the planum temporale may trigger the frontoparietal attentional system leading to spatial reorienting (den Ouden et al., 2012). Our design and the sluggishness of the BOLD response make it difficult to dissociate between these two explanations for the frontoparietal activations. Future EEG/MEG studies may be able to disentangle whether the expression of prediction error signals in the planum temporale may subsequently trigger attentional reorienting in the frontoparietal system.

Critically, however, we observed activation increases only for auditory stimuli when presented in the auditory unexpected hemifield, but not for visual stimuli. In fact, even the visual cortex showed activation increases only for unexpected auditory stimuli potentially mediated via direct connectivity from auditory areas or top-down modulation from parietal cortices.

Likewise, activation increases for spatially expected stimuli were observed selectively for audition in the medial prefrontal cortex that has previously been implicated in forming representations consistent with one's expectations (Summerfield et al., 2006). Hence, in line with the notion of predictive coding, higher order areas such as the medial prefrontal cortex form representations when stimuli match our spatial expectations, while sensory and potentially frontoparietal areas signal a prediction error when our spatial expectations are violated (Rao and Ballard, 1999; Friston, 2005). Critically, spatial expectations and prediction error signals were formed in a modality-specific fashion selectively for audition, where stimulus probability was explicitly manipulated. In fact, we did not observe any significant positive or negative expectation effects for visual stimuli anywhere in the brain even at an uncorrected threshold of $p < 0.2$ at the cluster

level. These results suggest that the neural systems can form and compute spatial expectations that encode the probability of stimulus occurrence separately for different sensory modalities.

Finally, we asked separately for audition and vision whether spatial attention and expectations influence these neural responses in an additive or interactive fashion. Recent accounts of predictive coding suggest that attention may increase the precision of prediction errors potentially leading to an increase in prediction error signals (i.e., BOLD-response enhancement for unexpected relative expected stimuli) in the attended hemifield (Feldman and Friston, 2010; Auksztulewicz and Friston, 2015). However, contrary to this prediction, spatial attention and expectation did not interact in the auditory modality but influenced neural responses in this system in an additive fashion. Our conjunction analysis over spatial attention and expectation revealed a dorsal and ventral frontoparietal network that was jointly recruited by spatial reorienting as well as expectation violations in audition (which may in turn trigger spatial reorienting). By contrast, in primary visual areas we observed a significant interaction between spatial attention and expectation selectively for visual stimuli (ROI analysis; Table 4). Activations for visual stimuli were greater when attention and expectation were directed to different hemifields than to the same hemifield. This activation profile mimics the pattern that we observed for behavioral response times and can be found at a lower threshold of significance throughout the motor system (e.g., primary motor cortex and cerebellum). It may thus be most likely mediated by top-down influences from response selection processes onto sensory cortices (van Elk et al., 2010; Gutteling et al., 2013, 2015). The interaction between attention and expectation in our study highlights processes of expectation (or stimulus history/probability) that depend on whether these stimuli were task-relevant (i.e., targets) or irrelevant (i.e., non-targets). It thereby contributes to the growing literature that reveals the importance of selection history (i.e., the probability of targets vs non-targets or distractors) on spatial (and other) priority maps (Awh et al., 2012; Lamy and Kristjánsson, 2013; Chelazzi and Santandrea, 2018; Theeuwes, 2018).

In conclusion, our results demonstrate that spatial attention and expectation engage partly overlapping neural systems yet differ in their modality-specificity. Attentional resources were controlled interactively across audition and vision within a widespread right-lateralized frontoparietal system. By contrast, spatial expectations and prediction error signals were formed in the planum temporale and frontoparietal cortices selectively for auditory stimuli where stimulus probability was explicitly manipulated.

