Audiovisual adaptation is expressed in spatial and decisional codes

Máté Aller¹,²,⁵✉, Agoston Mihalik¹,³,⁵ & Uta Noppeney¹,⁴

The brain adapts dynamically to the changing sensory statistics of its environment. Recent research has started to delineate the neural circuitries and representations that support this cross-sensory plasticity. Combining psychophysics and model-based representational fMRI and EEG we characterized how the adult human brain adapts to misaligned audiovisual signals. We show that audiovisual adaptation is associated with changes in regional BOLD-responses and fine-scale activity patterns in a widespread network from Heschl’s gyrus to dorsolateral prefrontal cortices. Audiovisual recalibration relies on distinct spatial and decisional codes that are expressed with opposite gradients and time courses across the auditory processing hierarchy. Early activity patterns in auditory cortices encode sounds in a continuous space that flexibly adapts to misaligned visual inputs. Later activity patterns in frontoparietal cortices code decisional uncertainty consistent with these spatial transformations. Our findings suggest that regions within the auditory processing hierarchy multiplex spatial and decisional codes to adapt flexibly to the changing sensory statistics in the environment.
Throughout life the brain needs to adapt dynamically to changes in the environment and the sensorium. Changes in the sensory statistics evolve across multiple timescales ranging from milliseconds to years and bring cues from different sensory modalities into conflict. Most notably, physical growth, ageing or entering a room with reverberant acoustics can profoundly alter the sensory cues that guide the brain’s construction of spatial representations. To maintain auditory and visual-spatial maps in co-registration the brain needs to constantly recalibrate the senses.1

Recalibrating auditory and visual-spatial maps is particularly challenging because the two sensory systems encode space not only in different reference frames (i.e. eye vs. head-centred) but also in different representational formats2-4. In vision, spatial location is encoded directly in the sensory epithelium and retinotopic maps of visual cortices (‘place code’)5. In audition, azimuth location is computed mainly from interaural time and level differences in the brain stem.2 In primate auditory cortices, sound location is thought to be encoded by activity differences between two neuronal populations, broadly tuned to ipsi- or contra-lateral hemifields (i.e. ‘hemifield code’)6-9. Less is known about the coding principles at successive processing stages in parietal or prefrontal cortices, in which the hemifield code may be converted into a place code with narrow spatial tuning functions comparable to vision (e.g. ventral intraparietal area).10

Mounting behavioural research illustrates the brain’s extraordinary ability for cross-sensory plasticity. Most prominently, exposure to synchronous, yet spatially misaligned audiovisual signals biases the observer’s perceived sound location towards a previously presented visual stimulus—a phenomenon coined ventriloquist aftereffect11-19. This cross-sensory adaptation to intersensory disparities emerges at multiple timescales ranging from milliseconds13,18,20 to minutes16,17,19 or even days21.

Yet, the underlying neural circuitries, mechanisms and representations that support cross-sensory plasticity remain unknown. The frequency selectivity of audiovisual recalibration has initially pointed towards early stages in tonotopically organized auditory cortices (see ref. 16 but ref. 14). By contrast, the involvement of hybrid spatial reference frames that are neither eye- nor head-centred suggested a pivotal role for parietal cortices or inferior colliculus.3,22

Importantly, recalibration may arise at multiple levels affecting spatial and choice-related computations. While the former alters the neural encoding of sound location irrespective of the task, the latter affects the read-out of decisional choices from those neural representations. The dissociation of the two is challenging for behavioural research that is forced to estimate recalibration from behavioural responses. Likewise, previous neuroimaging studies were not able to disentangle the two, because they employed a spatial localization task that maps each sound location onto one particular response choice, thereby conflating spatial and choice-related processes13,19,20. To dissociate changes in spatial and decisional representations, we need neuroimaging experiments that map different sound locations onto identical decisional choices14 (i.e. classification tasks).

In this work, we show that changes in audiovisual statistics shape neural representations in a spatial classification task on auditory stimuli. Combining model-based analyses of regional BOLD-responses and fine-scaled EEG and fMRI activity patterns we reveal that audiovisual recalibration relies on distinct spatial and decisional codes that are expressed with opposite gradients and time courses across the auditory processing hierarchy. Our findings suggest that regions within the auditory processing hierarchy multiplex spatial and decisional codes to adapt flexibly to the changing sensory statistics in the environment.

**Results**

For each participant, we performed psychophysics, functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) experiments for 13 days (Fig. 1a). Each experiment included unsensory auditory pre-adaptation, left (i.e. VA) and right (i.e. AV) audiovisual adaptation, and auditory postVA/ postAV-adaptation phases. During pre- and post-adaptation phases, participants were presented with auditory stimuli at 7 spatial locations (±12°, ±5°, ±2° and 0°) along the azimuth. They performed a left-right spatial classification task only on 22% of ‘response trials’ that were randomly interspersed in ‘non-response trials’ (Fig. 1b). During the audiovisual adaptation phases, participants were presented with a sound in synchrony with a visual stimulus that was displaced in separate sessions by 15° to the left or right of the sound location. To focus on implicit perceptual recalibration, observers performed a non-spatial task during the adaptation phases: they detected small changes in the contrast to the visual stimulus that occurred on 10% ‘response trials’ (see Fig. 1b). Direct comparison of postVA and postAV-adaptation trials enables the assessment of recalibration unconfounded by time and learning effects.

**Behavioural results.** Throughout all experiments, observers were attentive and followed task instructions as indicated by a sensitivity index (d’) of >4 across all experimental phases (Supplementary Table 1). Observers responded to >88% of ‘response trials’ (i.e. ‘hits’) both in the pre-/post-adaptation (i.e. 22% response trials) and the adaptation phases (i.e. 10% response trials, Supplementary Table 1). They made only <1.43% responses to ‘non-response trials’ (i.e. ‘false alarms’).

Likewise, eye movement analyses during the psychophysics experiment showed that fixation was well maintained on 94.02 ± 1.00% (mean ± SEM) of the trials throughout the entire experiment with no significant differences in eye movement indices between postVA- and postAV-adaptation phases (i.e. % saccades, % eye blinks and post-stimulus mean horizontal fixation position, see Supplementary Note 1: Eye movement results). Potential recalibration effects are therefore not confounded by eye movements.

To assess whether exposure to misaligned visual signals recalibrates observers’ sound representations, we fitted cumulative Gaussians as psychometric functions to the percentage ‘perceived right’ responses in the auditory post-adaptation phases (Fig. 2a). Consistent with previous research15-17 observers’ perceived sound locations shifted towards the displaced visual stimulus that was presented in the prior adaptation phases, causing the psychometric function to shift in the opposite direction (see Supplementary Fig. 1). Bayesian model comparison confirmed that the recalibration model that allows for shifts in the point of subjective equality (PSE) across pre-, postVA- and postAV-adaptation phases was exceedingly more likely than the static model in which the PSE values were constrained to be equal (random effects analysis with Akaike criterion as an approxima-

**tions to the model evidence yielded protected exceedance probabilities >0.90 in each of the psychophysics, fMRI and EEG experiments; Supplementary Table 2). Moreover, paired t-tests of the PSE values (from the recalibration model) between postVA- and postAV-adaptation phases revealed significantly more positive PSE for postVA-adaptation than for postAV-adaptation phases (one-tailed p-values for the second level bootstrap-based paired t-tests: psychophysics: t(14) = 11.4, p < 0.0001; fMRI: t(4) = 9.4 p = 0.010; EEG: t(4) = 9.2, p = 0.011) (Fig. 2b). This was also consistently reflected in participants’ mean response times across the three experiments (see Supplementary Fig. 2). Collectively, our behavioural results confirmed that prior

---

**ARTICLE**

**NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-022-31549-0**

---

2 NATURE COMMUNICATIONS | (2022) 13:3924 | https://doi.org/10.1038/s41467-022-31549-0 | www.nature.com/naturecommunications
Exposure to disparate audiovisual signals recalibrates observers' spatial classification responses to sounds.

fMRI: spatial encoding and recalibration indices. Using multivariate decoding we identified brain areas that (i) encoded sound location and (ii) recalibrated this encoded sound location. Guided by previous research we focused on five regions of interest (ROI): Heschl's gyrus (HG), higher auditory cortex (hA, mainly planum temporale), intraparietal sulcus (IPS), inferior parietal lobule (IPL) and frontal eye-field (FEF). In each of those ROIs we trained a linear support vector regression model (SVR, LIBSVM) in a four-fold cross-validation scheme to learn the mapping from BOLD-response patterns of the pre-adaptation phase to external auditory locations. We used this learnt mapping to decode the sound location from the activation patterns of the pre-, postVA- and postAV-adaptation phases. Focusing selectively on the pre-adaptation trials, we observed significantly better than chance decoding accuracies (i.e. 'spatial encoding index', the Fisher z-transformed Pearson correlation coefficient between the true and the decoded sound locations) across all ROIs along the auditory spatial processing hierarchy with a maximal decoding accuracy in planum temporale (t-values and FDR-adjusted one-tailed p-values for the second level bootstrap-based one-sample t-tests: HG: t(4) = 2.47, p = 0.028; hA: t(4) = 6.58, p = 0.020; IPS: t(4) = 4.04, p = 0.020; IPL: t(4) = 4.27, p = 0.020; FEF: t(4) = 2.84, p = 0.023, see Fig. 3a).

