Genetic influences on brain representations of natural audiovisual experiences

Satoshi Nishida (s-nishida@nict.go.jp)
National Institute of Information and Communications Technology https://orcid.org/0000-0003-0555-518X

Shunsuke Toyoda
Graduate School of Frontier Biosciences, Osaka University https://orcid.org/0000-0002-5291-7422

Chika Honda
Osaka University

Mikio Watanabe
Osaka University https://orcid.org/0000-0001-5971-9296

Miina Ollikainen
Duke University, Durham, NC, USA https://orcid.org/0000-0003-3661-7400

Eero Vuoksimaa
University of Helsinki

Jaakko Kaprio
University of Helsinki https://orcid.org/0000-0002-3716-2455

Shinji Nishimoto
National Institute of Information and Communications Technology https://orcid.org/0000-0001-8015-340X

Article

Keywords: sensory input, genetics, brain representations, audiovisual signaling

Posted Date: September 24th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-902535/v1

Abstract

Natural sensory inputs in everyday situations induce unique experiences that vary between individuals, even when inputs are identical. This experiential uniqueness stems from the representations of sensory signals in each brain. We investigated whether genetic factors control individual differences in sensory representations in the brain by studying the brain representations of natural audiovisual signals in twin-pairs. We measured the brain response to natural movies in twins using functional magnetic resonance imaging and quantified the genetic influence on the multivoxel-pattern similarity of movie clip representations between each twin. The whole-brain analysis revealed a genetic influence on the multivoxel-pattern similarity in widespread brain regions, which included the occipitotemporal sensory cortices as well as the frontoparietal association cortices and subcortical structures. Our findings suggest that genetic factors exhibit an effect on natural audiovisual signaling by controlling audiovisual representations in the brain.

Introduction

The natural audiovisual inputs that we receive in everyday situations induce subjective perceptual experiences that are highly diverse across individuals. The diversity of perceptual experiences is associated with individual differences in brain representations of audiovisual signals, as revealed in previous studies on the synchronization of brain activity across individuals[1–4]. These studies demonstrated that the inter-subject correlation of functional magnetic resonance imaging (fMRI) signals predicts the inter-subject similarity of experiences during video viewing and story listening. Although individual differences in audiovisual representations are characterized by environmental factors, such as familiarity[5], they may also be affected by genetic factors. However, no study to date investigated whether genetic factors contribute to the neural substrates underlying the formation of brain representations of natural audiovisual signals.

Previous twin studies that used task-based fMRI have shown that genetic factors control a wide range of brain functions, such as vision[6,7], executive function[8–10], mental calculation[11], and language[12,13]. These studies employed strictly controlled parametric tasks using simple abstract sensory stimuli that are traditionally used in psychology research. In contrast, natural audiovisual signaling involves multimodal integration and inter-regional network interactions in the brain, which likely yield more complex neural processing than that produced by traditional tasks[14]. Therefore, whether and how genetic factors control natural audiovisual signaling in the brain remain unclear. Addressing this issue is important to gain a deeper understanding of the genetic basis of adaptive brain functions.

To uncover the genetic contribution to brain representations of natural audiovisual signals, we conducted an fMRI experiment in which twin participants viewed a series of natural movie clips with audio while inside the scanner (Fig. 1). We defined a clip-wise representational dissimilarity matrix (CRDM) that was calculated by the pairwise correlation distances between instantaneous multivoxel activation patterns evoked by each movie clip, which acted as a fingerprint of the audiovisual representation in each brain region of each participant. The CRDM was obtained for each anatomically segmented brain region (Fig. 1a). Then, genetic effects on natural audiovisual representations were examined by comparing the CRDM similarities between pairs of monozygotic (MZ) and dizygotic (DZ) twins (Fig. 1b).

Results

Brain regions involved in natural audiovisual signaling

First, to identify which brain regions are involved in the signaling of natural audiovisual information contained in the movie clips, we evaluated the reproducibility of CRDMs in each brain region. We calculated the Pearson correlation of CRDMs in each region across repetitions of identical movie clip stimulation. If a given region is involved in audiovisual signaling, the Pearson correlation of CRDMs will be significantly higher than zero. The whole-brain analysis revealed that although the Pearson correlation of CRDMs varied between regions (Fig. 2), we observed correlation coefficients higher than zero in all cortical and subcortical regions (Wilcoxon test, $P < 10^{-9}$ after false discovery rate [FDR] correction for multiple comparisons [$q = 0.05$]; Supplementary Data 1). This result indicates that all brain regions were involved in audiovisual signaling during natural movie viewing.

Of all the cortical regions, correlation coefficients were highest in the occipital, superior temporal, and posterior parietal cortices compared with other regions (Fig. 2a), which was unsurprising because these cortical regions are involved in audiovisual sensory processing[15,16]. Moreover, this observation is consistent with previous reports that showed that audiovisual signaling during movie viewing is more reproducible in sensory regions than in high-order regions[17,18]. The correlation coefficients of the subcortical regions appeared to be relatively lower compared with the cortical regions; the highest correlation was observed in the brainstem and bilateral thalamus (Fig. 2b). This observation is also plausible because both the brainstem and thalamus include structures that are involved in audiovisual processing[19,20] (e.g., superior/inferior colliculus and lateral/medial geniculate nucleus, respectively). These results indicate that CRDM is a reliable and appropriate measure to capture whole brain natural audiovisual representations.

Variability of natural audiovisual representations across brain regions

Given that functions differ between brain regions, brain representations involved in natural audiovisual signaling should not be uniform across regions; rather, they should vary from region to region. To examine representational variability across regions, we visualized the global structure of CRDM variation

across regions in a low-dimensional representational space. We calculated Pearson correlation distances of CRDMs between all possible pairs of cortical and subcortical regions. Then, a three-dimensional (3D) representational space was estimated from the correlation distances using uniform manifold approximation and projection[21] (UMAP; see Methods), an unsupervised dimension reduction method.

The 3D representations of CRDMs for each region were visualized in UMAP space (Fig. 3a), where brain regions with more similar CRDM representations are located closer together and indicated by more similar colors. Three groups of CRDM representations existed that were located apart from other representations (representations in yellow, green, and magenta in the bottom right, middle right, and top left of Fig. 3a, respectively). Each group corresponded to audiovisual representations that are associated with a specialized function of the brain.

The cortical mapping of the 3D representations of CRDMs revealed that the three groups were localized in different parts of the cortex (Fig. 3b). The representations shown in yellow were mainly localized in the occipital and posterior temporal cortices, which are involved in visual processing[15]. The representations in green were primarily localized in the superior temporal cortex, which is involved in auditory processing[16]. The representations in magenta were mainly localized in the prefrontal and anterior temporal cortices, which are involved in higher-level cognitive functions[22,23]. Thus, the three CRDM representations appropriately linked with functional differentiation of the cortex. Furthermore, these observations did not depend on the dimension reduction method because similar trends were also observed using a more conventional method (Supplementary Fig. 1). These results indicate that audiovisual representations during natural movie viewing vary across brain regions and that CRDMs can effectively capture the variability of audiovisual representations.

Genetic effects on natural audiovisual signaling

We determined which brain regions exhibit natural audiovisual representations that are affected by genetic factors. The genetic effect was evaluated by comparing the similarity of audiovisual representations of MZ twin-pairs with that of DZ twin-pairs (Fig. 1b). When twin-pair representational similarity in a region was significantly higher in MZ than in DZ pairs, the audiovisual representations in that region were regarded as being affected by genetic factors. To test this, we computed the twin-pair Pearson correlation of CRDMs separately for MZ and DZ twins.

The twin-pair correlation of individual brain regions are biased due to differences in noise levels of fMRI signals across regions. To eliminate such bias, we corrected twin-pair correlations using the noise

ceiling[24], which is the upper-most possible bound of CRDM correlations across individuals when the noise is contained in fMRI data (see Methods). The noise ceiling computed from our fMRI data varied from region to region (Supplementary Fig. 2 and Data 2). Because the twin-pair correction divided by noise ceiling enabled fair comparisons of twin-pair similarity across brain regions, we used the divided correlation as a measure of twin-pair similarity.