Future studies need to investigate the extent to which the modality-specificity of spatial expectations depends on the statistical structure of the multisensory environment. For instance, in our experiment auditory and visual signals never occurred together thereby promoting an encoding of signal probability separately for each sensory modality. We therefore need to assess the impact of correlations between auditory and visual signals on the encoding of signal probability. Moreover, given the highly factorial nature of our design, we manipulated signal probability only in audition and assessed the generalization to vision. The reverse experiment (i.e., manipulating signal probability in vision) could reveal potential differences in the encoding and generalization of signal probability between audition and vision (see related discussion about asymmetric links of attentional resources by Spence and Driver, 1997; Ward et al., 2000). Because auditory events are typically transient and visual objects permanent, the brain may have developed different strategies for encoding signal

probabilities across the senses. Finally, future studies may manipulate stimulus probability via probabilistic cues rather than stimulus frequency to further characterize the neural mechanisms mediating prediction and prediction error signals (e.g., relationship between expectations/predictions and repetition suppression/priming; Wiggs and Martin, 1998).

References

Ahveninen J, Jääskeläinen IP, Raij T, Bonmassar G, Devore S, Hämäläinen M, Levänen S, Lin FH, Sams M, Shinn-Cunningham BG, Witzel T, Belliveau JW (2006) Task-modulated "what" and "where" pathways in human auditory cortex. Proc Natl Acad Sci U S A 103:14608–14613.

Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L (2010) Stimulus predictability reduces responses in primary visual cortex. J Neurosci 30:2960–2966.

Ashburner J, Friston KJ (2005) Unified segmentation. Neuroimage 26:839-851.

Auksztulewicz R, Friston K (2015) Attentional enhancement of auditory mismatch responses: a DCM/MEG study. Cereb Cortex 25:4273–4283.

Awh E, Belopolsky AV, Theeuwes J (2012) Top-down versus bottom-up attentional control: a failed theoretical dichotomy. Trends Cogn Sci 16:437–443.

Blignaut P (2009) Fixation identification: the optimum threshold for a dispersion algorithm. Atten Percept Psychophys 71:881–895.

Brainard DH (1997) The psychophysics toolbox. Spat Vis 10:433–436.

Brefczynski JA, DeYoe EA (1999) A physiological correlate of the "spotlight" of visual attention. Nat Neurosci 2:370–374.

Bressler DW, Fortenbaugh FC, Robertson LC, Silver MA (2013) Visual spatial attention enhances the amplitude of positive and negative fMRI responses to visual stimulation in an eccentricity-dependent manner. Vision Res 85:104–112.

Bressler SL, Tang W, Sylvester CM, Shulman GL, Corbetta M (2008) Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. J Neurosci 28:10056–10061.

Brunetti M, Belardinelli P, Caulo M, Del Gratta C, Della Penna S, Ferretti A, Lucci G, Moretti A, Pizzella V, Tartaro A, Torquati K, Olivetti Belardinelli M, Romani GL (2005) Human brain activation during passive listening to sounds from different locations: an fMRI and MEG study. Hum Brain Mapp 26:251–261.

Carrasco M (2011) Visual attention: the past 25 years. Vision Res 51:1484–1525.

Chelazzi L, Santandrea E (2018) The time constant of attentional control: short, medium and long (Infinite?). J Cogn 1:27.

Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3:201–215.

Corbetta M, Patel G, Shulman GL (2008) The reorienting system of the human brain: from environment to theory of mind. Neuron 58:306–324.

den Ouden HEM, Kok P, de Lange FP (2012) How prediction errors shape perception, attention, and motivation. Front Psychol 3:548.

Didelot A, Mauguière F, Redouté J, Bouvard S, Lothe A, Reilhac A, Hammers A, Costes N, Ryvlin P (2010) Voxel-based analysis of asymmetry index maps increases the specificity of ¹⁸F-MPPF PET abnormalities for localizing the epileptogenic zone in temporal lobe epilepsies. J Nucl Med 51:1732–1739.

Doherty JR, Rao A, Mesulam MM, Nobre AC (2005) Synergistic effect of combined temporal and spatial expectations on visual attention. J Neurosci 25:8259–8266.

Doricchi F, Macci E, Silvetti M, Macaluso E (2010) Neural correlates of the spatial and expectancy components of endogenous and stimulus-driven orienting of attention in the Posner task. Cereb Cortex 20:1574–1585.