To assess whether these encoded sound locations adapt to displaced visual stimuli, we computed the recalibration index (RI) as the difference between the fraction of 'decoded right' (i.e. positive azimuth) in the postAV-adaptation phase minus the pre- and postVA- and postAV-adaptation phases. Focusing selectively on the pre-adaptation trials, we observed significantly better than chance decoding accuracies (i.e. 'spatial encoding index', the Fisher z-transformed Pearson correlation coefficient between the true and the decoded sound locations) across all ROIs along the auditory spatial processing hierarchy with a maximal decoding accuracy in planum temporale (t-values and FDR-adjusted one-tailed p-values for the second level bootstrap-based one-sample t-tests: HG: t(4) = 2.47, p = 0.028; hA: t(4) = 6.58, p = 0.020; IPS: t(4) = 4.04, p = 0.020; IPL: t(4) = 4.27, p = 0.020; FEF: t(4) = 2.84, p = 0.023, see Fig. 3a).

To assess whether these encoded sound locations adapt to displaced visual stimuli, we computed the recalibration index (RI) as the difference between the fraction of 'decoded right' (i.e. positive azimuth) in the postAV-adaptation phase minus the pre- and postVA- and postAV-adaptation phases. Focusing selectively on the pre-adaptation trials, we observed significantly better than chance decoding accuracies (i.e. 'spatial encoding index', the Fisher z-transformed Pearson correlation coefficient between the true and the decoded sound locations) across all ROIs along the auditory spatial processing hierarchy with a maximal decoding accuracy in planum temporale (t-values and FDR-adjusted one-tailed p-values for the second level bootstrap-based one-sample t-tests: HG: t(4) = 2.47, p = 0.028; hA: t(4) = 6.58, p = 0.020; IPS: t(4) = 4.04, p = 0.020; IPL: t(4) = 4.27, p = 0.020; FEF: t(4) = 2.84, p = 0.023, see Fig. 3a).

To assess whether these encoded sound locations adapt to displaced visual stimuli, we computed the recalibration index (RI) as the difference between the fraction of 'decoded right' (i.e. positive azimuth) in the postAV-adaptation phase minus the pre- and postVA- and postAV-adaptation phases. Focusing selectively on the pre-adaptation trials, we observed significantly better than chance decoding accuracies (i.e. 'spatial encoding index', the Fisher z-transformed Pearson correlation coefficient between the true and the decoded sound locations) across all ROIs along the auditory spatial processing hierarchy with a maximal decoding accuracy in planum temporale (t-values and FDR-adjusted one-tailed p-values for the second level bootstrap-based one-sample t-tests: HG: t(4) = 2.47, p = 0.028; hA: t(4) = 6.58, p = 0.020; IPS: t(4) = 4.04, p = 0.020; IPL: t(4) = 4.27, p = 0.020; FEF: t(4) = 2.84, p = 0.023, see Fig. 3a).
postVA-adaptation phase. If the decoded sound location shifts towards the disparate visual signal during the AV- and VA-adaptation phase, we would expect a positive recalibration index. Consistent with this conjecture we observed a RI that was significantly greater than zero in nearly all ROIs (t-values and FDR-adjusted p-value of second level bootstrap-based one-sample t-tests, HG: t(4) = 3.55, p = 0.042, hA: t(4) = 3.33, p = 0.010, IPL: t(4) = 4.31, p = 0.026, FEF: t(4) = 2.49, p = 0.047, and a trend in IPS: t(4) = 1.81, p = 0.086, see Fig. 3a).

These results provide initial evidence that regions along the dorsal auditory processing hierarchy show an overall effect of recalibration averaged across all sound locations.

fMRI: neurometric functions. Similar to our behavioural analysis, we fitted cumulative Gaussians as neurometric functions to the percentage ‘decoded right’ (i.e. positive azimuth) separately for the pre-, postVA- and postAV-adaptation phases at the group level (see Fig. 3b). Consistent with our behavioural findings the neurometric functions shifted rightwards after VA-adaptation and leftwards after AV-adaptation. Bayesian model comparison provided strong evidence for the recalibration model that allows for changes in PSE values between pre, postVA- and postAV-adaptation phases relative to a static model in which the PSE values were constrained to be equal (Akaike criterion > 2.3 (evidence in favour of the recalibration model) in all ROIs; HG: 10.6, hA: 15.5, IPS: 12.7, IPL: 8.0, FEF: 5.6; Supplementary Table 3).

fMRI: representational similarity analysis. Our linear decoding results indicate that auditory representations throughout the processing stream are altered by prior AV recalibration. Critically, our results so far are agnostic about the coding principles underlying this pervasive cross-sensory plasticity. Next, we therefore characterized the geometry of the neural representations across the seven sound locations using representational similarity analysis (RSA). For visualization, we projected the group-level representational dissimilarity matrices (RDMs, averaged across participants) with non-classical multidimensional scaling (MDS) onto a single dimension to incorporate the spatial organization along the azimuth (i.e. as physical space, see Fig. 3c). Particularly, the higher auditory cortex (hA, including planum temporale) and IPS encoded the seven sound locations largely consistent with the physical distances of the sound locations. For instance, in hA, IPS and IPL the MDS distance is greater for the sound locations 12° and 5° than for the locations 5° and 0°. This was also reflected in the Spearman’s rank-correlation (R_s) between MDS projected and true spatial locations (hA: R_d(5) = 1; IPS: R_d(5) = 1; IPL: R_d(5) = 1). By contrast, the representational geometry in HG and FEF does not fully correspond to the physical sound distances (HG: R_d(5) = 0.86; FEF: R_d(5) = 0.54; see Fig. 3c bottom row). Across all regions, the MDS shows spatial shifts to the left (resp. right) after VA- (resp. AV)-adaptation. Again, the similarity between neural representations in hA, IPS, and IPL obeyed the physical order of the sound locations (hA:
Fig. 3 fMRI multivariate decoding and representational dissimilarity matrices. a Across-subjects’ mean (+SEM, n = 5 subjects) spatial encoding index (left bars, Fisher z-transformed correlation coefficient) and across-subjects’ mean (+SEM, n = 5 subjects) recalibration index (right bars, difference in fraction ‘decoded right’ (i.e. positive azimuth) for postAV-adaptation vs. postVA-adaptation) in the ROIs as indicated. FDR-corrected one-tailed p-values from one-sample bootstrap-based t-tests for spatial encoding index: HG: p = 0.028; hA: p = 0.020; IPS: p = 0.020; IPL: p = 0.020; FEF: p = 0.024; and recalibration index: HG: p = 0.042; hA: p = 0.010; IPL: p = 0.025; FEF: p = 0.046; IPS: p = 0.086. Asterisks denote alpha level *p ≤ 0.05, **p ≤ 0.01. The individual data points are overlaid as black circles. b Neurometric function fitted to the fMRI decoded sound locations across ROIs for pre-, postVA- and postAV-adaptation phases. c RDMs (across-subjects’ mean) showing the Mahalanobis distances for the fMRI decoded sound locations across ROIs. Top row: 7 × 7 RDM for pre-adaptation phase. Bottom row: 14 × 14 matrix for postVA- and postAV-adaptation. The drawings on the left indicate the organization of the RDMs of spatial locations for pre- (top) as well as for postVA- and postAV-adaptation phases (bottom). Using MDS we projected each RDM onto a single dimension (i.e. representing space along the azimuth). For illustration purposes, we vertically offset the MDS projections for postVA- and postAV-adaptation phases particularly in hA and IPS, while a more complex relationship arises in IPL and FEF. HG Heschl’s gyrus, hA higher auditory cortex, IPL inferior parietal lobule, IPS intraparietal sulcus, FEF frontal eye-field, RDM representational dissimilarity matrix, MDS multidimensional scaling, FDR false discovery rate. Source data are provided as a Source Data file.

R_{0}(12) = 0.91; IPS: R_{0}(12) = 0.92; IPL: R_{0}(12) = 0.93), while this is less clear for HG and FEF (HG: R_{0}(12) = 0.68; FEF: R_{0}(12) = 0.41). For instance, in FEF the neural representations for the ‘−12° sound’ are shifted towards the centre (see Fig. 3c, FEF in bottom row).

