nature communications

Article

https://doi.org/10.1038/s41467-025-62551-x

Invariant inter-subject relational structures in high order human visual cortex

Received: 20 May 2024

Accepted: 22 July 2025

Published online: 27 August 2025

Check for updates

Ofer Lipman[1], Shany Grossman[2,3], Doron Friedman[4], Yacov Hel-Or[1] & Rafael Malach[5]

It is a fundamental of behavior that different individuals see the world in a largely similar manner. This is an essential basis for humans' ability to cooperate and communicate. However, what are the neural properties that underlie these inter-subject commonalities of our visual world? Finding out what aspects of neural coding remain invariant across individuals' brains will shed light not only on this fundamental question but will also point to the neural coding scheme at the basis of visual perception. Here, we address this question by obtaining intracranial recordings from three groups of patients taking part in a visual recognition task (overall 19 patients and 244 high-order visual contacts included in the analyses) and examining the neural coding scheme that was most consistent across individuals' visual cortex. Our results highlight *relational coding* – expressed by the set of similarity distances between profiles of pattern activations—as the most consistent representation across individuals. Alternative coding schemes, such as activation pattern coding or linear coding, failed to achieve similar inter-subject consistency. Our results thus support *relational coding* as the central neural code underlying individuals' shared perceptual content in the human brain.

A fundamental aspect of our visual world, an aspect that is so basic that it is often taken for granted, is that different individuals appear to see the world in a similar manner[1]. Such inter-subject agreement is essential for our ability to point to visual targets using language and to succeed in cooperative actions. The obvious challenge to visual neuroscience research is then to explain these inter-subject commonalities of visual perception in terms of cross-individual congruency in the neural codes underlying our visual experiences. Simply put, we can ask: what is kept invariant in the neural coding of visual information as we move from one individual's brain to another? It should be noted that this question is relevant not only to explaining our shared visual world but also touches on the very basis of the neural coding mechanism that underlies the emergence of the rich array of our visual percepts.

Experimental and conceptual notions about the coding schemes in human visual cortex have progressed from the analysis of the functional tuning of cortical regions and individual neurons[2,3], through activation patterns[4], and finally, relational codes, in which the representation is defined by the distances between activation patterns[5–7]. We addressed this issue by studying neural responses in a large cohort of patients who underwent intra-cranial recordings in the course of a clinical diagnosis of epilepsy. This method provides highly detailed information, both spatially and temporally, of human neural responses. Neuronal responses, or activations, were defined as the power of the high-frequency broadband of the local field changes measured in each intra-cranially recording contact and time-locked to the stimulus presentation (see Methods for details). These high-frequency changes in power have been shown previously to index the aggregate neuronal firing in the vicinity of the recording contact[8–10]. However, the aggregate signal measured in this work should be strictly distinguished from the activity measured in individual neurons. In the experiment, the patients were presented with the same set of pictures of different categories while their neuronal responses were recorded, allowing

[1]Efi Arazi School of Computer Science, Reichman University, Herzliya, Israel. [2]Institute of Psychology, University of Hamburg, Hamburg, Germany. [3]Max Planck Insitute for Human development, Berlin, Germany. [4]Sammy Ofer School of Communications, Reichman University, Herzliya, Israel. [5]Department of Brain Sciences, Weizmann Institute of Science, Rehovot, Israel. e-mail: rafi.malach@gmail.com

comparison across contacts from a similar brain area across patients. In the rest of the manuscript, we will term these functional congruencies between patients "inter-subject invariance".

Formally, we denote by a_j^i the activation of a contact j when the subject is viewing an image i. Having a group of contacts, $\mathcal{P}$, of a particular subject in a particular cortical area, we define the population profile of these contacts as an activation pattern $\mathbf{a}_\mathcal{P}^i$ where the dimensionality of $\mathbf{a}_\mathcal{P}^i$ equals the number of contacts in the group $\mathcal{P}$ (see Fig. 1). We assume two groups of contacts, $\mathcal{P}$ and $\mathcal{Q}$, from two different brains in the same cortical area. We find the optimal pairing using a bipartite graph matching algorithm. This algorithm identifies the best one-to-one matching between the two groups, with the pairing criterion being the distance between the matched contacts' associated tuning curves, defined as 1-Pearson correlation. However, it should be emphasized that due to inherent and potentially major differences between the anatomical and functional organizations across individual brains, this does not necessarily imply that similar anatomical locations across brains should have identical functional properties. This problem is somewhat alleviated by the application of optimal matching between contacts across different brains (within the same cortical area) based on their corresponding tuning curves. We further assume for now that the two groups include the same number of contacts.