Downing CJ (1988) Expectancy and visual-spatial attention: effects on perceptual quality. J Exp Psychol Hum Percept Perform 14:188–202.

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25:1325–1335.

Eimer M (1999) Can attention be directed to opposite locations in different modalities? An ERP study. Clin Neurophysiol 110:1252–1259.

Eimer M, Schröger E (1998) ERP effects of intermodal attention and cross-modal links in spatial attention. Psychophysiology 35:313–327.

Feldman H, Friston KJ (2010) Attention, uncertainty, and free-energy. Front Hum Neurosci 4:215.

Friston K (2005) A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci 360:815–836.

Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ (1995) Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp 2:189–210.

Gardner WG, Martin KD (1995) HRTF measurements of a KEMAR. J Acoust Soc Am 97:3907–3908.

Geng JJ, Behrmann M (2002) Probability cuing of target location facilitates visual search implicitly in normal participants and patients with hemispatial neglect. Psychol Sci 13:520–525.

Geng JJ, Behrmann M (2005) Spatial probability as an attentional cue in visual search. Percept Psychophys 67:1252–1268.

Griffiths TD, Warren JD (2002) The planum temporale as a computational hub. Trends Neurosci 25:348–353.

Gutteling TP, Park SY, Kenemans JL, Neggers SF (2013) TMS of the anterior intraparietal area selectively modulates orientation change detection during action preparation. J Neurophysiol 110:33–41.

Gutteling TP, Petridou N, Dumoulin SO, Harvey BM, Aarnoutse EJ, Kenemans JL, Neggers SF (2015) Action preparation shapes processing in early visual cortex. J Neurosci 35:6472–6480.

Indovina I, Macaluso E (2007) Dissociation of stimulus relevance and saliency factors during shifts of visuospatial attention. Cereb Cortex 17:1701–1711.

Kincade JM, Abrams RA, Astafiev SV, Shulman GL, Corbetta M (2005) An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. J Neurosci 25:4593–4604.

Kok P, Jehee JF, de Lange FP (2012a) Less is more: expectation sharpens representations in the primary visual cortex. Neuron 75:265–270.

Kok P, Rahnev D, Jehee JF, Lau HC, de Lange FP (2012b) Attention reverses the effect of prediction in silencing sensory signals. Cereb Cortex 22:2197–2206.

Krumbholz K, Nobis EA, Weatheritt RJ, Fink GR (2009) Executive control of spatial attention shifts in the auditory compared to the visual modality. Hum Brain Mapp 30:1457–1469.

Lamy DF, Kristjánsson A (2013) Is goal-directed attentional guidance just intertrial priming? A review. J Vis 13(3):14 1–19.

Leitão J, Thielscher A, Tünnerhoff J, Noppeney U (2015) Concurrent TMS-fMRI reveals interactions between dorsal and ventral attentional systems. J Neurosci 35:11445–11457.

Leitão J, Thielscher A, Lee H, Tuennerhoff J, Noppeney U (2017) Transcranial magnetic stimulation of right inferior parietal cortex causally influences prefrontal activation for visual detection. Eur J Neurosci 46:2807–2816.

Lipschutz B, Kolinsky R, Damhaut P, Wikler D, Goldman S (2002) Attention-dependent changes of activation and connectivity in dichotic listening. Neuroimage 17:643–656.

Macaluso E (2010) Orienting of spatial attention and the interplay between the senses. Cortex 46:282–297.

Macaluso E, Patria F (2007) Spatial re-orienting of visual attention along the horizontal or the vertical axis. Exp Brain Res 180:23–34.

Macaluso E, Frith CD, Driver J (2002) Supramodal effects of covert spatial orienting triggered by visual or tactile events. J Cogn Neurosci 14:389–401.

McDonald JJ, Teder-Sälejärvi WA, Hillyard SA (2000) Involuntary orienting to sound improves visual perception. Nature 407:906–908.

Nobre AC, Gitelman DR, Dias EC, Mesulam MM (2000) Covert visual spatial orienting and saccades: overlapping neural systems. Neuroimage 11:210–216.