Twin-pair similarity of CRDMs in each region was averaged across MZ and DZ pairs, separately (Fig. 4; see Supplementary Fig. 3 for the results without noise-ceiling correction). Similarity was significantly higher than zero in all cortical and subcortical regions (Wilcoxon test, $P < 10^{-9}$, FDR corrected; see Supplementary Data 3 and 4 for details and the results without noise-ceiling correction, respectively). For the cortical regions, pair-average similarity was higher in the occipital, superior temporal, posterior parietal, and dorsolateral prefrontal (dlPFC) cortices compared with other regions, in both MZ and DZ pairs (Fig. 4a and c). This is consistent with previous reports that audiovisual signals in sensory regions exhibit higher similarity across individuals than those in high-order regions[1,25–27]. In the subcortical regions, similarity appeared to be higher in the caudate, thalamus, and brainstem compared with other regions, in both MZ and DZ pairs (Fig. 4b and d). MZ and DZ pairs showed highly similar patterns of mean twin-pair similarity values across brain regions (Pearson correlation, $r = 0.939$, $P < 10^{-75}$). These results indicate that natural audiovisual representations evaluated with CRDMs in individual brain regions are reproducible, even across individuals.

We observed significant CRDM similarity differences between MZ and DZ pairs in widespread cortical and subcortical regions (Wilcoxon test, $P < 0.05$, FDR corrected; Supplementary Data 5). In cortical regions, significant genetic effects were present in almost all regions of the occipital, temporal, parietal, and ventral prefrontal cortices. In contrast, regions of the dlPFC, the dorsomedial cortex (dmPFC), and the insular cortex exhibited weak or no significant genetic effects. Of the subcortical structures, significant genetic effects were present in all regions except for the right thalamus, right hippocampus, and left amygdala.

We used Falconer's heritability[28] to evaluate the degree of genetic effects in each cortical and subcortical region for the comparison between regions (see Methods). Of all the cortical regions with significant genetic effects, the occipital and temporal cortices exhibited relatively high heritability while the prefrontal and parietal cortices exhibited relatively low heritability, except for the orbitofrontal cortex (OFC; Fig. 5a and Supplementary Data 5; see Supplementary Fig. 4 and Data 6 for the results without noise-ceiling correction). Among all the subcortical regions with significant genetic effects, the right pallidum exhibited the highest heritability, although subcortical heritability was generally lower than cortical heritability. We observed similar results even when using Holzinger's heritability[29] as an alternative measure of genetic

effects (Supplementary Fig. 5) and even when using intraclass correlation[30] as an alternative measure of twin-pair similarity (Supplementary Fig. 6 and Data 7).

The brain regions with relatively high heritability contained the cortical regions in which susceptibility artifacts during fMRI scanning frequently appear, including the mediotemporal cortex and OFC[31] (Fig. 5a). Indeed, these regions had relatively low noise ceiling, which corresponds to high noise levels (Supplementary Fig. 2). It might be argued that the observed strong genetic effects in these regions were caused by the overestimation of contingent differences between MZ- and DZ-pair similarities due to the correction with low noise ceiling. However, this possibility can be excluded for the following reasons. First, irrelevant to noise-ceiling correction, these regions consistently exhibited significant genetic effects in contrast to the cortical regions with lower noise levels, such as dlPFC and dmPFC. Second, the heritability was not significantly correlated with the noise-ceiling values across brain regions (Supplementary Fig. 7; Pearson correlation, $r = -0.122$, $P = 0.120$). Thus, it is unlikely that the observed cortical pattern of genetic effects is explained solely by the influence of low noise ceiling due to susceptibility artifacts.

Taken together, these results indicate that genetic effects on natural audiovisual representations were present in widespread brain regions, including prefrontal and subcortical regions. In particular, these effects were relatively strong in the occipital, temporal, and orbitofrontal regions.

Discussion

The present study investigated whether individual variability of natural audiovisual representations in the brain is affected by genetic factors. For this purpose, we introduced CRDMs as a measure of natural audiovisual representations in each brain region. We demonstrated that CRDM is a reliable measure of audiovisual representations that is reproducible both within and across individuals. Moreover, we showed that it can effectively capture representational variability across different brain regions. Next, the comparison of twin-pair CRDM similarities between MZ and DZ pairs revealed significant genetic effects on natural audiovisual representations in widespread brain regions. We observed genetic effects not only in the sensory regions but also in the frontoparietal association cortices and subcortical regions. These findings suggest that genetic factors affect natural audiovisual signaling by controlling individual differences of audiovisual representations across the whole brain.

The low-dimensional visualization of CRDMs revealed that the brain possesses a continuous representational space in which different audiovisual representations are smoothly distributed across brain regions (Fig. 3). Specifically, three distinctive groups of brain regions were identified in representational space. One was located in the occipitotemporal visual region, another in the superior

temporal auditory region, and the other in the prefrontal region with the anterior temporal region. It is worth noting that the remaining regions are not necessarily unrelated to natural audiovisual signaling (Fig. 2). Below, we will discuss the genetic effects on natural audiovisual representations based primarily on these distinctive representational groups.

In the occipitotemporal visual regions, we found that genetic effects were high across all regions (Fig. 5). These regions included early visual areas in the occipital cortex and category-selective areas in the posterior inferior temporal cortex. Our results showed that heritability was lower in the category-selective regions than in the early visual areas. Previous studies explored the genetic effects on visual signaling in these regions using strictly controlled tasks[6,7,32]. A magnetoencephalography study revealed that gamma-band synchronization in early visual areas induced by grating stimuli exhibits a high genetic contribution ratio close to 90%[32]. Other studies reported that although fMRI responses in category-selective visual areas, such as the fusiform face and visual word form areas, are controlled by genetic factors, heritability is low[6,7]. Furthermore, it was previously reported that representations in category-selective regions are easily manipulated by environmental factors, such as expertise[33,34]. Therefore, consistent with these findings, our results suggest that although representations in visual areas are strongly affected by genetic factors, these effects are weakened with increased representational complexity.

In the superior temporal auditory region, we found that genetic effects were present but were lower than those observed in the occipitotemporal visual region (Fig. 5a). Moreover, the superior temporal gyrus, which includes the primary auditory cortex[16], exhibited lower heritability than the region around the superior temporal sulcus, which is involved in complex auditory processing, such as speech comprehension[35]. Previous studies reported that the neural basis of speech perception is genetically controlled[7,36]. However, to our knowledge, a neuroimaging study that investigated the genetic contributions to non-speech auditory signaling in the brain does not exist, although several studies reported high heritability of auditory-processing ability at the behavioral level[37,38]. Taken together, genetic effects on non-speech auditory representations may be weak, at least in natural audiovisual signaling.

We found that genetic effects varied across regions in the prefrontal area (Fig. 5a). Within the prefrontal area, genetic effects were observed in the ventrolateral cortex (vlPFC) and OFC. In contrast, the dlPFC and dmPFC exhibited weak or no genetic effects despite their involvement in natural audiovisual signaling (Fig. 2a). Previous studies examined genetic contributions to activation patterns in each of these cortices separately using different behavioral tasks that were strictly controlled for a specific cognitive function, such as semantic processing[39], emotion[40], working memory[9,10], and conflict monitoring[8]. These studies

separately reported genetic contributions in the vlPFC[10,39], dlPFC[9], and dmPFC[8] but not in the OFC[40], which is not consistent with our results. However, because each of these prefrontal regions is involved in different cognitive functions[22], the genetic effects on activation patterns in the same region may depend on the specific cognitive function being investigated. Therefore, our findings suggest that the cognitive functions related to natural audiovisual signaling are genetically controlled in the vlPFC and OFC but not in the dlPFC or dmPFC.