We examined the level of inter-subject invariance under three hypothetical coding schemes, which are illustrated graphically in Fig. 2. The most straightforward scheme we considered is that perceptual information is encoded by the activation pattern of cortical neurons. This coding scheme, denoted activation pattern coding, is arguably the most commonly considered in cortical neurophysiology[11–13] and is illustrated in Fig. 2. Here the assumption is that the direction of an activation pattern elicited by a particular visual stimulus should be preserved across subjects. Using the above notations, for two brains with activation patterns $\boldsymbol{a}_\mathcal{P}^i$ and $\boldsymbol{a}_\mathcal{Q}^i$ respectively, we expect that $\mathbf{a}_\mathcal{P}^i \propto \mathbf{a}_\mathcal{Q}^i$, for each image i; namely, the activation patterns $\mathbf{a}_\mathcal{P}^i$ and $\mathbf{a}_\mathcal{Q}^i$ should be identical or proportional to each other, for each image i. In other words, we say that the two activation patterns are *directional invariant*.

We have previously presented data compatible with this very basic scheme, showing, using fMRI, an inter-subject congruency in the tuning properties of cortical sites during naturalistic visual stimulation[14,15]. However, it should be emphasized that due to the limited spatio-temporal resolution of fMRI, these findings are only relevant to the coarser aspects of the cortical visual representations – i.e., regions selective to general visual categories and entire cortical areas. Going beyond these coarse measures necessitates brain recording methods of higher spatio-temporal resolution; hence, our choice is to use direct intracranial recordings in patients[16–18].

An important point that needs emphasizing is that, as illustrated in Fig. 2a, the *activation pattern coding* hypothesis does not necessarily predict similar overall activations across individuals; only that their activation profiles – represented by their directions in the vector space – should be preserved. Thus, the "distance" between two activation patterns $\mathbf{a}_\mathcal{P}$ and $\mathbf{a}_\mathcal{Q}$ is defined by their cosine distance:

$$dist(\mathbf{a}_\mathcal{P}, \mathbf{a}_\mathcal{Q}) = 1 - (\mathbf{a}_\mathcal{P} \cdot \mathbf{a}_\mathcal{Q}) / (|\boldsymbol{a}_\mathcal{P}||\boldsymbol{a}_\mathcal{Q}|) \quad (1)$$

where $\mathbf{a}_\mathcal{P} \cdot \mathbf{a}_\mathcal{Q}$ denotes an inner product between vectors, and $|\mathbf{a}_\mathcal{P}|$ indicates the L_2 norm. The assumption that the overall magnitude of activation is not a critical parameter in establishing perceptual content has recently received experimental support in explaining the perceptual stability of stationary images despite drastic changes in overall activation magnitudes[13,19]. Consequently, in the *activation pattern coding*, where we assume directional invariance, we expect that the $dist(\mathbf{a}_\mathcal{P}^i, \mathbf{a}_\mathcal{Q}^i)$ should be close to zero, for each image i.

While *activation pattern* has been considered the most basic cortical coding scheme, in recent years, there has been a growing interest in the possibility of an alternative coding scheme, denoted relational coding, echoing classical structuralist principles[6,20]. Relational coding is a theory that proposes that the brain represents information by the *correlations between the activation patterns* elicited by different *objects* or concepts, rather than by the actual patterns themselves. This theory suggests that the brain encodes information about the world by creating representations of the relationships between objects and concepts. In this model, visual information is not encoded by the activation pattern on its own but rather by the set of similarities and dissimilarities – also termed distances (see refs. 4,12–15) – that appear between the activation patterns when activated by different visual contents[12–16]. However, it should be emphasized that this study examines the relational coding as an invariant representation between different brains, while the identity of specific representations may differ across individuals. In terms of the vector space of activation patterns, with relational coding we would

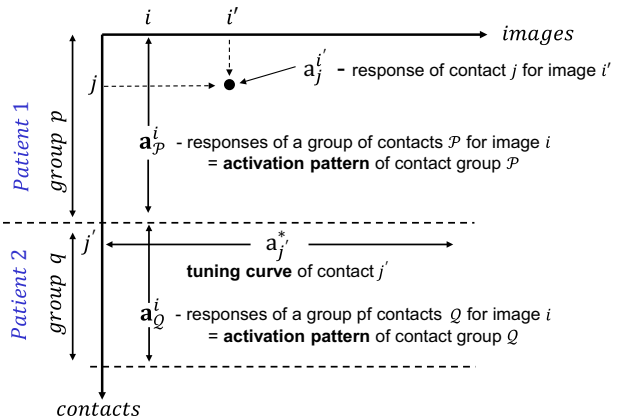

Fig. 1 | Population profiles of intracranial recordings captured as a matrix. Each row indicates a particular contact response to all visual stimuli (images), which is defined as the tuning curve of this contact. Each column indicates the responses of all contacts for a particular stimulus. A group of contacts, $\mathcal{P}$ or $\mathcal{Q}$, indicates a set of contacts in a particular subject and in a particular cortical area. A set of responses for a particular group, representing a profile of pattern activation, is rearranged in a vectorial form denoted as an activation pattern. For example, the activation pattern $\mathbf{a}_\mathcal{P}^i$ indicates the responses of all contacts in group $\mathcal{P}$ when the subject is viewing stimulus (image) i.