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113.

Posner MI (1980) Orienting of attention. Q J Exp Psychol 32:3–25.

Posner MI, Snyder CR, Davidson BJ (1980) Attention and the detection of signals. J Exp Psychol 109:160–174.

Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci 2:79–87.

Rohe T, Noppeney U (2015) Cortical hierarchies perform Bayesian causal inference in multisensory perception. PLoS Biol 13:e1002073.

Rohe T, Noppeney U (2016) Distinct computational principles govern multisensory integration in primary sensory and association cortices. Curr Biol 26:509–514.

Rohe T, Noppeney U (2018) Reliability-weighted integration of audiovisual

signals can be modulated by top-down control. eNeuro 5:ENEURO. 0315-17.2018.

Santangelo V, Macaluso E (2012) Spatial attention and audiovisual processing. In: The new handbook of multisensory processing (Stein BE, ed), pp 359–370. Cambridge, MA: MIT.

Santangelo V, Olivetti Belardinelli M, Spence C, Macaluso E (2009) Interactions between voluntary and stimulus-driven spatial attention mechanisms across sensory modalities. J Cogn Neurosci 21:2384–2397.

Shomstein S, Yantis S (2006) Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. J Neurosci 26:435–439.

Shulman GL, Astafiev SV, Franke D, Pope DL, Snyder AZ, Abraham Z, Mcavoy MP, Corbetta M (2009) Interaction of stimulus-driven reorienting and expectation in ventral and dorsal frontoparietal and basal ganglia-cortical networks. J Neurosci 29:4392–4407.

Spence C, Driver J (1996) Audiovisual links in endogenous covert spatial attention. J Exp Psychol Hum Percept Perform 22:1005–1030.

Spence C, Driver J (1997) Audiovisual links in exogenous covert spatial orienting. Percept Psychophys 59:1–22.

Spence C, Pavani F, Driver J (2000) Crossmodal links between vision and touch in covert endogenous spatial attention. J Exp Psychol Hum Percept Perform 26:1298–1319.

Stekelenburg JJ, Vroomen J (2012) Electrophysiological correlates of predictive coding of auditory location in the perception of natural audiovisual events. Front Integr Neurosci 6:26.

Summerfield C, Egner T (2009) Expectation (and attention) in visual cognition. Trends Cogn Sci 13:403–409.

Summerfield C, Egner T, Greene M, Koechlin E, Mangels J, Hirsch J (2006) Predictive codes for forthcoming perception in the frontal cortex. Science 314:1311–1314.

Summerfield C, Monti JM, Trittschuh EH, Mesulam MM, Egner T (2008) Neural repetition suppression reflects fulfilled perceptual expectations. Nat Neurosci 11:1004–1006.

Theeuwes J (2018) Visual selection: usually fast and automatic; seldom slow and volitional. J Cogn 1:29.

Thirion B, Pinel P, Mériaux S, Roche A, Dehaene S, Poline JB (2007) Analysis of a large fMRI cohort: statistical and methodological issues for group analyses. Neuroimage 35:105–120.

Tootell RB, Hadjikhani N, Hall EK, Marrett S, Vanduffel W, Vaughan JT, Dale AM (1998) The retinotopy of visual spatial attention. Neuron 21:1409–1422.

van Elk M, van Schie HT, Neggers SF, Bekkering H (2010) Neural and temporal dynamics underlying visual selection for action. J Neurophysiol 104:972–983.

Ward LM, McDonald JJ, Lin D (2000) On asymmetries in cross-modal spatial attention orienting. Percept Psychophys 62:1258–1264.

Wiggs CL, Martin A (1998) Properties and mechanisms of perceptual priming. Curr Opin Neurobiol 8:227–233.

Wu CT, Weissman DH, Roberts KC, Woldorff MG (2007) The neural circuitry underlying the executive control of auditory spatial attention. Brain Res 1134:187–198.

Zuanazzi A, Noppeney U (2018) Additive and interactive effects of spatial attention and expectation on perceptual decisions. Sci Rep 8:6732.