Model-based fMRI analysis: dissociating perceptual and decisional codes. MDS revealed subtle differences in representational geometry across regions that may arise from the mixing of multiple representational components. Most notably, recalibration may affect spatial and decisional (i.e. choice-related) representations. To arbitrate between these hypotheses, we compared a spatial, a decisional and a combined spatial + decisional uncertainty model as explanations for the regional mean BOLD response and/or the fine-scale activation patterns using pattern component modelling (PCM).
Pattern component modelling is a relatively recent methodological approach that, similar to related multivariate analysis approaches (e.g. multivariate multiple regression or multivariate analysis of variance) models multivariate observations (i.e. multivoxel activity patterns) by a set of explanatory variables (e.g. experimental conditions) under Gaussian assumptions. Critically, it does not fit individual voxel weights but second-order parameters that determine the similarity structure or distribution of activity profiles over voxels as incorporated in the second-moment matrix. By marginalizing over all possible voxel-specific feature weights PCM obtains the marginal likelihood of the data given the model that can be used for model comparison (for further details see refs. 32,33). PCM thus allows us to adjudicate whether the activity distribution over voxels is better accounted for by spatial, decisional or combined spatial + decisional coding.

To model spatial/perceptual representations we used the hemifield model (see Fig. 4a) that encodes sound location in the relative activity of two subpopulations of neurons, each broadly tuned to the ipsi- or contra-lateral hemifield6,34. Because the hemifield and place code models make near-indistinguishable predictions for the pattern similarity structure over the relatively central locations used in the current study, we applied the hemifield model as a generic spatial model to all regions (for details, see Supplementary Methods: Comparison of spatial hemifield and place code models). The decisional uncertainty model (see Fig. 4a) codes observers’ decisional uncertainty as a non-linear function of the distance between the spatial estimate and the decisional classification boundary35 (see Methods for further details).

Because a region may combine spatial and decisional coding, we first assessed whether (i) the spatial model (S), (ii) the decisional uncertainty model (D) or (iii) the combined spatial + decisional uncertainty model (S + D) were the best explanations for our data from the pre-adaptation phase. In a second step, we used the combined spatial + decisional uncertainty model to assess the contributions of the spatial and decisional components to recalibration by comparing: (i) spatial + decisional with no recalibration in either model component (S + D), (ii) spatial with recalibration + decisional (S + D), (iii) spatial + decisional with recalibration (S + D), and (iv) spatial with recalibration + decisional with recalibration (S + D) as explanations for the data from the pre-, postVA- and postVA-adaptation phase. Each
model component accounted for recalibration by shifting the ‘encoded sound locations’ by a constant ±2.3° (i.e. across-subjects’ mean in behavioural PSE shift) to the right (postVA-adaptation) vs. left (postAV-adaptation). As shown in Fig. 4, the spatial and decisional uncertainty models make distinct predictions for the regional mean BOLD-response (Fig. 4b) and the similarity structure of activity patterns over sound locations (Fig. 4c). In supplementary analyses, we also explored whether additional model components that account for observers’ binary left/right decisional choices account for additional variance in our data (see Supplementary Note 2: Assessment of decisional choice model in fMRI and EEG). In the main text we focus mainly on the
Regional mean BOLD-response—linear mixed-effects modelling. The spatial model predicts that the regional mean BOLD-response of—for instance—a left hemisphere region increases for stimuli towards the right hemisphere. Likewise, its response to all sounds irrespective of location should be greater after right recalibration. By contrast, the decisional uncertainty model predicts a mean BOLD-response that peaks at the decisional boundary and tapers off with a greater distance from the boundary. As a result of recalibration, this peak shifts towards the left or right, because the brain’s spatial estimates have been recalibrated leading to a change in the relative distance between the spatial estimates and the decision boundary (see Fig. 4b). For instance, the encoded spatial estimate for a sound stimulus at the physical location of 5° may be shifted towards the left after left recalibration (i.e. visual signal displaced towards the left of the auditory location). As a result, the decisional boundary and spatial location that is associated with the greatest decisional uncertainty is shifted towards the right after left recalibration (see Supplementary Fig. 1).

In hA the mean BOLD-response increased progressively for stimuli along the azimuth as expected under the spatial model (Fig. 5a). By contrast, in IPS, IPL and FEF the BOLD-response peaked for physical 0° sound location in the pre-adaptation phase, left sound locations (i.e. <0°) in the postVA-adaptation and right sound locations (i.e. >0°) in the postVA-adaptation phase as expected under the decisional uncertainty model. This visual impression was confirmed by Bayesian model comparison of the three (i.e. spatial, decisional, spatial + decisional) linear mixed-effects (LME) models that included the simulated activity of the respective models as fixed effects predictors (with the additional constant term) and were fitted to the pre-adaptation phase alone. The Bayes factors provided strong evidence for the spatial model in hA and moderate evidence for the decisional uncertainty model in IPS (see Supplementary Table 4). In HG, IPL and FEF the most parsimonious baseline model outperformed or was on par with the other models suggesting that the regional BOLD-response in those regions did not provide reliable information about sound location in the pre-adaptation runs. For the recalibration analysis, we designed each LME model using the simulated activity of the spatial and/or decisional uncertainty models as fixed effects predictors (in addition to a constant regressor) for the 7 (sound locations) × 3 (pre-, postVA- and postVA-adaptation) conditions. We observed a small trend toward spatial coding in hA (i.e. S_R + D and S_R + D_h > S + D and S + D_h), but moderate evidence towards decisional coding in IPL and strong evidence in IPS and FEF (i.e. S_R + D_h and S + D_h > S + D and S_R + D, Fig. 5b bottom row, Supplementary Table 4). This suggests that the mean regional BOLD-response follows the predictions of the hemifield model only in hA. By contrast in IPS, IPL and FEF the regional mean BOLD-response mainly reflects activity associated with the decisional uncertainty invoked by mapping the recalibrated spatial estimates onto observers’ decisional boundary. Adding predictors encoding observers’ decisional spatial choices did not increase the model evidence in any of those regions (see Supplementary Note 2: Assessment of decisional choice model in fMRI and EEG).

Fine-scale activation pattern—pattern component modelling. Using pattern component modelling (PCM) we investigated whether spatial and decisional codes contributed to the fine-scale voxel-level activation patterns and the recalibration effects. Consistent with our analysis of the regional mean BOLD-response we fitted and compared a baseline model that assumes independence of activation patterns across conditions with the spatial, decisional, and combined spatial + decisional PCM models. All these models were fitted to the 7 (sound locations) in the pre-adaptation conditions. In a second step we used the combined spatial + decisional uncertainty model to assess whether recalibration was expressed in spatial, decisional, or spatial + decisional codes using data from the pre-, postVA- and postVA-adaptation phases (see Fig. 5c). We fitted all models (e.g. spatial and/or decisional etc.) and a fully flexible model using a leave-one-subject-out cross-validation scheme. The fully flexible model accommodates any possible covariance structure of activity profiles and enables us to assess whether a particular target model accounts for the key representational structure in the observations given the measurement noise and inter-subject variability (see Fig. 5c).

As shown in Fig. 5c top row, the Loge-Bayes factors (averaged across participants ±SEM) provided strong evidence for the spatial relative to the decisional uncertainty model in HG and hA (see Supplementary Table 5 for exact values). In HG, the spatial component alone was sufficient to reach the threshold set by the fully flexible model. In all other regions, i.e. hA, IPS, IPL and FEF, the combined spatial + decisional uncertainty model outperformed the single-component models suggesting that in those regions spatial and decisional representations together contribute to adaptive coding. Moreover, the decisional component became progressively more dominant along the dorsal processing hierarchy. In FEF the decisional uncertainty model even outperformed the spatial model.

The recalibration results further emphasized these opposite gradients for spatial and decisional coding along the auditory processing stream. As shown in Fig. 5c bottom row, in HG and hA the models expressing recalibration in spatial coding (S_R + D and S_R + D_h) outperformed those without recalibration in spatial coding (S + D and S + D_h). The evidence was strong in hA, but
not conclusive in HG. In hA, the combined model incorporating recalibration of spatial and decisional coding outperformed more parsimonious models. Conversely in IPS, IPL and FEF we observed strong evidence for the models that expressed recalibration in the decisional code \((S + D_{R} \text{ and } S_{R} + D_{R})\) relative to those that did not \((S + D \text{ and } S_{R} + D_{R})\), see Supplementary Table 5 for exact values). Again, in IPL and IPS the combined model \((S_{R} + D_{R})\) outperformed the more parsimonious models.

In summary, our advanced model-based representational analyses demonstrate that audiovisual adaptation relies on plastic changes in spatial and decisional uncertainty representations. While all regions (apart from HG) multiplexed spatial and decisional uncertainty representations, the combined model incorporating recalibration of spatial and decisional coding outperformed more parsimonious models in hA and IPS, IPL and FEF.
decisional coding, their contributions arose with opposite gradients along the dorsal auditory hierarchy. The spatial model provided a better explanation for activations in HG and hA, the decisional uncertainty model dominated in IPS, IPL and FEF. Adding a decisional choice component (i.e. observers’ left/right choices) substantially increased the model evidence only in IPS for spatial coding in pre-adaptation runs and in IPS and hA for recalibration (see Supplementary Note 2: Assessment of decisional choice model in fMRI and EEG, Supplementary Methods: Comparison of response time and decisional choice models, and Supplementary Tables S6 and S7). These results suggest that IPL may be more involved in computing decisional uncertainty, while IPS codes observers’ decisional spatial choice, which in turn influences spatial coding in hA via feedback connectivity. Crucially, only characterizing the representational geometry of the neural representations enabled us to reveal that spatial and decisional codes evolve with opposite gradients across regions. This highlights the critical importance to move beyond simple linear decoding analyses.