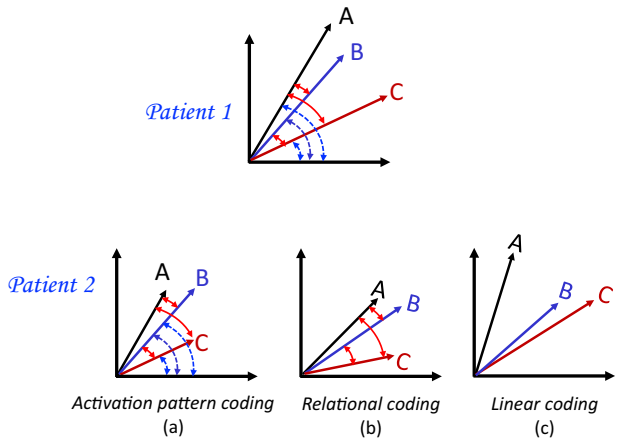

Fig. 2 | A schematic illustration of the three possible inter-subject encoding schemes examined. a Activation pattern coding, retaining the orientation of activation patterns between subjects, and accordingly also the intra-angle distances between activation patterns (blue dashed arrows and red solid arrows). **b** Relational coding, retaining the intra-angle distances between activation patterns (red solid arrows). **c** Linear coding, retaining the linear relations between triplets of colinear activation patterns (see text). Axes x,y represent activities in contact 1 and in contact 2. A, B, C denote 3 distinct images.