The anterior temporal cortex (ATC) had audiovisual representations that were similar to ones in the prefrontal area (Fig. 3) and also showed relatively strong genetic effects (Fig. 5a). The ATC is considered as an extended part of the limbic system[41], including the OFC, and indeed has a strong connection with the OFC[42,43]. A variety of cognitive functions, such as semantic memory[44,45], social cognition[46], and emotion[47,48], are associated with the ATC. Of them, the involvement in emotional processing is in common with the OFC[49], which also showed relatively strong genetic effects in our results (Fig. 5a). This finding may suggest that the neural substrates of emotional signaling during audiovisual processing are substantially controlled by genetic factors. To our knowledge, no previous study has reported the genetic contribution of audiovisual processing in the ATC. Thus, our findings provide clues to the nature of the functional and genetic roles of the ATC in audiovisual processing.

The low-dimensional visualization of CRDMs revealed that the distributions of the three subcortical regions, the left and right accumbens and the right pallidum, were distinct from that of the other subcortical regions (Fig. 3). Notably, they were contained within the group of prefrontal regions. The accumbens has strong projections from prefrontal regions[50] and is involved in cognitive functions similar to those of the OFC, such as emotion[51] and value-based decision making[52,53]. Therefore, it is plausible that audiovisual representations in the accumbens are analogous to those in prefrontal regions and exhibit genetic effects. However, it was surprising that only the right pallidum exhibited audiovisual representations similar to those in prefrontal regions and the strongest genetic effect among subcortical regions (Fig. 5b). Although the ventral pallidum receives projections from the accumbens[50], no study, to our knowledge, explored the lateralization of these projections or functions of the pallidum. The lateralization was maintained even when a different dimensional reduction method was used (Supplementary Fig. 1). Our findings provide important insights into the neural substrates of audiovisual signaling in the subcortex.

Natural audiovisual stimulation induces complex, integrative neural signaling that involves multimodal integration and inter-regional network interactions[14]. However, evidence suggests that natural audiovisual stimulation rather induces brain signaling that is more reproducible within and across individuals than

that induced by strictly controlled stimulation[54]. This enables efficient investigation of the individual differences in brain functions[55] and how they are associated with pathological abnormalities[56,57]. With the use of natural audiovisual stimulation, we were able to capture reliable audiovisual representations across the brain and comprehensively uncover genetic influences on audiovisual representations. To our knowledge, this is the first study to report genetic effects on natural audiovisual signaling and quantify heritability across brain regions. Our findings provide novel insights into the genetic basis underlying natural audiovisual experiences in everyday situations.

Methods

Participants

Thirty-nine healthy Japanese twin-pairs were recruited from the registry established by the Center for Twin Research at Osaka University[58,59]. Twenty-two pairs were MZ twins (15 female and 7 male pairs; mean age ± standard deviation [SD] = 26.0 ± 4.3 years), and 17 pairs were DZ twins (6 female-female, 4 male-male, and 7 female-male pairs; mean age ± SD = 26.6 ± 5.3 years). All had normal or corrected-to-normal vision. Written informed consent was obtained from all participants. The experimental protocol was approved by the ethics and safety committees of the National Institute of Information and Communications Technology, Osaka University, and University of Yamanashi.

MRI data collection

Functional and anatomical MRI data were collected using a 3T Siemens TIM Trio scanner (Siemens, Germany) with a 32-channel Siemens volume coil. Functional data were collected using a multiband gradient echo EPI sequence[60] (repetition time [TR] = 2,000 ms; echo time [TE] = 30 ms; flip angle = 62°; voxel size = 2 × 2 × 2 mm; matrix size = 96 × 96; field of view [FOV] = 192 × 192 mm; number of slices = 72; multiband factor = 3). Anatomical data were collected using a T1-weighted MPRAGE sequence (TR = 2,530 ms; TE = 3.26 ms; flip angle = 9°; voxel size = 1 × 1 × 1 mm; matrix size = 256 × 256; FOV = 256 × 256 mm; number of slices = 208).

Experimental design

The audiovisual stimuli consisted of natural color movies obtained from Vimeo (https://vimeo.com). Movies were clipped to 10–20 s in length, and the movie clips were sequentially concatenated in a pseudo-random order. This procedure produced seven sets of 610-s movies. Each set of movies were displayed during separate scans. The initial 10 s of each scan were dummy scans to avoid hemodynamic transients caused by movie onset and were not used for analysis. Six of the seven movie sets consisted of 225 unique movie clips (1 h in total). The remaining one set consisted of four repetitions of 11 unique

movie clips (150 s in total) for testing the reproducibility of brain representations (see Representational reproducibility). The 236 unique movie clips included a wide variety of movie categories (Supplementary Table 1). Participants viewed the movies displayed on a projector screen (23.8 × 13.5 degrees of visual angle at 24 Hz) and listened to the audio through MR-compatible headphones while inside the scanner. Participants were not asked to perform any explicit task.

fMRI data preprocessing

For the functional data, motion correction was performed using the statistical parameter mapping toolbox (SPM8, http://www.fil.ion.ucl.ac.uk/spm/software/ spm8/). For each subject, all volumes were aligned to the first image of the first functional run. Low-frequency fMRI response drift was eliminated by subtracting the median-filtered signals (with a 120-s window) from the raw signals. Then, the response for each voxel was normalized by subtracting the mean response and scaling to the unit variance. FreeSurfer[61,62] was used to identify cortical surfaces and subcortical structures from anatomical data, which were then registered to the voxels of functional data. Each voxel was assigned to a cortical or subcortical region based on cortical[63] and subcortical[64] segmentation atlases.

CRDM

The structure of the natural audiovisual representations in each cortical or subcortical region was evaluated using CRDMs, which were calculated using the following procedure. First, the timeseries of fMRI responses of each participant were separated into time windows of each clip after temporally shifting with a delay of 4 s to correct for the hemodynamic response. For each time window of each clip, correlation distances (1–Pearson correlation coefficient) of the multivoxel response patterns in each region were calculated for all possible pairs of 2-s fMRI time samples, which yielded a CRDM for that clip. This procedure was repeated for all clips and in all regions, which produced a set of CRDMs for each participant. This method is a form of representational similarity analysis[65,66], which allows comparisons of different brain representations between individuals.

Representational reproducibility

For the set of movies that comprised four repetitions of 11 movie clips (see Experimental design), we evaluated the reproducibility of natural audiovisual representations in each brain region using the Pearson correlation of the upper triangular section of the CRDMs for each pair of repetitions. Specifically, Pearson correlation coefficients were calculated for each clip for all four repetitions and then averaged. Then, the mean coefficients were averaged across the 11 clips, which produced one coefficient for each

region in each participant as a measure of representational reproducibility. If the coefficient was larger in any given region, this indicated that reproducibility was higher in that region.

Representational variability across regions

The variability of natural audiovisual representations across brain regions was examined using the Pearson correlation distances of the upper triangular section of the CRDMs for all possible pairs of brain regions. The correlation distances were calculated for each movie clip and then averaged across all clips. Based on the average correlation distances, a low-dimensional representational space was estimated using UMAP[21] to visualize the global structure of CRDM variability across regions. UMAP is a recently developed nonlinear dimension reduction method. This method can effectively estimate low-dimensional representations of data samples from their pairwise distances while preserving the global structure of the data samples. In this study, the learning of UMAP was implemented using the official python library (https://github.com/lmcinnes/umap) with the following parameters: number of components = 3, size of local neighborhood = 15, and minimum distance of embedded points = 0.1. Accordingly, we obtained 3D representations of CRDMs. To verify the consistency of visualization results across methods, a more conventional nonlinear dimension reduction method called isomap[67] was also used to estimate 3D representations of CRDMs using the following parameters: number of components = 3 and number of nearest neighbors = 15.

Twin-pair representational similarity

The twin-pair similarity of natural audiovisual representations was evaluated using the Pearson correlation of the upper triangular section of the CRDMs for each clip within each twin-pair. Then, Pearson correlation coefficients of the CRDMs for each clip were averaged across MZ or DZ pairs, which produced a coefficient for each of the 236 movie clips for the MZ and DZ pairs separately. The CRDMs for the 11 clips that were repeatedly presented four times were obtained from the fMRI signals averaged across the repetitions. The mean coefficients for the MZ and DZ pairs were used as a measure of twin-pair representational similarity.