**EEG: spatial encoding and recalibration indices.** Using EEG, we characterized how spatial and decisional coding evolved post-stimulus. Consistent with our fMRI analysis, we first computed the spatial encoding and recalibration indices for EEG activity patterns across time. We trained linear support vector regression models (SVR, LIBSVM) on the EEG activity patterns of the pre-adaptation trials in overlapping 50 ms time windows sliding from −100 to 500 ms post-stimulus and generalized them to trials from pre- and post-adaptation trials in a four-fold cross-validation scheme. The cluster-based bootstrap test on the Fisher z-transformed Pearson correlation coefficient between the true and the decoded sound locations revealed a significant cluster extending from 110 to 500 ms post-stimulus (one-tailed \( p = 0.01 \)) corrected for multiple comparisons within the entire [−50 to 500] ms window. The decoding accuracy was significantly better than chance from about 110 ms post-stimulus, rose steadily and peaked at about 355 ms (Fig. 6a). Likewise, we assessed the recalibration index (RI) within the time window that showed a significant effect of spatial encoding, i.e. [110−500] ms post-stimulus. In the cluster-based bootstrap test the RI was significantly positive in two clusters from [185−285] ms (one-tailed \( p = 0.019 \)) and [335−470] ms (one-tailed \( p = 0.005 \)) post-stimulus (Fig. 6b). Moreover, because previous ERP analyses suggested that the N100 potential is affected by recalibration, we performed a temporal ROI analysis selectively on the EEG activity pattern averaged within the N100 time window (i.e. [70−130] ms). This temporal ROI analysis showed that the decoding accuracy was significantly better than chance (mean ± SEM: 0.125 ± 0.059; \( t(4) = 2.12, \) one-tailed \( p = 0.0406 \)) and the RI was significantly greater than zero (mean ± SEM: 5.963 ± 1.743; \( t(4) = 3.42, \) one-tailed \( p = 0.0019 \)).

**Model-based EEG analysis: dissociating perceptual and decisional codes.** Combining EEG and PCM we investigated whether spatial and decisional representations evolved with different time courses. For this, we compared the null, spatial, decisional and combined spatial + decisional uncertainty models as explanations for EEG activity patterns over spatial locations in four consecutive time windows: [50−150], [150−250], [250−350] and [350−450] ms post-stimulus (Fig. 6c). From [50−150] ms the spatial model alone provided a sufficient explanation of the EEG data; however, including the decisional uncertainty model as well did explain moderately more evidence. From [150−250] ms and beyond the spatial + decisional uncertainty model performed better than the spatial model with LogeBFs of at least 6.1 (Supplementary Table 8). These results suggest that the decisional code becomes progressively more important at later processing stages. This representational gradient across post-stimulus time was also apparent in our recalibration results (Fig. 6c, bottom row). Spatial recalibration was expressed mainly via spatial coding (i.e. \( S_R + D_R \) and \( S_R + D_A \) better than \( S + D \) and \( S + D_A \) by LogeBFs of at least 2.4, see Supplementary Table 8) from 150 to 250 ms. From 250 ms onwards, recalibration relied jointly on spatial and decisional codes (i.e. \( S_R + D_R \) explained the most variance consistently), as in our fMRI results in later stages. Consistent with the fMRI results, adding the decisional choice component explained additional variance in the EEG activity patterns from 250 to 450 ms (see Supplementary Note 2: Assessment of decisional choice model in IMRI and EEG, Supplementary Methods: Comparison of response time and decisional choice models, and Supplementary Table 9). Collectively, our EEG results revealed that spatial codes are more prominent in earlier processing stages and decisional coding in later ones. Critically, at later stages recalibration was expressed jointly in spatial and decisional codes.

**Fusing EEG and fMRI in pattern component modelling.** The fMRI results revealed a mixture of spatial and decisional codes in several ROIs across the dorsal processing hierarchy. To link fMRI and EEG results more directly, we fused them in PCMs in which the second-moment matrices of the BOLD-response activation patterns in one of our five ROIs (i.e. HG, hA, IPS, IPL or FEF) formed the predictor for the distribution of EEG activity patterns.
Discussion

This study demonstrates that the brain recalibrates the senses by flexibly adapting spatial and decisional codes, which are expressed with opposite gradients along the auditory processing hierarchy. Early activity patterns in planum temporale encode sound locations in a continuous space that dynamically shifts towards misaligned visual inputs. Later activity patterns in frontoparietal cortices mainly code choice-related uncertainty in line with these representational changes.

Our behavioural results provide robust evidence that the brain recalibrates auditory space to keep auditory and visual maps in co-representation.

NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-022-31549-0 | www.nature.com/naturecommunications

For spatial coding in human auditory cortices. Yet, we applied it as a model not only to auditory but also to frontoparietal areas because its pattern similarity structure over the limited (i.e. [−12° to 12°]) sound locations is indistinguishable from that of the competing place code model (see Supplementary Fig. 3).

Crucially, Bayesian model comparison showed a functional gradient along the auditory processing stream: the spatial model provided a better explanation for the regional mean and fine-scale activation patterns in planum temporale, while the decisional uncertainty model performed better in frontoparietal cortices. As shown in Fig. 5a, the regional BOLD-response increased linearly for stimuli along the azimuth in planum temporale, but followed an inverted U-shaped function that adapts to audiovisual conflicts in frontoparietal cortices (Fig. 5b top and bottom row). Likewise, the recalibration of the fine-scale activation patterns was better captured by the spatial model in planum temporale, but by the decisional uncertainty model in frontoparietal cortices. Yet, despite this predominance of decisional coding, IPS, IPL and even FEF multiplexed spatial and decisional patterns (Fig. 5c top row).

This spatial–decisional gradient across the cortical hierarchy was mirrored in their temporal evolutions. Recalibration of spatial coding was reflected in early EEG activity patterns ([150–250] ms), while adaptation of decisional uncertainty was more prominent in later activity (after 250 ms, Fig. 6c bottom row). We could even directly link the functional gradients across time and regions by fusing fMRI and EEG via pattern component modeling: the BOLD-response patterns in hA predicted EEG activity patterns mainly at early ([150–250] ms) stages and IPL at later stages from 250 ms onwards (Fig. 6d).

Collectively, our fMRI and EEG results demonstrate that the brain adapts flexibly to the changing statistics of the environment. They also highlight the critical importance to move beyond decoding analyses, as the ability to decode a feature value (e.g. sound location) from activity patterns does not imply that this feature is well represented. Instead, in order to define the features or mixtures of features (e.g. spatial, decisional) that are encoded in activations, we need to characterize the fine-grained representational geometry and assess the explanatory power of different sets of features via model comparison. Only the latter was able to show that later recalibration effects in frontoparietal cortices reflect mainly choice-related rather than genuine spatial coding. The expression of recalibration in distinct spatial and decisional codes opens the intriguing possibility that recalibration may not always impact both codes together. Instead, the representations and neural circuitries involved in recalibration may depend on the particular task and/or the duration of recalibration. For instance, while our study focused on implicit long-term recalibration using a non-spatial task, other studies have shown rapid recalibration effects when observers performed spatial localization tasks during the adaptation and/or post-adaptation phases. It remains unknown whether these short-term recalibration effects involve plastic changes in the spatial representations or selectively alter choice-related activity. In line with this notion, accumulating research suggests that recalibration relies on multiple mechanisms that operate over different timescales. Future studies thus need to disentangle the impact of recalibration duration and task on spatial and decisional coding.

So far, we have emphasized that spatial and decisional codes are expressed with opposite gradients across the cortical hierarchy and post-stimulus time. Yet, Bayesian model comparison also indicated that the combined spatial + decisional uncertainty model significantly outperformed the more parsimonious decisional or spatial models throughout the cortical hierarchy. Likewise, a benefit was observed for the pattern component models in which recalibration was expressed in both spatial and decisional
codes. This mixture of spatial and decisional codes may result from true multiplexing of multiple neural codes concurrently within the same region or from a methodological limitation of fMRI, namely the sluggishness of the BOLD-response whereby neural codes that are present at different latencies nevertheless mix in the BOLD-response patterns. The EEG-fMRI fusion approach suggests that early neural activity until 250 ms is dominated by spatial coding in higher-order auditory cortices. By contrast, later neural activity reflects a mixture of spatial and decisional coding. We therefore suggest that later activity in all higher-order cortical regions (i.e. hA, IPS, IPL and PEF) multiplexed perceptual and decisional codes and adapt both codes to recalibrate the senses\(^{12}\). Moreover, because multiplexing of spatial and decisional codes arose mainly later from 250 ms post-stimulus, the coding of decisional uncertainty in auditory cortices (hA) is likely to reflect top-down influences from fron-totoparietal areas\(^{48,49}\).