https://doi.org/10.1038/s41467-025-62551-x

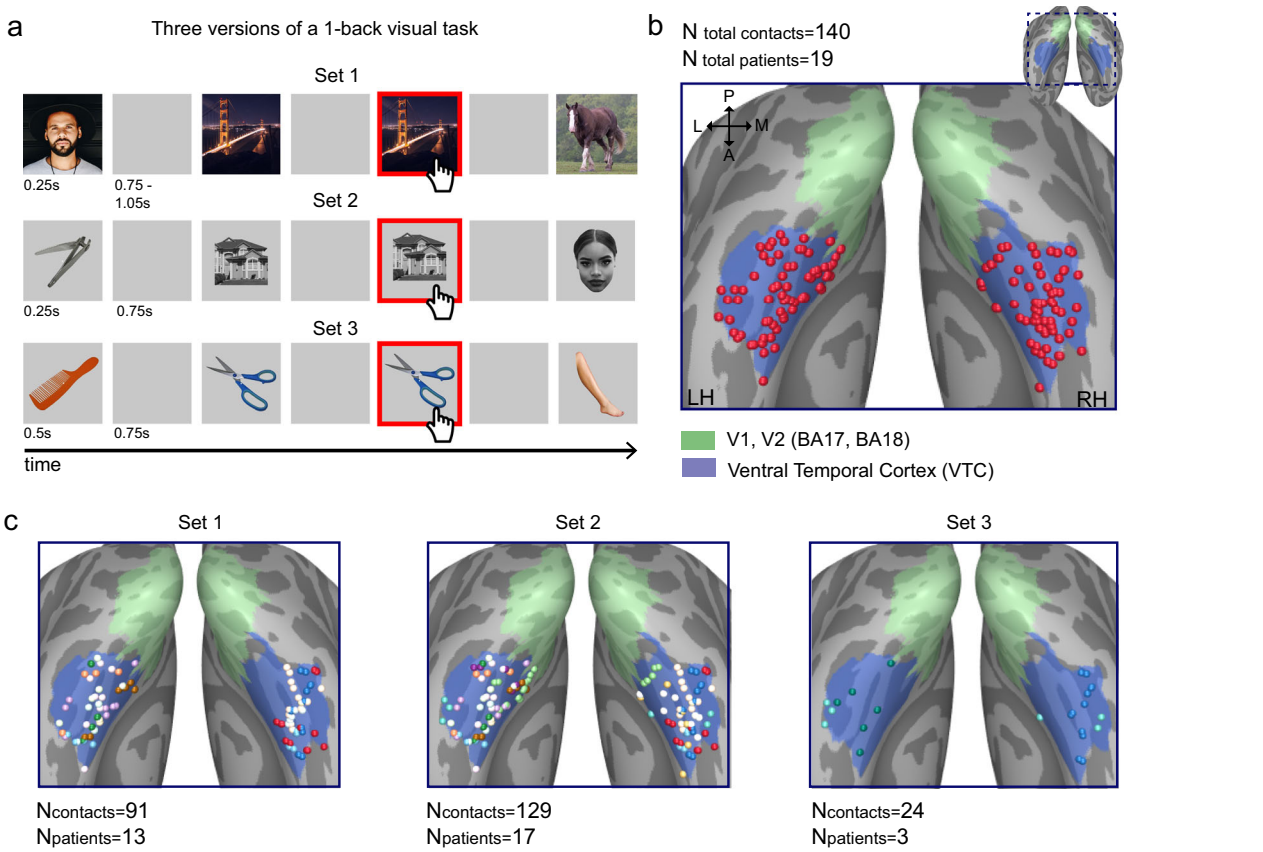

Fig. 3 | Experimental procedure and contacts' localization. a Three sets followed the 1-back visual recognition task in which patients were asked to press a button on rare occasions of image repeats. Presented images in sets 1 and 2 are variants similar to images originally used in the epxeriment due to copy rights limitations. **b** Locations of contacts included in the analyses pooled across all 3 sets (red circles), after being projected onto a common cortical surface template (FSAverage). High-order (VTC: Ventro-temporal cortex) and low-order (V1,V2) visual areas are denoted by blue and green shades, respectively. **c** Contacts' localization presented separately for each of the 3 sets. Contacts from different patients are denoted by different colors. Note that 14 patients took part in 2 out of the 3 tasks, leading to the total number of unique patients (19) being smaller than their sum across the 3 sets. LH/RH: left/right hemisphere. L, M, A, P: lateral, medial, anterior, posterior. Images of nail clipper, comb and scissors are taken from the BOSS opensource data base (Brodeur MB, Guérard K, Bouras M (2014) Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos. PLoS ONE 9(9): e106953.).

expect different individuals to preserve the mutual distances between activation patterns and not necessarily to preserve the original orientations of the activation patterns. Formally speaking, we say that activation patterns of two individuals are similar *up to rotation* (orthogonal transformation in high dimensional space, which is an angle-preserving transformation) while relaxing the inter-subject congruency between the activation patterns themselves. Using the above notations, it gives that for two brains with activation patterns $\mathbf{a}^i_{\mathcal{P}}$ and $\mathbf{a}^i_{\mathcal{Q}}$ respectively, we expect that there is an orthogonal transformation R, such that $R\mathbf{a}^i_{\mathcal{P}} \approx \mathbf{a}^i_{\mathcal{Q}}$, for each image i. Namely, the activation patterns $\mathbf{a}^i_{\mathcal{P}}$ and $\mathbf{a}^i_{\mathcal{Q}}$ are identical up to orthogonal transformation R. This yields that the angles between activation patterns in one subject are preserved in other subjects:

$$dist\left(\mathbf{a}^i_{\mathcal{P}}, \mathbf{a}^j_{\mathcal{P}}\right) = dist\left(\mathbf{a}^i_{\mathcal{Q}}, \mathbf{a}^j_{\mathcal{Q}}\right) \text{for each pair of stimuli } i,j. \quad (2)$$

This invariance is a direct outcome of the property that orthogonal transformation is angle-preserving. This prediction is illustrated in Fig. 2 and termed relational coding. Some support for the notion that this indeed may be the invariant coding across individuals has been recently obtained using fMRI. Thus[21,22], have shown that aligning the inter-pattern distances provides high consistency across individual brains, which is transferable across different visual stimuli.

Finally, we may continue along this line and increase the model complexity to the next step assuming that the perceptual code is invariant up to general linear transformation between individuals, namely, there is a general matrix M, such that $M\mathbf{a}^i_{\mathcal{P}} \approx \mathbf{a}^i_{\mathcal{Q}}$ for each image i. Intuitively, a linear transformation is a mapping that takes input vectors, and by stretching, shrinking, rotating, or flipping them, outputs new vectors (while straight grid lines before the transformation are kept straight after the transformation). This model is illustrated in Fig. 2 Panel (c). In that case, we may expect the angle distances between activation patterns to vary between individuals—while preserving a fixed, linear transformation between them.

To explore the above three alternative predictions, we have assembled a large group (19 patients implanted with 244 visually responsive contacts) of intracranial recordings in patients with the aim of examining which of the three coding models is best conserved across individual brains.

Results

To examine inter-subject similarity of visual responses, patients undergoing intracranial recordings for clinical purposes (see Fig. 3 and Methods), were presented with identical sets of visual images of different categories – such as faces, houses, and tools (for more details, see Methods), while their neuronal responses, measured as changes in high-frequency broad band power of the iEEG signal (see Methods) were recorded. Our experiment is based on a cohort of 19 patients that had contacts in high order visual cortex, suggested by a large body of prior research to be the central hub of perceptual content[6]. A total of