Conventional twin studies have employed intraclass correlation[30], instead of standard Pearson correlation, to assess twin-pair similarity[11,68]. Hence, the present study also calculated the intraclass correlation of CRDMs within each twin-pair using the double-entry method[30] as an alternative measure of the twin-pair similarity of natural audiovisual representations. However, although the results were similar between standard and intraclass correlations, our main conclusions were derived from the results with standard correlation but not those with intraclass correlation for the following two reasons. First,

intraclass correlation has been conventionally applied to univariate data[11,68] but not multivariate data as used in this study (i.e., CRDMs). Second, the noise-ceiling correction of twin-pair similarity, which is important to compare the similarity fairy across brain regions (see the next section), was originally introduced for standard correlation[24], and it is not straightforward that the correction can be extended to intraclass correlation. Thus, intraclass correlation was used for a supplementary analysis in this study.

Correction using the noise ceiling

Measurement noise contained in fMRI signals vary across brain regions that can be due to physical factors, such as the position-dependent sensitivity change of a receiver coil[69]. This affects the regional difference of representational similarity between individuals and prevents fair comparisons of similarity across brain regions. To solve this issue, similarity was corrected using the noise ceiling[24], which is the upper-most possible bound of similarity in each region when position-dependent measurement noise is contained in fMRI signals. The noise ceiling was obtained from the Pearson (or intraclass) correlation between single-participant and participant-average CRDMs that was computed after each participant's CRDM was z-scored[24]. The correlation coefficients were calculated for each single-participant CRDM of each clip and averaged across all clips and all participants. This procedure yielded one noise-ceiling value for each brain region. The MZ- and DZ-pair CRDM similarity for each clip in each brain region was divided by the corresponding noise-ceiling value of that region.

Genetic effects

Because twins from MZ pairs have the same genomic sequence, while twins from DZ pairs share, on average, 50% of segregating genes, the similarity differences between these pairs are considered to reflect the genetic effects on natural audiovisual representations. The environmental exposures shared by twins are assumed to be similar in MZ and DZ twins[70,71]. Accordingly, the genetic effect in each brain region was examined by comparing between MZ- and DZ-pair CRDM similarities, corrected with the noise ceiling. A Wilcoxon sign-rank test was used to statistically test the difference between the sets of MZ- and DZ-pair CRDM similarities for the 236 clips in each brain region. The *p* values were corrected using FDR for multiple comparisons (q = 0.05). When MZ-pair CRDM similarity in a region was significantly higher than DZ-pair one, the audiovisual representations in that region were regarded as having genetic effects.

The degree of genetic effect on audiovisual representations in each region was quantified using Falconer heritability[28], given by $2 \times (R_{MZ} - R_{DZ})$, and Holzinger's heritability[29], given by $(R_{MZ} - R_{DZ}) / (1 - R_{DZ})$. R_{MZ} and R_{DZ} are the clip-averaged CRDM similarities (Pearson or intraclass correlation coefficients) for MZ and DZ pairs, respectively. Previous twin studies have typically calculated the same types of heritability

from twin-pair similarity evaluated with univariate intraclass correlation[7,11,68]. In contrast, this study measured twin-pair similarity using Pearson or intraclass correlation calculated from multivariate CRDMs. Therefore, comparing heritability results of our study with those of other studies seems unfair and is not recommended. We used heritability only for our regional comparisons of genetic effects on audiovisual representations.

Data availability

The source data of the figures are provided with this paper and the data set generated during the current study is publicly available in the GitHub repository (https://github.com/***).

Code availability

The code implementing the analyses in the current study is publicly available in the GitHub repository (https://github.com/***).

Declarations

Acknowledgments: We thank all the twins who participated in this study. We also thank members of our groups for their support and helpful discussions. The work was supported by JSPS KAKENHI Grant-in-Aid for Early-Career Scientists (18K18141 to S.Nishida; 19K16195 to S.T.), for Challenging Exploratory Research (26640012 to S.T.), for JSPS Fellows (15J02806 to S.T.), and for Young Scientists A (15H05311 to S.Nishimoto); Grant-in-Aid for Scientific Research on Innovative Areas (Integrative research toward elucidation of generative brain systems for individuality) (17H05944), and JST ERATO (JPMJER1801) from MEXT, the Brain Sciences Project of the Center for Novel Science Initiatives (CNSI), National Institutes of Natural Sciences (NINS) (BS281003), the Kowa Life Science Foundation, and the Takeda Science Foundation. S.T. was supported by postdoctoral fellowships from JSPS, the TOYOBO Biotechnology Foundation, and the Uehara Memorial Foundation. The Osaka University Center for Twin Research was supported by grants from the Ministry of Education, Culture, Sports, Science, and Technology.

Author contributions: S. Nishida, S.T., C.H., M.W., and S. Nishimoto designed the research. S. Nishida. and S.T. performed the experiments. S. Nishida performed the analysis. S. Nishida, S.T., C.H., M.W., N.O., E.V., J.K., and S. Nishimoto wrote the manuscript.

Competing Interest: The authors declare no competing interest.

References

1. Nummenmaa, L. *et al.* Emotions promote social interaction by synchronizing brain activity across individuals. *Proc. Natl. Acad. Sci.* **109**, 9599–9604 (2012).

2. Jääskeläinen, I. P. *et al.* Brain hemodynamic activity during viewing and re-viewing of comedy movies explained by experienced humor. *Sci. Rep.* **6**, 27741 (2016).

3. Nguyen, M., Vanderwal, T. & Hasson, U. Shared understanding of narratives is correlated with shared neural responses. *Neuroimage* **184**, 161–170 (2019).

4. Saalasti, S. *et al.* Inferior parietal lobule and early visual areas support elicitation of individualized meanings during narrative listening. *Brain Behav.* **9**, e01288 (2019).

5. Charest, I., Kievit, R. A., Schmitz, T. W., Deca, D. & Kriegeskorte, N. Unique semantic space in the brain of each beholder predicts perceived similarity. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 14545–14570 (2014).

6. Polk, T. A., Park, J., Smith, M. R. & Park, D. C. Nature versus nurture in ventral visual cortex: A functional magnetic resonance imaging study of twins. *J. Neurosci.* **27**, 13921–13925 (2007).

7. Pinel, P. *et al.* Genetic and environmental influences on the visual word form and fusiform face areas. *Cereb. Cortex* **25**, 2478–2493 (2015).

8. Matthews, S. C. *et al.* Heritability of anterior cingulate response to conflict: An fMRI study in female twins. *Neuroimage* **38**, 223–227 (2007).

9. Blokland, G. A. M. *et al.* Quantifying the heritability of task-related brain activation and performance during the N-back working memory task: A twin fMRI study. *Biol. Psychol.* **79**, 70–79 (2008).

10. Koten, J. W. *et al.* Genetic contribution to variation in cognitive function: an fMRI study in twins. *Science* **323**, 1737–1740 (2009).

11. Pinel, P. & Dehaene, S. Genetic and environmental contributions to brain activation during calculation. *Neuroimage* **81**, 306–316 (2013).

12. Sommer, I. E. C., Ramsey, N. F., Mandl, R. C. W. & Kahn, R. S. Language lateralization in monozygotic twin pairs concordant and discordant for handedness. *Brain* **125**, 2710–2718 (2002).

13. Sakai, K. L., Miura, K., Narafu, N. & Muraishi, Y. Correlated functional changes of the prefrontal cortex in twins induced by classroom education of second language. *Cereb. Cortex* **14**, 1233–1239

(2004).

14. Sonkusare, S., Breakspear, M. & Guo, C. Naturalistic stimuli in neuroscience: Critically acclaimed. *Trends Cogn. Sci.* **23**, 699–714 (2019).

15. Grill-Spector, K. & Malach, R. The human visual cortex. *Annu. Rev. Neurosci.* **27**, 649–677 (2004).

16. Brewer, A. A. & Barton, B. Maps of the auditory cortex. *Annu. Rev. Neurosci.* **39**, 385–407 (2016).

17. Golland, Y. *et al.* Extrinsic and intrinsic systems in the posterior cortex of the human brain revealed during natural sensory stimulation. *Cereb. Cortex* **17**, 766–777 (2007).