In conclusion, our results show that audiovisual adaptation relies on spatial and decisional coding. Crucially, the expression of spatial and decisional codes evolves with different times courses and opposite gradients along the auditory processing hierarchy. Early neural activity in planum temporale encoded sound location within a continuous auditory space. Later frontotoparietal activity mainly encoded observers’ decisional uncertainty that flexibly adapts in accordance with these transformations of auditory space.

**Methods**

This study was conducted in accordance with the Declaration of Helsinki and approved by the research ethics committee of the University of Birmingham (approval number: ERN_11_0470AP4).

**Participants.** Fifteen participants (10 females, mean age = 22.1; SD = 4.1) participated in the psychophysics study. Five of those participants (4 females, mean age = 22.2; SD = 5.1, one author of the study, A.M.) completed the fMRI and EEG experiments (for full selection criteria see Supplementary Methodic: Participants). All participants had no history of neurological or psychiatric illnesses, had a normal or corrected-to-normal vision and had normal hearing. Participants gave informed written consent to participate in the study and were compensated with £6 per hour for behavioural and £8 per hour for fMRI and EEG sessions.

**Stimuli.** The auditory stimulus consisted of a burst of white noise with a duration of 50 ms and 5 ms on/off ramp delivered at a 75 dB sound pressure level. To create a virtual auditory signal, the noise was convolved with spatially specific head-related transfer functions thereby providing both binaural and monaural cues (a\(^{12}\)).

Air-MRI, namely the sluggishness of the BOLD-response whereby neural codes that are present at different latencies nevertheless mix in the BOLD-response patterns. The EEG-fMRI fusion approach suggests that early neural activity until 250 ms is dominated by spatial coding in higher-order auditory cortices. By contrast, later neural activity reflects a mixture of spatial and decisional coding. We therefore suggest that later activity in all higher-order cortical regions (i.e. hA, IPS, IPL and PEF) multiplexed perceptual and decisional codes and adapt both codes to recalibrate the senses\(^{12}\). Moreover, because multiplexing of spatial and decisional codes arose mainly later from 250 ms post-stimulus, the coding of decisional uncertainty in auditory cortices (hA) is likely to reflect top-down influences from fron-totoparietal areas\(^{48,49}\).

In conclusion, our results show that audiovisual adaptation relies on spatial and decisional coding. Crucially, the expression of spatial and decisional codes evolves with different times courses and opposite gradients along the auditory processing hierarchy. Early neural activity in planum temporale encoded sound location within a continuous auditory space. Later frontotoparietal activity mainly encoded observers’ decisional uncertainty that flexibly adapts in accordance with these transformations of auditory space.

**Methods**

This study was conducted in accordance with the Declaration of Helsinki and approved by the research ethics committee of the University of Birmingham (approval number: ERN_11_0470AP4).

**Participants.** Fifteen participants (10 females, mean age = 22.1; SD = 4.1) participated in the psychophysics study. Five of those participants (4 females, mean age = 22.2; SD = 5.1, one author of the study, A.M.) completed the fMRI and EEG experiments (for full selection criteria see Supplementary Methodic: Participants). All participants had no history of neurological or psychiatric illnesses, had a normal or corrected-to-normal vision and had normal hearing. Participants gave informed written consent to participate in the study and were compensated with £6 per hour for behavioural and £8 per hour for fMRI and EEG sessions.

**Stimuli.** The auditory stimulus consisted of a burst of white noise with a duration of 50 ms and 5 ms on/off ramp delivered at a 75 dB sound pressure level. To create a virtual auditory signal, the noise was convolved with spatially specific head-related transfer functions thereby providing both binaural and monaural cues (a\(^{12}\)).

**Eye movement recording.** Eye movement recordings were calibrated in the recommended field of view (32° horizontally and 24° vertically) for the EyeLink 1000 Plus system (SR Research Ltd.) with the desktop mounted at a sampling rate of 1400 x 1050 pixels at a frame rate of 60 Hz. The screen was visible to the subject through a mirror mounted on the magnetic resonance (MR) head coil and the eye-to-screen distance was 68 cm. Auditory stimuli were delivered via a pair of MR-compatible headphones (MR Confor HP-VS03, Cambridge Research Systems Ltd.). Participants responded using a two-button MR-compatible keypad (LXPAD v 5±1.6, NATA Technologies). Exact audiovisual onset timing in adaptation trials was confirmed by recording visual and auditory signals concurrently with a photodiode and a microphone.

**Experimental setup.** In all experiments, visual and auditory stimuli were presented using PsychoBox version 3.0.1151\(^{82}\) under MATLAB R2011b (MathWorks Inc.) on a MacBook Pro running Mac OSX 10.6.8 (Apple Inc.). In the psychophysics and EEG experiments, participants were seated at a desk with their heads rested on a chinrest. Two accessory rods were mounted on the chinrest serving as forehead rest and allowing stable and reliable head positioning. Visual stimuli were presented at a viewing distance of 60 cm via a gamma-corrected 24° LCD monitor (ProLite B2483HS, iyama Corp.) with a resolution of 1920 x 1080 pixels at a frame rate of 60 Hz. Auditory stimuli were delivered via circumaural headphones (HD 280 Pro, Sennheiser electronic GmbH & Co. KG) in the psychophysics experiment and via earphones (E-A-RTONE GOLD, 3M Company Auditory Systems) in the EEG experiment. Participants used a standard USB keyboard for responding. In the fMRI experiment, visual stimuli were back-projected to a plexiglass screen using a D-ILA projector (DILA-SX21, JVC, JVC/KENWOOD UK Ltd.) with a resolution of 1400 x 1050 pixels at a frame rate of 60 Hz. The screen was visible to the subject through a mirror mounted on the magnetic resonance (MR) head coil and the eye-to-screen distance was 68 cm. Auditory stimuli were delivered via a pair of MR-compatible headphones (MR Confor HP-VS03, Cambridge Research Systems Ltd.). Participants responded using a two-button MR-compatible keypad (LXPAD v 5±1.6, NATA Technologies). Exact audiovisual onset timing in adaptation trials was confirmed by recording visual and auditory signals concurrently with a photodiode and a microphone.

**Signal detection measures.** Participants responded on a fraction of ‘response trials’, i.e. 22% in auditory pre- and post-adaptation and 10% in AV- and VA-adaptation. This enabled us to assess their performance with the signal sensitivity measure d’. The formula for calculating d’ is given by:

\[
d’ = Z(p_{hit}) - Z(p_{false alarm}) \tag{1}
\]

where \(p_{hit}\) and \(p_{false alarm}\) are the hit and false alarm rates, respectively. Hits are ‘response trials’, on which observers gave a response. False alarms are ‘non-response trials’, in which observers gave a response. 100% hit rate and 0% false alarm rate were approximated by 99.999% and 0.001%, respectively, to enable the calculation of Z-scores.
Psychometric functions. For the pre-, postVA- and postAV-adaptation phases we fitted cumulative Gaussians as psychometric functions (PF) to the percentage ‘correct right’ responses (the response time to a function stimulus at a location (±12°, ±5°, ±2° and 0°) (see Fig. 2). To account for potential overdispersion in the data across sessions, we used the beta-binomial model53,54. The beta-binomial model assumes that the percentage of right responses at a particular stimulus location is not fixed throughout the entire experiment but a beta-distribution random variable with variance determined by the scaling factor η (between 0 and 1). The models were fitted individually to the behavioural data of each participant with maximum likelihood estimation55 and using the beta-binomial model56,57. To enable reliable parameter estimates for each participant, we employed a multi-condition fitting approach. This approach constrains (i) the just noticeable difference (JND or slope parameter) were set equal across all conditions (i.e., pre-, postVA- and postAV-adaptation phases); (ii) guess and lapse rates were set equal to each other and (iii) equal across all conditions. Furthermore, (iv) we constrained the fitted guess and lapse rate parameters to be within 0 and 1 and (v) the variance of the beta-distribution (i.e. η in the beta-binomial model) was set equal across all conditions. For statistical inference, we assessed the goodness of fit of the cumulative Gaussian for each condition and participant using a likelihood ratio test. This likelihood ratio test compares the likelihood of participants’ responses given the model that is constrained by a cumulative Gaussian function (i.e. our ‘target model’) to the likelihood given by a so-called ‘saturated model’ that models observers’ responses with one parameter for each stimulus location in each condition. The resulting likelihood ratio for the original dataset is then compared with a null distribution of likelihood ratios generated by parametrically bootstrapping the data (5000×) from the ‘target model’ fitted to the original dataset and refitting the ‘target’ and ‘saturated’ models. Since the likelihood ratio for the original dataset was not any smaller in any of the participants than 5% of the parametrically bootstrapped likelihood ratios (i.e. p>0.05), we inferred sufficient goodness of fit for all participants.