https://doi.org/10.1038/s41467-025-62551-x

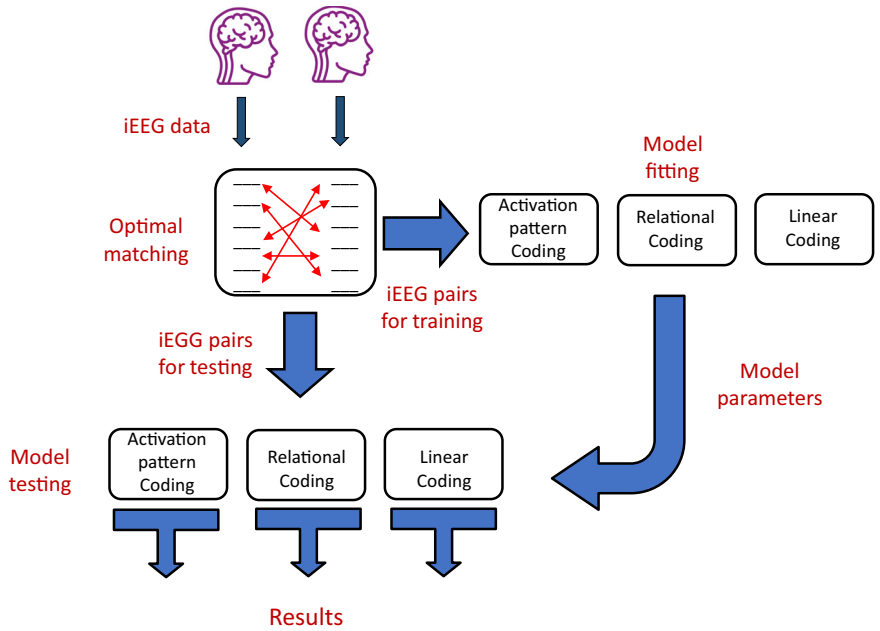

Fig. 4 | A schematic illustration of the procedure used to test the three possible encoding schemes. For each pair of patients, we first optimally match their recorded iEEG contacts using a bipartite graph matching algorithm. Next, we determine the model parameters for each encoding scheme (model fitting) using a portion of the stimuli (training set). Finally, we test the model's performance on the remaining stimuli (test set). This procedure is repeated multiple times, each time with different training and test sets (k-fold cross-validation). Brain icon image taken from Flaticon.com (https://www.flaticon.com/free-icon/brain_2491347).

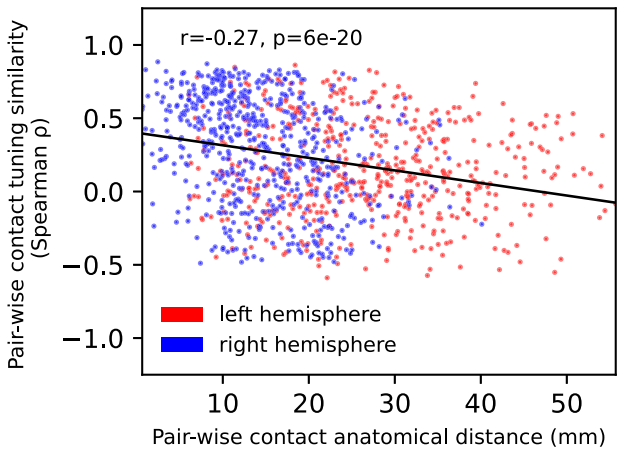

Fig. 5 | Functional similarity between pairs of contacts from different patients as a function of their anatomical distances. The *Y*-axis depicts the similarity of the tuning curve measured by Spearman correlation between contact pairs, while the *X*-axis depicts their anatomical Euclidean distance in mm (pairs taken from 19 patients with 244 contacts). Anatomical distances between contact pairs were estimated by first projecting all individual brains onto a standard surface template using SUMA[44] and computing the Euclidean distance between vertices corresponding to contact pairs on a standard pial surface. Only contact pairs localized in the same hemispheres and in high-order visual areas were evaluated. Note the significant but low correlation (two-sided one-sample *t*-test on the measured Pearson's r), indicating a high degree of local scatter in functional similarity across patients. Source data are provided in Source Data file.

244 contacts were examined in these patients (see Methods for inclusion criteria).

In the experiment (Fig. 3), visual images were presented for 250 ms (sets 1 and 2) or 500 ms (set 3) interspersed with blank periods of 750–1050 ms durations in which only the central fixation point was present. The patients' task was a 1-back recognition task, in which patients had to indicate rare images (7% for set 1, 12% for set 2, and 10% for set 3) repeats via a button press. Image repeats (i.e., the 1-backs) as well as false alarm trials were discarded from all analyses.

To assist in following the complex sequence of analysis and modeling, Fig. 4 depicts a flow-chart of the different stages.