18. Hasson, U., Yang, E., Vallines, I., Heeger, D. J. & Rubin, N. A hierarchy of temporal receptive windows in human cortex. *J. Neurosci.* **28**, 2539–2550 (2008).

19. Paxinos, G. & Huang, X.-F. *Atlas of the human brainstem*. (Elsevier, 2013).

20. Jones, E. G. *The thalamus*. (Springer Science & Business Media, 2012).

21. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv* 1802.03426 (2018).

22. Fuster, J. *The prefrontal cortex*. (Academic Press, 2015).

23. Olson, I. R., Plotzker, A. & Ezzyat, Y. The Enigmatic temporal pole: a review of findings on social and emotional processing. *Brain* **130**, 1718–1731 (2007).

24. Nili, H. *et al.* A toolbox for representational similarity analysis. *PLoS Comput. Biol.* **10**, e1003553 (2014).

25. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634–1640 (2004).

26. Hasson, U., Furman, O., Clark, D., Dudai, Y. & Davachi, L. Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. *Neuron* **57**, 452–462 (2008).

27. Naci, L., Cusack, R., Anello, M. & Owen, A. M. A common neural code for similar conscious experiences in different individuals. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 14277–14282 (2014).

28. Falconer, D. S. *Introduction to quantitative genetics*. (Oliver and Boyd, 1960).

29. Holzinger, K. J. The relative effect of nature and nurture influences on twin differences. *J. Educ. Psychol.* **20**, 241–248 (1929).

30. Furr, R. M. The double-entry intraclass correlation as an index of profile similarity: meaning, limitations, and alternatives. *J. Pers. Assess.* **92**, 1–15 (2010).

31. Ojemann, J. G. *et al.* Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. *Neuroimage* **6**, 156–167 (1997).

32. van Pelt, S., Boomsma, D. I. & Fries, P. Magnetoencephalography in twins reveals a strong genetic determination of the peak frequency of visually induced gamma-band synchronization. *J. Neurosci.* **32**, 3388–3392 (2012).

33. Bilalić, M., Grottenthaler, T., Nägele, T. & Lindig, T. The faces in radiological images: Fusiform face area supports radiological expertise. *Cereb. Cortex* **26**, 1004–1014 (2016).

34. Martens, F., Bulthé, J., van Vliet, C. & Op de Beeck, H. Domain-general and domain-specific neural changes underlying visual expertise. *Neuroimage* **169**, 80–93 (2018).

35. Blomert, L. The neural signature of orthographic-phonological binding in successful and failing reading development. *Neuroimage* **57**, 695–703 (2011).

36. Babajani-Feremi, A. Neural mechanism underling comprehension of narrative speech and its heritability: Study in a large population. *Brain Topogr.* **30**, 592–609 (2017).

37. Morell, R. J. *et al.* A twin study of auditory processing indicates that dichotic listening ability is a strongly heritable trait. *Hum. Genet.* **122**, 103–111 (2007).

38. Brewer, C. C. *et al.* Heritability of non-speech auditory processing skills. *Eur. J. Hum. Genet.* **24**, 1137–1144 (2016).

39. Araki, T. *et al.* Language-related cerebral oscillatory changes are influenced equally by genetic and environmental factors. *Neuroimage* **142**, 241–247 (2016).

40. Côté, C. *et al.* Individual variation in neural correlates of sadness in children: A twin fMRI study. *Hum. Brain Mapp.* **28**, 482–487 (2007).

41. Mesulam, M. M. Paralimbic (mesocortical) areas. in *Principles of behavioral and cognitive neurology* 49–54 (Oxford University Press New York, 2000).

42. Kondo, H., Saleem, K. S. & Price, J. L. Differential connections of the temporal pole with the orbital and medial prefrontal networks in macaque monkeys. *J. Comp. Neurol.* **465**, 499–523 (2003).

43. Kondo, H., Saleem, K. S. & Price, J. L. Differential connections of the perirhinal and parahippocampal cortex with the orbital and medial prefrontal networks in macaque monkeys. *J. Comp. Neurol.* **493**, 479–509 (2005).

44. Correia, J. *et al.* Brain-based translation: fMRI decoding of spoken words in bilinguals reveals language-independent semantic representations in anterior temporal lobe. *J. Neurosci.* **34**, 332–338 (2014).

45. Malone, P. S., Glezer, L. S., Kim, J., Jiang, X. & Riesenhuber, M. Multivariate pattern analysis reveals category-related organization of semantic representations in anterior temporal cortex. *J. Neurosci.* **36**, 10089–10096 (2016).

46. Reniers, R. L. E. P., Völlm, B. A., Elliott, R. & Corcoran, R. Empathy, ToM, and self–other differentiation: An fMRI study of internal states. *Soc. Neurosci.* **9**, 50–62 (2014).

47. Dolan, R. ., Lane, R., Chua, P. & Fletcher, P. Dissociable temporal lobe activations during emotional episodic memory retrieval. *Neuroimage* **11**, 203–209 (2000).

48. Hsieh, S., Hornberger, M., Piguet, O. & Hodges, J. R. Brain correlates of musical and facial emotion recognition: Evidence from the dementias. *Neuropsychologia* **50**, 1814–1822 (2012).

49. Pessoa, L. On the relationship between emotion and cognition. *Nat. Rev. Neurosci.* **9**, 148–58 (2008).

50. Floresco, S. B. The nucleus accumbens: an interface between cognition, emotion, and action. *Annu. Rev. Psychol.* **66**, 25–52 (2015).

51. Cardinal, R. N., Parkinson, J. A., Hall, J. & Everitt, B. J. Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci. Biobehav. Rev.* **26**, 321–352 (2002).

52. Patel, S. R. *et al.* Single-neuron responses in the human nucleus accumbens during a financial decision-making task. *J. Neurosci.* **32**, 7311–7315 (2012).

53. Knutson, B., Rick, S., Wimmer, G. E., Prelec, D. & Loewenstein, G. Neural predictors of purchases. *Neuron* **53**, 147–156 (2007).

54. Hasson, U., Malach, R. & Heeger, D. J. Reliability of cortical activity during natural stimulation. *Trends Cogn. Sci.* **14**, 40–48 (2010).

55. Finn, E. S. *et al.* Idiosynchrony: From shared responses to individual differences during naturalistic neuroimaging. *Neuroimage* **215**, 116828 (2020).

56. Eickhoff, S. B., Milham, M. & Vanderwal, T. Towards clinical applications of movie fMRI. *Neuroimage* **217**, 116860 (2020).

57. Nishida, S. *et al.* Reduced intra- and inter-individual diversity of semantic representations in the brains of schizophrenia patients. *bioRxiv* 2020.06.03.132928 (2020) doi:10.1101/2020.06.03.132928.

58. Hayakawa, K. & Iwatani, Y. An overview of multidisciplinary research resources at the Osaka University Center for Twin Research. *Twin Res. Hum. Genet.* **16**, 217–220 (2013).

59. Honda, C., Watanabe, M., Tomizawa, R. & Sakai, N. Update on Osaka University Twin Registry: An Overview of Multidisciplinary Research Resources and Biobank at Osaka University Center for Twin

Research. *Twin Res. Hum. Genet.* **22**, 597–601 (2019).

60. Moeller, S. *et al.* Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn. Reson. Med.* **63**, 1144–1153 (2010).

61. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* **9**, 179–194 (1999).

62. Fischl, B., Sereno, M. I. & Dale, A. M. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage* **9**, 195–207 (1999).

63. Destrieux, C., Fischl, B., Dale, A. & Halgren, E. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage* **53**, 1–15 (2010).

64. Fischl, B. *et al.* Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron* **33**, 341–355 (2002).

65. Kriegeskorte, N., Mur, M. & Bandettini, P. Representational similarity analysis - connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2**, 4 (2008).

66. Kriegeskorte, N. & Kievit, R. A. Representational geometry: Integrating cognition, computation, and the brain. *Trends Cogn. Sci.* **17**, 401–412 (2013).

67. Tenenbaum, J. B., de Silva, V. & Langford, J. C. A global geometric framework for nonlinear dimensionality reduction. *Science* **290**, 2319–2323 (2000).