Next, we assessed whether AV- and VA-adaptation induced a shift in participants’ perceived auditory location by comparing a ‘static’ model, which constrains the EIs to be the same for pre-, postVA- and postAV-adaptation phases, against a ‘recalibration’ model, which includes three PSE values for the pre-, postVA- and postAV-adaptation PFSs. For each participant and model, we calculated the Akaike Information Criterion (AIC)59 according to the following formula:

\[
\text{AIC} = -2 \ln L - N
\]

where \(L\) stands for the likelihood of the model given the data, \(N\) is the number of parameters and \(n\) is the natural logarithm. The summed AIC values over participants for each model provided an approximation to the model evidence. We performed Bayesian model comparison at the level of the random effect as implemented in SPM12 to obtain the protected exceedance probability for the candidate models58. A uniform prior was used over candidate models.

fMRI data acquisition and analysis

fMRI data acquisition. We used a 3T Philips Achieva scanner to acquire both T1-weighted anatomical images (TR/TE/TI, 7.4/3.5/min. 989 ms; 176 slices; image spatial resolution, 3 × 3 × 3 mm\(^3\) voxels). On each of the 4 days we acquired 5 fMRI runs (20 auditory pre-adaptation runs and four audiovisual adaptation runs. Each fMRI run consisted of 70 fMRI volumes. We discarded to allow for T1 equilibration effects.

The resulting contrast images for the original dataset is then compared with all 21 sound conditions relative to fixation baseline (for identification of sound-responsive voxels). To determine whether a ROI encodes auditory spatial representations, we computed the Pearson correlation coefficients between the true and the decoded auditory locations for the pre-adaptation runs for each participant as a ‘spatial encoding index’.

To determine whether auditory spatial representations in a region of interest are recalibrated by misaligned visual signals, we binarized the predicted auditory locations into left vs. right predictions and compared the difference in the fraction of ‘decoded right responses’ (i.e. positive azimuth) between auditory postVA- and postAV-adaptation phases as ‘recalibration index’ (RI).

fMRI multivariate decoding—spatial encoding and recalculation indices. We extracted the voxel response patterns in a particular ROI from the pre-whitened and normalized parameter estimate images pertaining to the magnitude of the BOLD-response for each condition and run. To avoid motor confounds, we used the non-motion-corrected images to calculate a ‘target’ and a ‘tailed’ version of the BOLD-response to each of the four auditory stimuli. Because of the lower signal-to-noise ratio of the ‘tailed’ version of the BOLD-response, we limited the number of stimuli to three. We trained a multi-condition mapping between auditory stimuli by comparing a ‘static’ model with a ‘recalibration’ model using the Akaike Information Criterion (AIC)28, for details see Methods: Psychometric functions and Supplementary Methods: fMRI multivariate decoding—neurometric functions.

fMRI multivariate pattern analysis—representational dissimilarity analyses and multidimensional scaling. To generate 21 condition-wise dissimilarity images for the 7 auditory spatial locations × 3 (pre-, postVA- and postAV-adaptation) by averaging parameter estimate images across fMRI runs for each participant. We then characterized the geometry of spatial representations using representational dissimilarity matrices (RDMs)60 based on the Mahalanobis distance for each participant and for each subject separately for pre-adaptation as well as postVA- and postAV-adaptation phases (see Supplementary Methods: fMRI multivariate pattern analysis—representational dissimilarity analyses and multidimensional scaling).
Using non-classical multidimensional scaling (MDS) with non-metric scaling, we projected the group-level RDMs (i.e. averaged across participants) onto a one-dimensional space for visual inspection. To quantify how well the MDS projection of neural representations corresponds to the relative physical positions of stimuli, we computed the Spearman’s rank-correlation between the MDS projections and the corresponding spatial locations separately for pre- and post-adaptation phases in each ROI. As the MDS projection only reflects the relative distance between neural representations, it is agnostic about the absolute orientation of stimuli in physical space (i.e. whether ‘left’ or ‘right’ is associated with a negative or positive sign). Therefore, only the magnitude, not the sign of the rank-correlation is meaningful, hence we report the absolute value of rank-correlation coefficients. For illustration purposes, we flipped the MDS projections around zero consistently for all conditions in ROIs where the correlation between MDS projections and corresponding spatial locations was negative.

EEG data acquisition and analysis

EEG data acquisition. Continuous EEG signals were recorded from 64 channels using Ag/AgCl active electrodes arranged in 10–20 layout (ActiCap, Brain Products GmbH, Gilching, Germany) at a sampling rate of 1000 Hz with FCz as reference.

Between MDS projections and corresponding spatial locations was negative. The rank-correlation coefficient associated with a negative or positive sign). Therefore, only the magnitude, not the MDS projection values in postVA- and postAV-adaptation phases from the psychometric functions.

EEG pre-processing. Pre-processing was performed with the FieldTrip toolbox (http://www.fieldtrip.fu-berlin.de/). Raw data were high-pass filtered at 0.1 Hz, inspected for bad channels, re-referenced to the average of all channels, and low pass filtered at 50 Hz. Bad channels were identified by visual inspection, rejected, and interpolated using a spherical spline interpolation based on the neighbouring channels (number of rejected channels = 1.5%, min = 0.2% max = 4%)

Trials epochs for the unisensory auditory pre-adaptation and the auditory post-adaptation conditions were extracted between [−100 to 500] ms post-stimulus (i.e. the onset of the response cue on the response trials), baseline corrected and down-sampled to 200 Hz. Epochs containing artefacts within the time window of interest (i.e. between [0–500] ms post-stimulus) were identified based on visual inspection and rejected. Furthermore, based on eye-tracking data, trials were rejected if (i) they contained eye blinks or (ii) saccades or (iii) the eye gaze was away from the fixation cross by more than 2 degrees (% rejected trials across-subjects mean ± SEM: 8.2 ± 1.0%). Grand average ERPs were computed by averaging all trials for each condition first within each participant and then across participants.

For the multivariate analysis, we applied spatial multivariate noise normalization to the individual trials using a noise covariance matrix estimated separately for each time point and the optimal shrinkage method. Furthermore, the EEG activity patterns were divided by their Euclidean norm separately for each time point and trial for normalization. Since the EEG responses on trials with and without behavioural response were identical until 500 ms post-stimulus, we pooled over response and no response trials in all EEG analyses.

EEG multivariate decoding—spatial encoding and recalibration indices. Similar to our MDS approach, we trained a SVR in a four-fold stratified cross-validation scheme (C = 1, ν ∈ [0.527]) to learn the mapping from evoked potentials (averages of 16 randomly sampled trials) of the pre-adaptation run to external auditory space over 50 ms time windows, shifting in increments of 5 ms, from −100 to 500 ms post-stimulus. The learnt mapping was used to decode the sound location from the EEG activity patterns of the pre-adaptation examples in the remaining fold and all post-adaptation examples. To minimize sample variance, we averaged the decoding locations across 50 repetitions of this cross-validation procedure to compute the ‘spatial encoding’ and ‘recalibration’ indices for each 50 ms window. At the level of the random effect, we report results from a bootstrap-based t-test against zero corrected for multiple comparisons across [−50 to 500] ms using a cluster-based correction with an auxiliary cluster-defining height threshold of p < 0.05.

Based on our prior hypothesis that spatial location is encoded in the N100 potential, we also performed bootstrap-based t-tests on the spatial encoding and recalibration indices obtained from evoked potentials averaged within a [70–130] ms window.

Spatial hemifield and decisional uncertainty models for fMRI and EEG

We assessed a spatial and a decisional encoding model as explanations for the regional mean BOLD-response and fine-scale fMRI/EEG patterns.

Spatial hemifield model. The hemifield model encodes sound location in the relative activity of two subpopulations of neurons each broadly tuned either to the ipsi- or contra-lateral hemifield (Fig. 5a). We simulated 360 neurons with broad Gaussian tuning functions. The standard deviation was set to 64°. The means of the tuning functions were sampled uniformly from 80° to 100° for the neuronal population tuned to the contra-lateral hemifield and from −80° to −100° azimuth for the neuronal population tuned to the ipsilateral hemifield. Consistent with previous research the ratio of the ipsi- and contralaterally tuned neurons was set to 3:1 (see Supplementary Fig. 3a). Critically, additional simulations indicated that the ratio of ipsi/contralaterally tuned neurons affected only the predictions for the regional mean BOLD-responses but not those for the pattern similarity structure. In fact, the pattern similarity structure is nearly identical for simulations with the ratio between ipsi/contralaterally tuned neurons set to 30%/70% and 50%/50% (see Supplementary Material: Comparison of spatial hemifield and place code models).

For the pre-adaptation conditions we sampled neural responses from the seven sound locations in our paradigm. For the post-adaptation conditions, we sampled again from those seven locations for the non-recalibration model. For the recalibration model we sampled the neural responses from the above locations shifted by 2.3° to the right (postVA-adaptation) or left (postAV-adaptation).