The inter-subject similarity of anatomically neighboring contacts

The simplest model of inter-subject similarity assumes that contacts that are located in close homologous anatomical locations across patients should also show similar functional tuning compared to contacts that are anatomically located in disparate locations. To examine if this indeed was the case in our data, we compared the activation similarity of all possible pairs of contacts and across all patient pairs in high-order visual cortex. In our analysis, we calculated the inter-subject anatomical proximity of each pair of contacts – first by normalizing all patients' brains to a standard brain (see Methods). We then checked whether anatomical proximity is indeed translated to greater similarity in tuning curves (defined here as the pattern of the responses of a single contact to all visual stimuli, see Fig. 1) and across all individuals. The result of this large-scale analysis is depicted in Fig. 5. In line with our previous fMRI study (e.g. ref. 14), contact pairs that were located in close anatomical positions showed a significantly higher degree of functional correlation compared to distant pairs (Pearson's $r = -0.27$, $p < 0.001$, see Methods for more details). However, although significant, the correlation value was low, revealing a high degree of local scattering, which clearly violated the anatomical proximity rule.

Pairing contacts between subjects

Nevertheless, it could be argued that although homologous cortical locations may not be linked to identical tuning properties of individual contacts, such similarities in functional tunings may be apparent at the levels of groups of contacts, regardless of their precise location within the group, i.e., across activation patterns (see definition in Fig. 1) rather than individually matched homologous contacts. According to the hypothesis presented as "activation pattern coding" in Fig. 2, we do not

confine the search to exactly matched anatomical locations across individuals, but rather evaluate the similarity between contacts based on their corresponding activations.

In order to perform this location-free analysis, we derived an optimal matching between contacts across each pair of individuals in the following manner: As illustrated in Fig. 1, for a specific contact, say contact j, we define the tuning curve $\mathbf{a}_j{}^*$ as the vector of this contact responses to all visual images. To calculate the similarity between two tuning curves across two contacts, $\mathbf{a}_j{}^*$ and $\mathbf{a}_k{}^*$, we take the Spearman's rank correlations between $\mathbf{a}_j{}^*$ and $\mathbf{a}_k{}^*$, denoted by $S\left(\mathbf{a}_j{}^*, \mathbf{a}_k{}^*\right)$. For two groups of contacts, $\mathcal{P}$and$\mathcal{Q}$, each of which belongs to a different patient, we define the similarity score as the mean of the Spearman's correlations:

$$S_{\mathcal{P},\mathcal{Q}} = \frac{1}{n}\sum_{j\in\mathcal{P}} S\left(\mathbf{a}_j^*, \mathbf{a}_{j'}^*\right) \tag{3}$$

where for each contact j in group $\mathcal{P}$, we denote $j' = \Pi_{\mathcal{P}\to\mathrm{Q}}(j)$ as the matched contact in group $\mathcal{Q}$. We assume we have n contacts in each group, thus we normalize by n. The best match of the two groups can be sought such that the similarity score will be maximal:

$$\hat{\Pi} = argmax_{\Pi}\sum_{j\in\mathcal{P}} S(\mathbf{a}_j^*, \mathbf{a}_{j'}^*) \text{where} j' = \Pi_{\mathcal{P}\to\mathrm{Q}}(j) \tag{4}$$

This optimization can be formalized as a bipartite graph matching problem[23] that finds the optimal matching between contacts in group $\mathcal{P}$ (patient A) and contacts in group $\mathcal{Q}$ (patient B). If the number of contacts in the two brains is not identical, the solution provides the maximum number of possible pairs (that is, the contacts' count in the patient with the smaller number of contacts). To evaluate the role of the optimal pairing on the results, we conducted an identical analysis while randomly pairing the contacts across subjects (see methods). The results are depicted in Supplementary Fig. 2. As can be seen, the random pairing had a rather negligible effect on the relational and linear coding while substantially reducing the activation pattern-based correlations.

Inter-subject activation pattern Invariance

In order to compare the three hypotheses, we define a measurement between subjects that we term Inter-Subject Similarity Index (ISSI). For the first hypothesis, as outlined above, instead of selecting homologous cortical locations, we search for the best possible one-to-one mapping of contacts between each pair of subjects. Specifically, by permuting all possible combinations of contacts between each pair of patients, the maximum overall Spearman correlation over all possible pairs of patients was established (see Methods). We then calculate the ISSI scores for the activation patterns across all images as follows: Assuming a pair of patients with corresponding sets of contacts, $\mathcal{P}$ and $\mathcal{Q}$, where the permutation matrix P is the optimal permutation between $\mathcal{P}$ and $\mathcal{Q}$ according to the best matching procedure (the matrix P is extracted from the permutation mapping $\hat{\Pi}$ above). The ISSI between $\mathcal{P}$ and $\mathcal{Q}$ is calculated based on the Spearman correlation:

$$ISSI(\mathcal{P},\mathcal{Q}) = \frac{1}{|\mathcal{P}|}\sum_i S\left(P\mathbf{a}_{\mathcal{P}}^i, \mathbf{a}_{\mathcal{Q}}^i\right) \tag{5}$$

where in this case, the correlations are summed over a set of images (index i), and the Spearman correlation is calculated between activation patterns rather than tuning curves.