68. Bartley, A. J., Jones, D. W. & Weinberger, D. R. Genetic variability of human brain size and cortical gyral patterns. *Brain* **120**, 257–269 (1997).

69. Yang, Y. *et al.* Simultaneous perfusion and BOLD imaging using reverse spiral scanning at 3T: Characterization of functional contrast and susceptibility artifacts. *Magn. Reson. Med.* **48**, 278–289 (2002).

70. Boomsma, D., Busjahn, A. & Peltonen, L. Classical twin studies and beyond. *Nat. Rev. Genet.* **3**, 872–882 (2002).

71. Polderman, T. J. C. *et al.* Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* **47**, 702–709 (2015).

Figures

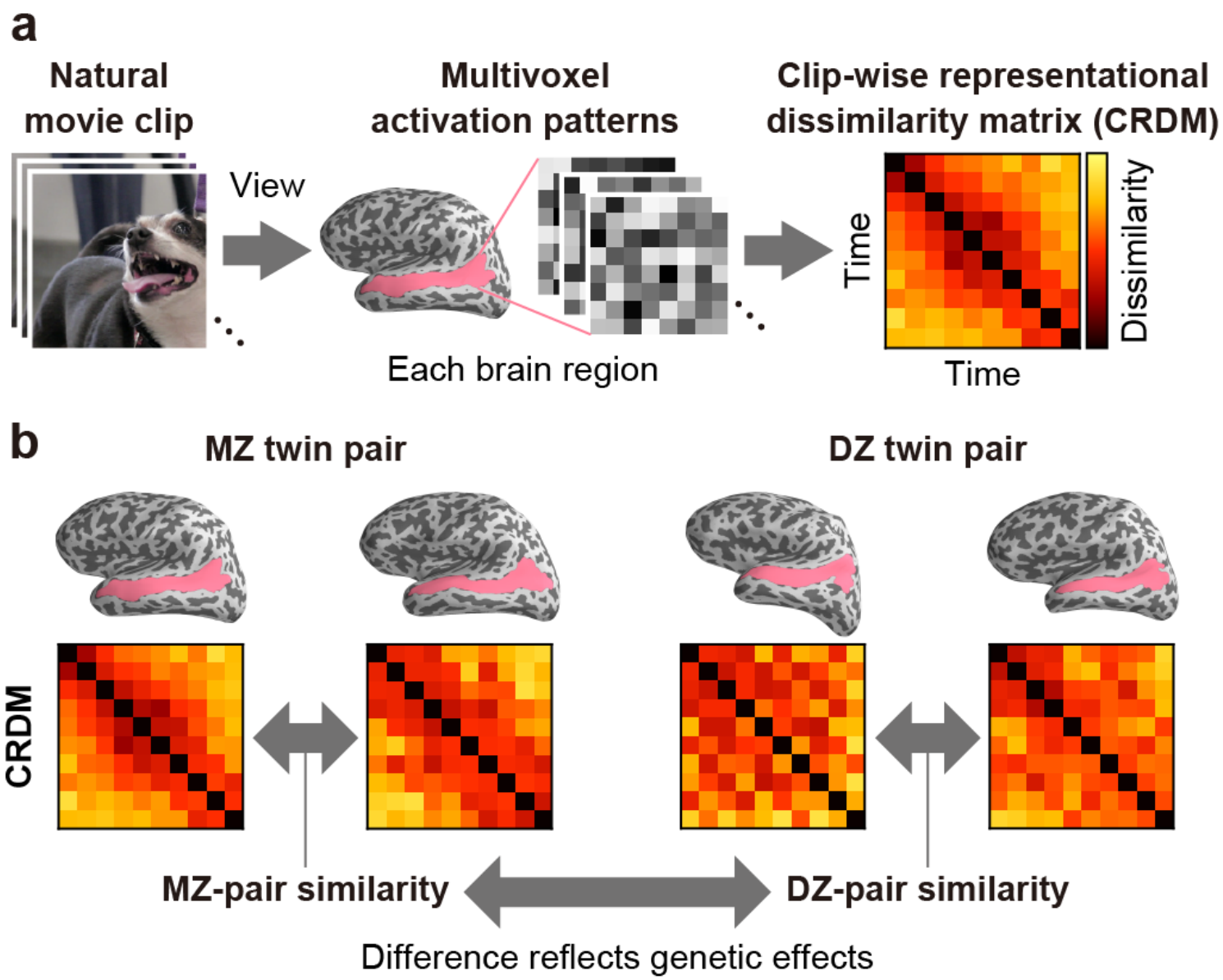

Figure 1

Experimental design. a, Characterization of natural audiovisual representations in the brain. fMRI signals were measured in each participant while they viewed a series of movie clips (each clip was 10–20 s in length, a total of 236 clips) that contained a variety of natural audiovisual information. Multivoxel activation patterns within each anatomically segmented brain region were used to obtain CRDMs. CRDMs were computed from the correlation distance between all possible time-sample pairs of the multivoxel activation patterns from each clip and were regarded as the character of audiovisual representations in each brain region. b, Testing for genetic effects on natural audiovisual representations. Twin-pair similarity of the CRDMs was calculated separately for MZ and DZ twin-pairs. The similarity of CRDMs between MZ and DZ pairs was compared to determine the genetic effects on audiovisual representations, given the greater genetic similarity of MZ pairs (identical genomic sequence) compared to DZ pairs (50% of segregating genes in common). The relevant environmental exposures are assumed to be equally shared in MZ and DZ pairs.