Decisinal uncertainty model. In the decisional uncertainty model the activity of a neuron encodes observers’ choice-related uncertainty that depends non-linearly on the distance between observers’ spatial estimates and their left-right spatial classification boundary according to:

\[
\text{Decisinal uncertainty} = \frac{|F(x; \mu, \sigma) - 0.5| - 0.5}{0.5}
\]

where \(F(x; \mu, \sigma)\) is the cumulative normal distribution function with mean \(\mu\) and standard deviation \(\sigma\) evaluated at spatial location \(x\).

The standard deviations of the cumulative normal distribution were set to 10° and their means were uniformly sampled between −1° and +1° (see Fig. 4a).

We simulated responses from 360 neurons. As exactly for the spatial model, we sampled neural responses from the seven sound locations for the pre-adaptation conditions and the post-adaptation conditions for the non-recalibration model. For the recalibration model we sampled the neural responses from the above locations shifted by 2.3° to the right (postVA-adaptation) or left (postAV-adaptation).

As shown in Fig. 4bc, the spatial and decisional uncertainty models make distinct predictions for the regional mean BOLD-response and the pattern similarity structure over the 21 conditions = 7 spatial locations × 3 phases (pre, postVA, postAV).

While decisional uncertainty depends on the distance of the noisy signal from the decision boundary in signal detection theory models (see above), experimentally it is closely related to observers’ confidence and response times—although differences between decisional uncertainty in first-order perceptual judgments and explicit confidence judgment are well-established.

We have therefore assessed to what extent the decisional uncertainty as computed above is related to observers’ response times for spatial locations during the pre- and post-adaptation phases. As expected, the predictions for the spatial locations in pre- and post-adaptation phases are highly correlated between the decisional uncertainty model and the response time models (Spearman’s rank-correlation (Rs); Rs(19) = 0.95, p < 0.001, see Supplementary Note 3: Pairwise correlations between the predictions of the spatial, decisional uncertainty, decisional choice and response time models; Response time analysis; Supplementary Figs. 2 and 4).

Regional mean BOLD-response: spatial and decisional linear mixed-effects models. Regional mean BOLD-response: For each of the 2 hemispheres (left, right) × 5 ROI (M1, M2, BA4/2, IPS, IPL, FEF) regions we selected the 20 most reliably responsive voxels, i.e. with the greatest t-values for all unisensory sound conditions relative to fixation. For each of those 10 regions we extracted the BOLD-response magnitude for each of the 7 locations × 3 phases (pre-, postVA- and postAV-adaptation) and formed the regional mean.

Linear mixed-effects modelling: To account for localization effects, we performed separate analyses for each region and hemisphere. Separately for each hemisphere we used the spatial (resp. decisional) model to generate predicted regional mean BOLD-responses for each of the 21 conditions = 7 spatial locations × 3 phases (pre-, postAV- and postVA-adaptation) by averaging activations of 360 simulated neurons.

We generated seven linear mixed-effects (LME) models that varied in their fixed effects predictors:

- Null LME: single intercept term.
- Spatial LME model (S): predictor from the spatial encoding model without recalibration and intercept term.
- Decisional LME model (D): predictor from the decisional uncertainty model without recalibration and intercept term.

The remaining LME models included spatial, decisional and intercept terms (i.e. the fixed effects regressors) and factorially manipulated whether the spatial and/or the decisional predictor modelled recalibration:

- (S + D): Spatial without recalibration + decisional without recalibration
- (S + D): Spatial with recalibration + decisional without recalibration
- (S + D): Spatial without recalibration + decisional with recalibration
- (S + D): Spatial with recalibration + decisional with recalibration

Subject-level effects were included as random effects. For each of the 2 hemispheres × 5 ROIs we fitted (1) the Null, S, D, S + D LME models to the pre-
adaptation phase (i.e. 7 conditions, see Fig. 5b top row) and (2) the Null, S + D, S + D, S + D, S + D, S + D, S + D, S + D, S + D, S + D. LME models were fitted using maximum-likelihood estimation and the Bayesian information criterion (BIC) was computed as:

\[
\text{BIC} = \log L - \frac{1}{2} \log N
\]

where \(L\) stands for the likelihood of the model given the data, \(N\) is the number of parameters, \(N\) is the number of observations, and \(\log\) is the natural logarithm. The figures show the natural logarithm of the Bayes factors (Log-Bayes factors) averaged across hemispheres for LME models relative to the null LME (i.e. BIC\(_{\text{null}}\) = BIC\(_{\text{LME}}\) for visualization details see Supplementary Methods: Plotting of regional mean BOLD-responses).

### Multivariate pattern: pattern component modelling of fMRI and EEG data

To assess whether spatial and/or decisional uncertainty models can explain the fine-scale fMRI or EEG activity patterns across the 7 (sound location) × 3 (pre, post-VA, post-AV) = 21 conditions, we combined Pattern Component Modelling (PCM) [32, https://github.com/jedidriksen/pcm_toolbox] and Bayesian model comparison. Like the more widely used representational similarity analyses, pattern component modelling allows one to investigate whether specific representational structures—as implied for instance by the spatial or decisional uncertainty model—are expressed in fMRI or EEG activity patterns. Critically, PCM goes beyond standard representational similarity analyses by allowing us to assess arbitrary mixtures of representational components such as a combined spatial + decisional uncertainty model (i.e. with unknown mixture weights, for further information see Supplementary Methods: Multivariate pattern—pattern component modelling of fMRI and EEG data).

Consistent with our LME analysis, we generated second-station matrices (‘pattern components’) as predictors for PCM based on the activations of 360 simulated neurons from the spatial and decisional uncertainty models, respectively.

We compared the following PCM models for fMRI and EEG data:

- Null PCM: all conditions are independent (i.e. the second-moment matrix is the identity matrix).
- Spatial PCM (S): activity patterns generated by the spatial model without recalibration.
- Decisional PCM (D): activity patterns generated by the decisional uncertainty model without recalibration.

### Combined (spatial + decisional) PCM: activity patterns are a weighted linear combination of the patterns generated by the spatial and the decisional uncertainty model

We factiorially manipulated whether the spatial and/or decisional uncertainty model accommodates audiovisual recalibration:

- (S + D) spatial component without recalibration and decisional component without recalibration.
- (S + D) spatial component with recalibration and decisional component without recalibration.
- (S + D) spatial component without recalibration and decisional component with recalibration.
- (S + D) spatial component with recalibration and decisional component with recalibration.

### Free (i.e. fully flexible) PCM imposes no constraints on the second-moment matrix and provides an upper benchmark

If a model performs at least as good as the cross-validated free PCM, it is sufficiently complex to capture all consistent variations in the data [33].

**Multivariate Pattern: Pattern Component Modelling of fMRI and EEG Data**

The marginal likelihood for each model and subject from the leave-one-subject-out cross-validation scheme was used as an approximation to the model evidence. We compared the models using the natural logarithm (\(\log\)) of the Bayes factors averaged across participants [32].

**Evidential categories for Log-Bayes factors**

Throughout the manuscript we report model comparison results using Bayes factors (BF), specifically the natural logarithm of the Bayes factors (\(\log BF\)) which are a convenient way of expressing evidence in favour of a given model with respect to a reference model. Although BFs and Log-BFs are defined on a continuous scale, it is useful to subdivide this continuous scale into discrete evidential categories [39].

In our manuscript we followed the heuristic classification scheme described by [39] first proposed by [37] expressing the level of evidence in favour of the tested model with respect to the null model. Anecdotal evidence: BF of 1–3 (Log BF of 0–1); Moderate evidence: BF of 3–10 (Log BF of 1.1–2.3); Strong evidence: BF of >10 (Log BF >2.3).

**Reporting summary**

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

The processed data files necessary to reproduce the results using the shared analysis code are available at figshare (https://doi.org/10.6084/m9.figshare.19469861.v2) [38]. The raw data are available for research purposes only upon request from the corresponding author (M.A. mate.aller@nrc-cbu.cam.ac.uk), because of constraints imposed by the ethics approval under which this study was conducted. Source data are provided with this paper.

**Code availability**

The analysis code for this manuscript is available at GitHub (https://github.com/allenmat/audiovisual_adaptation_fMRI-EEG) [39].