Critically, this analysis was performed on a randomly selected subset of images, that will be termed the training set (80% of the data). Once the contact set was selected based on this training set, the ISSI was calculated on the remaining images that were not part of the optimization process, which will be termed the test set.

The results of this analysis of the activation are presented in Fig. 6 (red curves) and in Fig. 7 (red bars). Figure 6 shows an example of ISSI values of the activation patterns across all pairs of patients, when they were presented with the category of animal images in set 1. We plot all ISSI values in an increasing order. The lower panel shows the ISSI for the training set of images, while the upper panel shows the ISSI values for the test set. As can be seen, there was a small positive and consistent similarity across individuals (bottom panel) that—importantly—was maintained in the independently measured test set (top panel). Supplementary Fig. 1 depicts the same analysis for the other image categories in set1, showing, again, a small but consistent increase in the ISSI values of the activation patterns across all object categories. Figure 7 depicts, for each category, the average ISSI values across all patient pairs for the activation pattern similarity analysis (red bars). Similar to the results of Fig. 6, we see a small and positive inter-subject similarity in this measure across all different categories, both in the training as well as the test sets.

Inter-subject relational coding Invariance

The second hypothesis we consider is that of *relational coding*. In this scheme, the neural representation of a concept is based on how similar and dissimilar this concept is from other encoded concepts. Formally, a *relational coding scheme* implies that the activation codes are invariant across individuals up to orthogonal transformations; namely, there is an orthogonal transformation R, such that, for each image i, $R\mathbf{a}_{\mathcal{P}}^i \approx \mathbf{a}_{\mathcal{Q}}^i$. The orthogonal transformation, R, which can be viewed as rotation in high-dimensional space, is a distance-preserving transformation, thus, if indeed the coding between individuals is identical up to rotation, we have:

$$dist\left(\mathbf{a}_{\mathcal{P}}^i, \mathbf{a}_{\mathcal{P}}^j\right) = dist\left(\mathbf{a}_{\mathcal{Q}}^i, \mathbf{a}_{\mathcal{Q}}^j\right) for\ each\ pair\ of\ images\ (i,j) \tag{6}$$

Here, the activation patterns are not preserved across individuals, but rather the inter (angle) distances between these vectors. At this point, it is worth mentioning that invariance up to permutation (the 1st hypothesis) is a special case of invariance up to rotation (permutation is a particular case of rotation). Therefore, it is expected that the ISSI score after optimal rotation will exceed the ISSI after optimal permutation (1st coding scheme). However, this will not necessarily be the case when the rotation is determined on a training set and evaluated on an independent, different, test set.

To examine whether indeed such relational coding was preserved across patients we conducted the following analysis. First, for each category of visual stimuli and each pair of patients we applied the optimal-match procedure as discussed above. After this step, in each pair of patients the same number of contacts that are optimally correlated were selected for further analysis. Next, we divided the set of all visual stimuli into training and test sets (on average, over 100 train/test splits, with each split 2/3 training and 1/3 test). Next, we computed the optimal rotation (orthogonal transformation) for each pair of patients on the training set, and then applied the evaluated rotation to the test set (see Methods). The optimal rotation, $\hat{R}$, was evaluated by solving (see Methods):

$$\hat{R} = argmin_R \sum_{i\in\, train\_set} ||R\boldsymbol{a}_{\mathcal{P}}^i - \boldsymbol{a}_{\mathcal{Q}}^i||_2 \tag{7}$$

To evaluate the similarity between the two patients after applying the optimal transformation, we calculated the ISSI measure (as it is more robust for monotonic value mapping) again on the test set:

$$ISSI(\mathcal{P},\mathcal{Q}) = \frac{1}{|test\_set|}\sum_{i\in\, test\_set} S\left(\hat{R}\mathbf{a}_{\mathcal{P}}^i, \mathbf{a}_{\mathcal{Q}}^i\right) \tag{8}$$