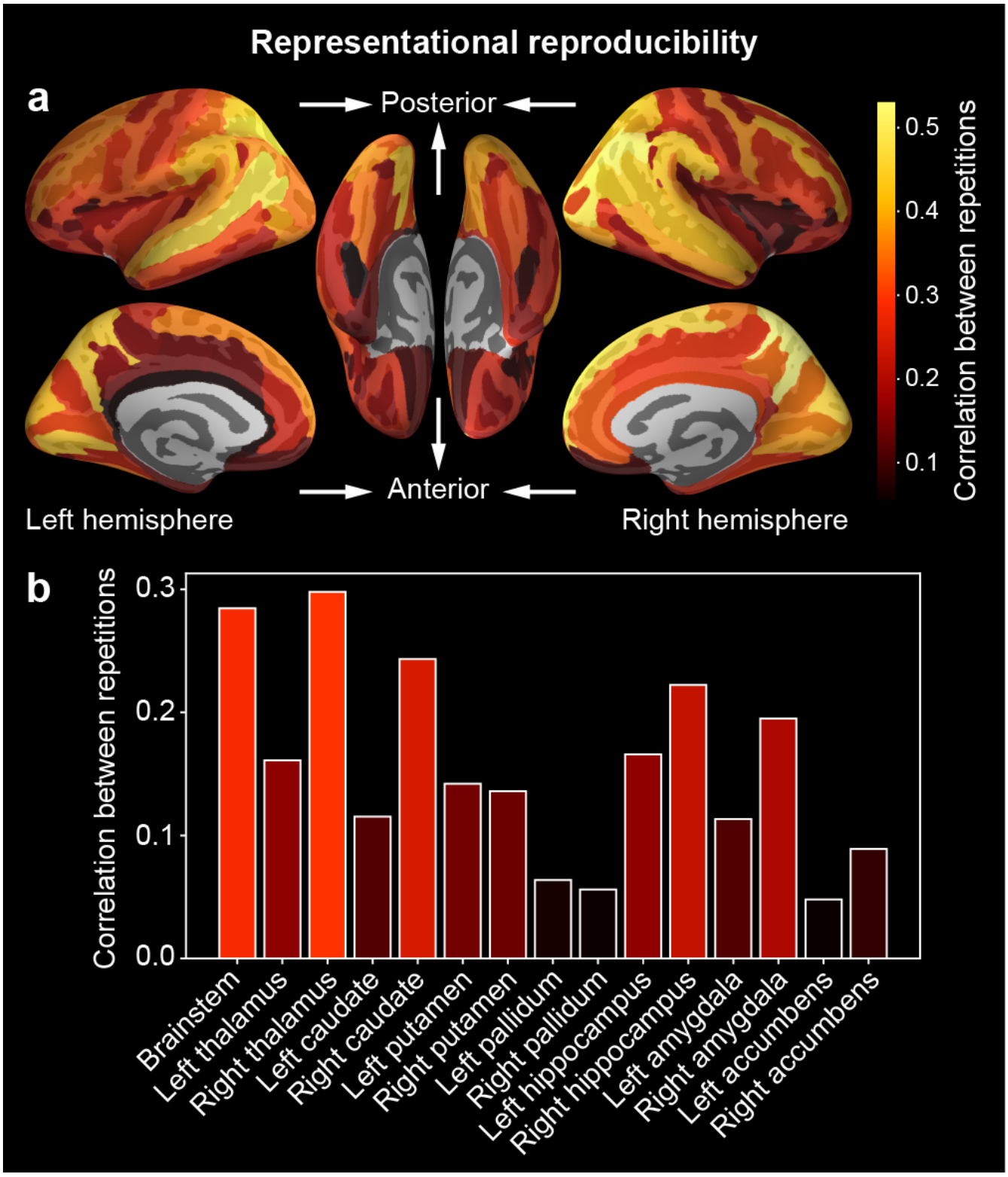

Figure 2

Representational reproducibility in each brain region. a, A cortical map of representational reproducibility. The reproducibility of audiovisual representations in each region was calculated using a Pearson correlation of the CRDMs of identical movie clip repetitions. The reproducibility averaged across all participants was mapped onto the cortical surface of a reference brain (top left: lateral view; bottom left: medial view; right: ventral view). Brighter locations indicate regions with larger reproducibility. b,

Reproducibility of CRDMs in subcortical regions. Each bar represents the Pearson correlation coefficient calculated for each subcortical region. The color of each bar was determined using the same colormap used in a. All cortical and subcortical regions showed significant reproducibility (P < 10−9, FDR corrected; See Supplementary Data 1).

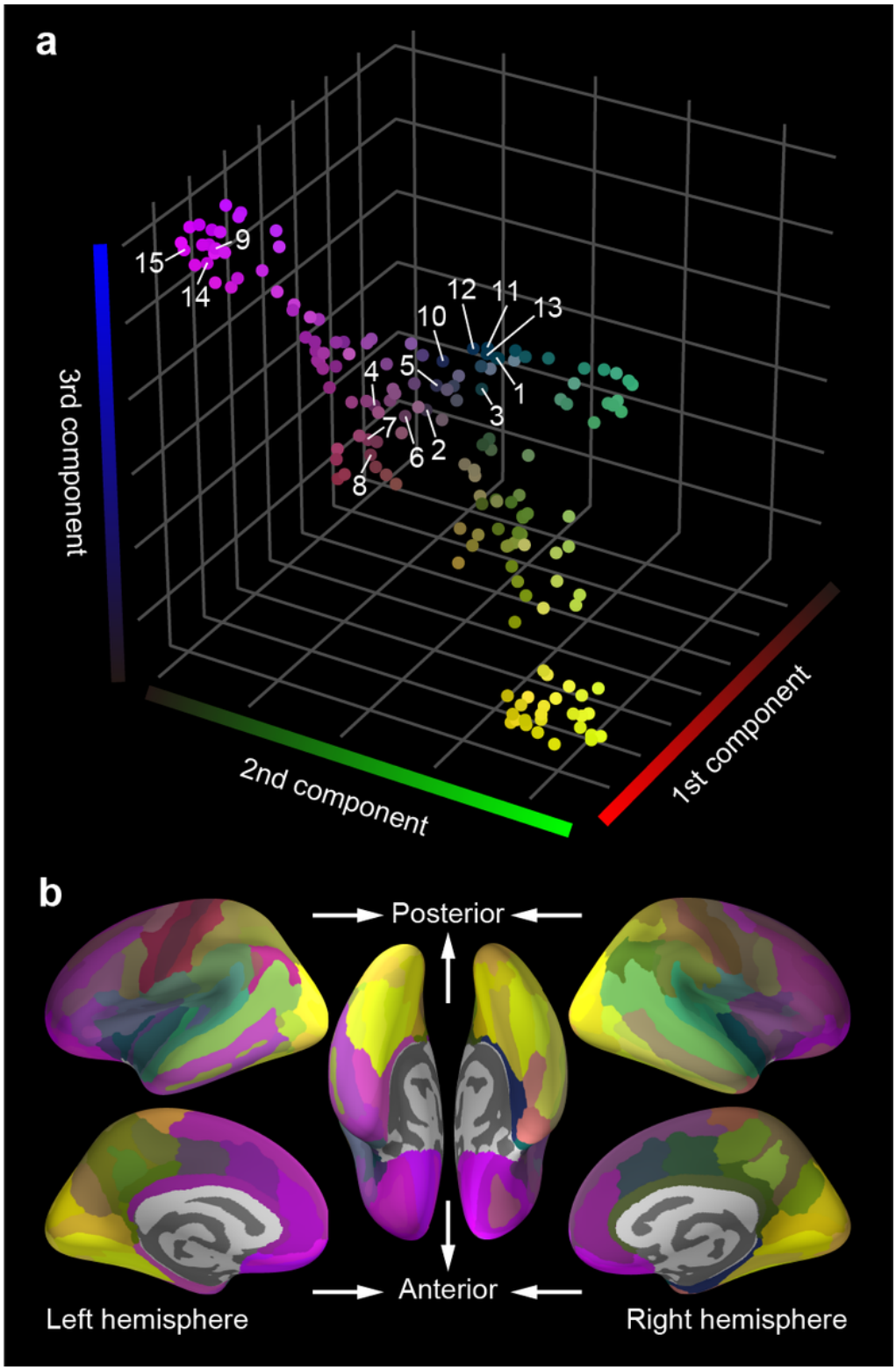

Figure 3

Representational variability across brain regions visualized using dimension reduction. a, Natural audiovisual representations in individual regions mapped onto a 3D representational space that was estimated from the similarity of CRDMs across regions using UMAP, an unsupervised dimension reduction method. This UMAP space preserves representational similarity between regions as proximity in the space. Each marker represents a CRDM in a single cortical or subcortical region. The color of each marker is determined by an RGB colormap based on the coefficients of three components (red, first component; green, second component; blue, third component) as indicated by the color bar along each axis. CRDMs of brain regions with similar colors demonstrate similar audiovisual representations. The markers for the subcortical regions are denoted by numbers 1–15 (1: brainstem; 2/3: left/right thalamus; 4/5: left/right caudate; 6/7: left/right putamen: 8/9; left/right pallidum; 10/11: left/right hippocampus; 12/13: left/right amygdala; 14/15: left/right accumbens). b, A cortical map of representational variability. The 3D representations of CRDMs in individual cortical regions were mapped onto the cortical surface using the same RGB colormap used in a.