Received: 24 March 2021; Accepted: 21 June 2022;

**References**

1. Chen, L. & Vroomen, J. Intersensory binding across space and time: a tutorial review. Atten. Percept. Psychophys. 75, 790–811 (2013).
2. Grothe, B., Pecka, M. & McAlpine, D. Mechanisms of sound localization in mammals. Physiol. Rev. 90, 983–1012 (2010).
3. Konno, N., Lin, I.-F., Shin-Cunningham, B. G. & Groh, J. M. Reference frame of the ventriloquism aftereffect. J. Neurosci. 29, 13809–13814 (2009).
4. Maier, J. X. & Groh, J. M. Multisensory guidance of orienting behavior. Hear. Res. 258, 106–112 (2009).
5. Wandell, B. A., Dumoulin, S. O. & Brewer, A. A. Visual field maps in human cortex. Neuron 56, 366–383 (2007).
6. McAlpine, D., Jiang, D. & Palmer, A. A. Neural code for low-frequency sound localization in mammals. Nat. Neurosci. 4, 396–401 (2001).
7. Ortiz-Rios, M. et al. Widespread and opponent fMRI signals represent sound location in macaque auditory cortex. Neuron 93, 971–983 (2017).
8. Salminen, N. H., May, P. J. C., Alku, P. & Titiminen, H. A population rate code of auditory space in the human cortex. PLoS ONE 4, e7600 (2009).
9. Stecker, G. C., Harrington, J. A. & Middlebrooks, J. C. Location coding by opponent neural populations in the auditory cortex. PLoS Biol. 3, 0520–0528 (2005).
10. Schlack, A., Sterbing-D’Angelo, S. J., Hartung, K., Hoffmann, K.-P. & BREMMER, F. Multisensory space representations in the macaque ventral intraparietal area. J. Neurosci. 35, 4616–4625 (2015).
11. Bertelson, P., Frissen, L., Vroomen, J. & Gelder, B. The aftereffects of ventriloquism: patterns of spatial generalization. Percept. Psychophys. 68, 428–436 (2006).
12. Bosen, A. K., Fleming, J. T., Allen, P. D., O’Neill, W. E. & Paige, G. D. Multiple time scales of the ventriloquism aftereffect. PLoS ONE 13, e0200930 (2018).
13. Radeau, M. & Bertelson, P. The after-effects of ventriloquism. Q. J. Exp. Psychol. 26, 63–71 (1974).
14. Recanzone, G. H. Rapidly induced auditory plasticity: the ventriloquism aftereffect. Proc. Natl Acad. Sci. USA 95, 869–875 (1998).
17. Woods, T. M. & Recanzone, G. H. Visually induced plasticity of auditory spatial perception in macaques. *Cereb. Biol. 14*, 1539–1564 (2004).
18. Wozny, D. R. & Shuman, L. Recalibration of auditory space following milliseconds of cross-modal discrepancy. *J. Neurosci. 31*, 4607–4612 (2011).
19. Zierul, B., Röder, B., Tempelmann, C., Bruns, P. & Noesselt, T. The role of auditory cortex in the spatial ventriloquism aftereffect. *NeuroImage 162*, 257–268 (2017).
20. Zink, T. & Kayer, C. Shared neural underpinnings of multisensory integration and trial-by-trial perceptual recalibration in humans. *eLife* 8, e47001 (2019).
21. Zwislocki, M. P., Van Opstal, A. J. & Paige, G. D. Plasticity in human sound localization induced by compressed spatial vision. *Nat. Neurosci. 6*, 175–181 (2003).
22. Mullett-Gillman, O. A., Cohen, Y. E. & Grob, J. M. Eye-centered, head-centered, and complex coding of visual and auditory targets in the intraparietal sulcus. *J. Neurophysiol. 94*, 2331–2352 (2005).
23. Michalka, S. W., Rosen, M. L., Kong, L., Shinn-Cunningham, B. G. & Somers, D. C. Auditory spatial coding in the intraparietal sulcus. *Cereb. Cortex* 26, 1302–1308 (2016).
24. Rauschecker, J. P. & Tian, B. Mechanisms and streams for processing of ‘what’ and ‘where’ in auditory cortex. *Proc. Natl Acad. Sci. USA 97*, 11800–11806 (2000).
25. van der Heijden, K., Rauschecker, J. P., de Gelder, B. & Formisano, E. Cortical mechanisms of spatial hearing. *Nat. Rev. Neurosci. 20*, 609–623 (2019).
26. Zatorre, R. J., Bouffard, M., Ahad, P. & Belin, P. Where is ‘where’ in the human auditory cortex? *Nat. Rev. Neurosci. 3*, 905–909 (2002).
27. Zhang, C.-G. & Winer, J. The LBISVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol. 2*, 1–27 (2011).
28. Akaike, H. A. New look at the Statistical Model Identification. In *Selected Papers of Hirotugu Akaike* (eds Parzen, E., Tanabe, K. & Kitagawa, G.) 215–222 (Springer New York, 1998).
29. Schönbrodt, F. D. & Wagenmakers, E.-J. Bayes factor design analysis: planning for coming evidence. *Psychon. Bull. Rev. 25*, 128–142 (2018).
30. Nili, H. et al. A toolbox for representational similarity analysis. *PLoS Comput. Biol. 10*, e1003553 (2014).
31. Cox, M. A. A. & Cox, T. F. Multidimensional Scaling. In *Handbook of Data Visualization* (eds. C. Harris, W. & Unwin, A.) 315–347 (Springer, 2008).
32. Diedrichsen, J., Yokoi, A. & Arbib, S. A. Pattern component modeling: a flexible approach for understanding the representational structure of brain activity patterns. *NeuroImage 180*, 119–133 (2018).
33. Friston, K. J., Diedrichsen, J., Holmes, E. & Zeidman, P. Variational representational similarity analysis. *NeuroImage 201*, 115986 (2019).
34. Middlebrooks, J. C., Clock, A. E., Xu, L. & Green, D. M. A panoramic code for sound location by cortical neurons. *Science 264*, 842–844 (1994).
35. Grinband, J., Hirsch, J. & Ferrera, V. P. A neural representation of categorization uncertainty in the human brain. *Neuron 49*, 757–763 (2006).
36. Cichy, R. M. & Oliva, A. A M/EEG-fMRI fusion primer: resolving human multisensory brain responses in space and time. *Neuron* 80, 1544–1557 (2013).
37. Perrin, F., Pernier, J., Bertrand, O. & Echallier, J. F. Spherical splines for scalp MEG-data. *Cereb. Cortex N. Y. N. 25*, 4493–4503 (2011).
38. Zylberberg, A., Fetsch, C. R. & Shadlen, M. N. The influence of evidence volatility on choice, reaction time and confidence in a perceptual decision. *eLife* 5, e17688 (2016).
39. Perrin, F., Bertrand, J. B. & Echallier, J. F. Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurophysiol. 72*, 184–187 (1989).
40. Wozny, D. R. & Shams, L. Recalibration of auditory space following multisensory recalibration. *Science 323*, 177–180 (2009).
41. Zaidel, A., Ma, W. J. & Angelaki, D. E. Supervised calibration relies on the recent history. *J. Neurosci. 31*, 8513 (2011).
42. Zyzner, A., Aertsen, A. M. & Fritsch, C. A. R. & Shadlen, M. N. The influence of evidence volatility on choice, reaction time and confidence in a perceptual decision. *eLife* 5, e17688 (2016).
43. Kiani, R., Cortelli, L. & Shadlen, M. N. A choice certainty is informed by both evidence and decision time. *Neuron 84*, 1329–1342 (2014).
44. Fleming, S. M. & Dew, N. D. Self-evaluation of decision-making: a general Bayesian framework for metacognitive computation. *Psychol. Rev. 124*, 91–114 (2017).
45. Schulz, L., Fleming, S. M. & Dayan, P. Metacognitive computations for information search: confidence in decision. *bioRxiv* https://doi.org/10.1101/2021.03.01.433342 (2021).
46. Bishop, C. M. *Pattern Recognition and Machine Learning* (Information Science and Statistics). (Springer-Verlag, 2006).
47. Jeffreys, H. *Theory of probability*. (Oxford University Press, 1961).
48. Aller, M., Mihalik, A. & Noppeney, U. Source data for research article *Audiovisual adaptation is expressed in spatial and decisional codes*. *f1ghare* https://doi.org/10.6084/m9.f1ghare.19469861.v2 (2022).
49. Aller, M., Mihalik, A. & Noppeney, U. *Audiovisual adaptation is expressed in spatial and decisional codes, allernat/audiospatial_adaptation_fMRI_EEG_av adaptation 0.1.1*. Zenodo https://doi.org/10.5281/zenodo.6572895 (2022).

**Acknowledgements**

This research was funded by the European Research Council (https://erc.europa.eu/; grant number: ERC-2012-StG_20111109 multisens).
Author contributions
Conceptualization: M.A., A.M. and U.N.; Methodology: M.A. and A.M.; Software: M.A. and A.M.; Formal analysis: M.A., A.M. and U.N.; Investigation: M.A. and A.M.; Resources: U.N.; Data curation: M.A. and A.M.; Writing—original draft: M.A., A.M. and U.N.; Writing—review and editing: M.A., A.M. and U.N.; Visualization: M.A. and A.M.; Supervision: U.N.; Funding acquisition: U.N.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-31549-0.

Correspondence and requests for materials should be addressed to Máté Aller.

Peer review information Nature Communications thanks Biyu He, William Penny, and the other, anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022