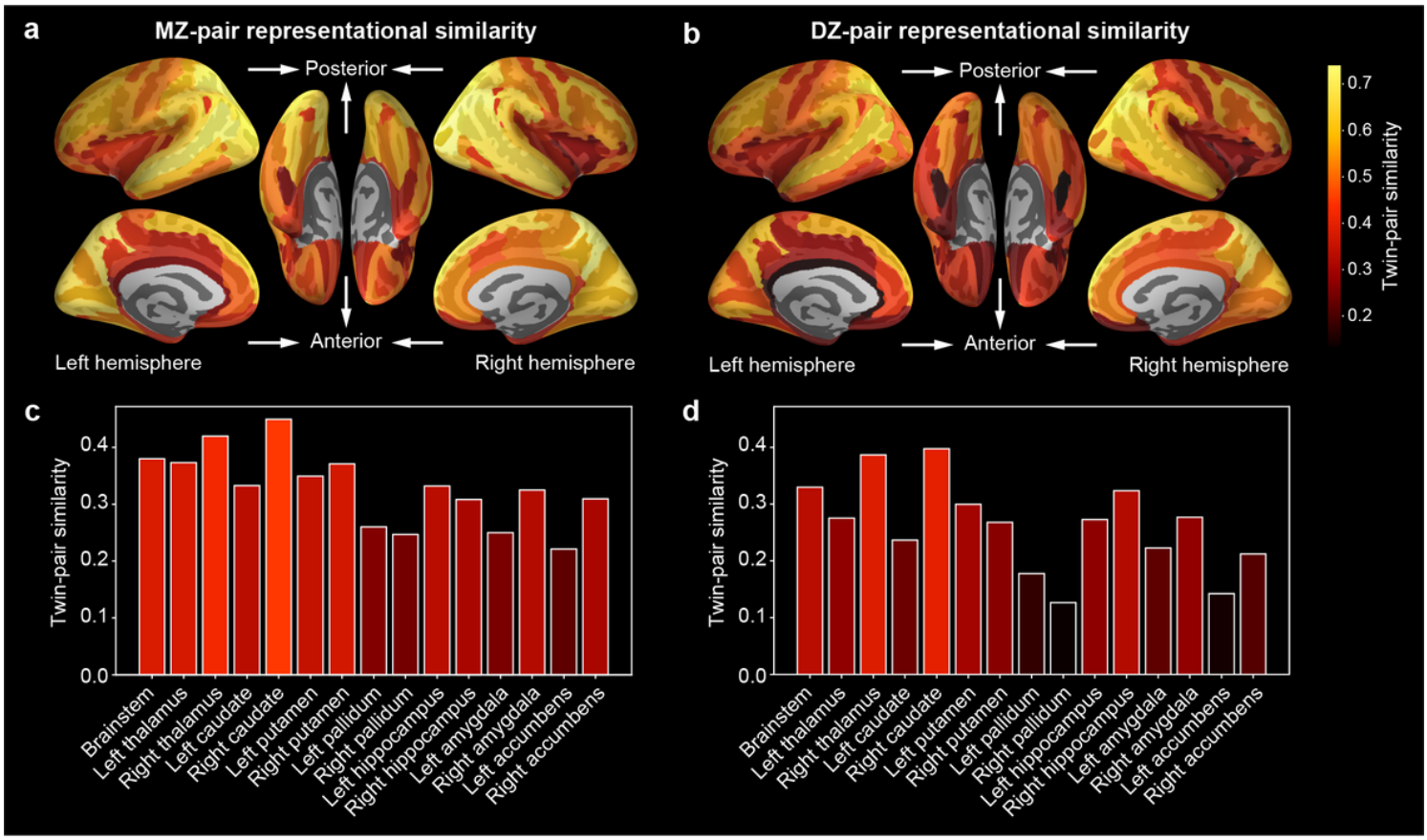

Figure 4

Twin-pair representational similarity in each brain region. a–b, Cortical mapping of twin-pair representational similarity. Twin-pair similarity of CRDMs corrected with the noise ceiling of each region was mapped onto the cortical surface separately for MZ (a) and DZ (b) pairs. Brighter locations indicate regions with higher similarity. The other conventions are the same as those in Fig. 2a. c–d, Twin-pair representational similarity in subcortical regions. Each bar represents corrected twin-pair CRDM similarity of each subcortical region for MZ (c) and DZ (d) pairs. The color of each bar in c and d is determined by

the colormap in a and b, respectively. All cortical and subcortical regions showed significant similarity (P < 10−9, FDR corrected; See Supplementary Data 3).

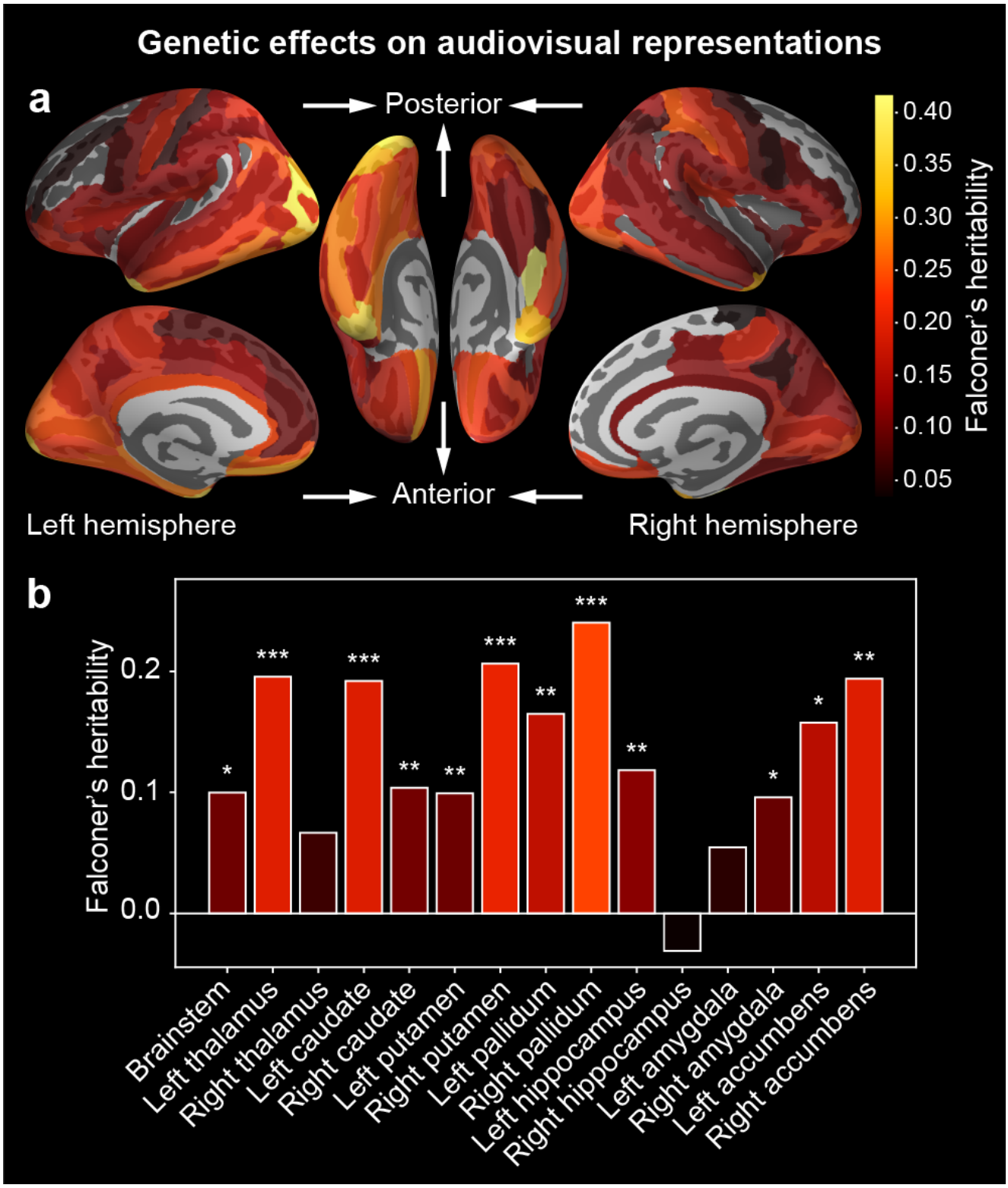

Figure 5

Genetic effects on the audiovisual representations of each brain region. a, Heritability of audiovisual representations in cortical regions. The Falconer's heritability calculated from the twin-pair similarity of CRDMs was mapped onto the cortical surface. Heritability is shown only in cortical regions with

significant genetic effects evaluated using significant differences between MZ- and DZ-pair CRDM similarities (Wilcoxson test, P < 0.05, FDR corrected). Brighter locations indicate regions with higher heritability. The other conventions are the same as those in Fig. 2. b, Heritability of audiovisual representations in subcortical regions. The colors of the bars correspond to those in the cortical map (a). Marks denote significant differences between MZ- and DZ-pair CRDM similarities (*P < 0.05, **P < 0.001, ***P < 0.00001, FDR corrected). See also Supplementary Data 5.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SuppData1.xlsx
- NishidaTwinsNaturalSuppInfo01.docx
- SuppData6.xlsx
- SuppData3.xlsx
- SuppData2.xlsx
- SuppData5.xlsx
- SuppData7.xlsx
- SuppData4.xlsx