https://doi.org/10.1038/s41467-025-62551-x

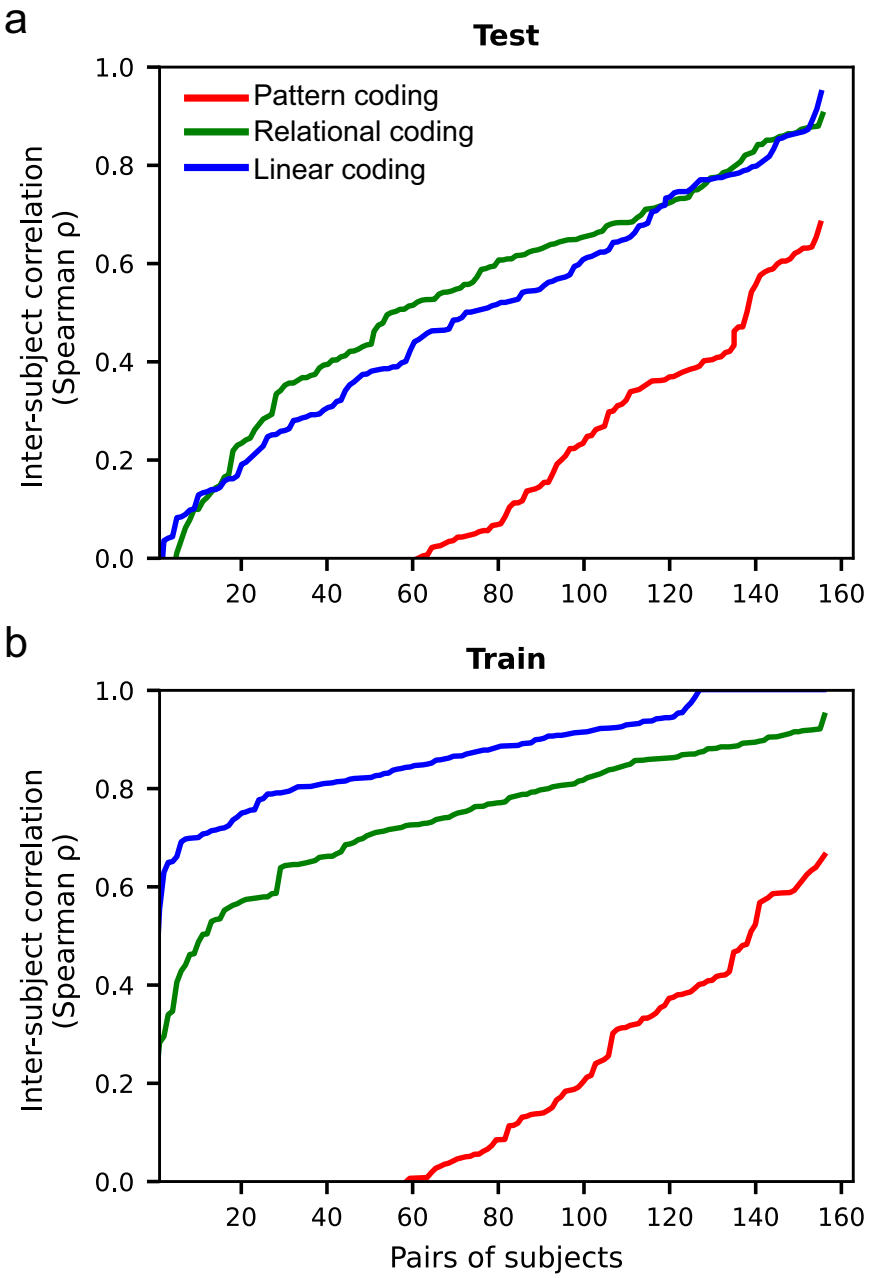

Fig. 6 | Inter-subject similarity index (ISSI) of the three neural coding schemes for the animals category (set1). a Independent test set. **b** Training set. In both panels, the *Y*-axis depicts the mean of Spearman correlation values between pairs of patients (ISSI). *X*-axis depicts the pair index. The ISSI values are sorted and shown in increasing order. Red, green, and blue colors indicate the ISSI for optimal permutation (activation pattern matching), rotation (relational code matching), and linear transformation matching, respectively. All three coding schemes show a high measure of inter-subject invariance; however, for the test set, the relational coding scheme (rotation) shows a clear enhancement compared to the other coding schemes. Source data are provided in Source Data file.

The results of this analysis are presented by green curves in Fig. 6 (lower and upper panels for training and test sets, respectively) for the category of animal images (and for all other categories in Supplementary Fig. 1), and by green bars in Fig. 7 for all categories. As can be seen, there was a robust and consistent increase in the inter-subject similarities when relational coding formed the basis of comparison as opposed to the activation pattern coding scheme. Such an increase is indeed to be expected for the training set; however, critically, a robust and significant increase was also evident when examining the test set (Wilcoxon signed-rank test between ISSI values for relational match and activation pattern match). The effect was consistent across all image categories (see Figs. 6, 7, and Supplementary Fig. 1).

Thus, we find that relational coding, as reflected in the rotational analysis, shows consistently and robustly, a better match across subjects and image categories compared to the activation pattern coding.

Inter-subject Linear Invariance

The observation that relational coding is matched across individual patients raises the possibility that perhaps this similarity will be further enhanced if we further relax the constraints on the image coding schemes. A straightforward next step in such a relaxation process can be applied by extending the transformation from orthogonal to linear transformations (see illustrations in Fig. 2); as mentioned, if the orthogonal transformation yields representation invariance up to

https://doi.org/10.1038/s41467-025-62551-x

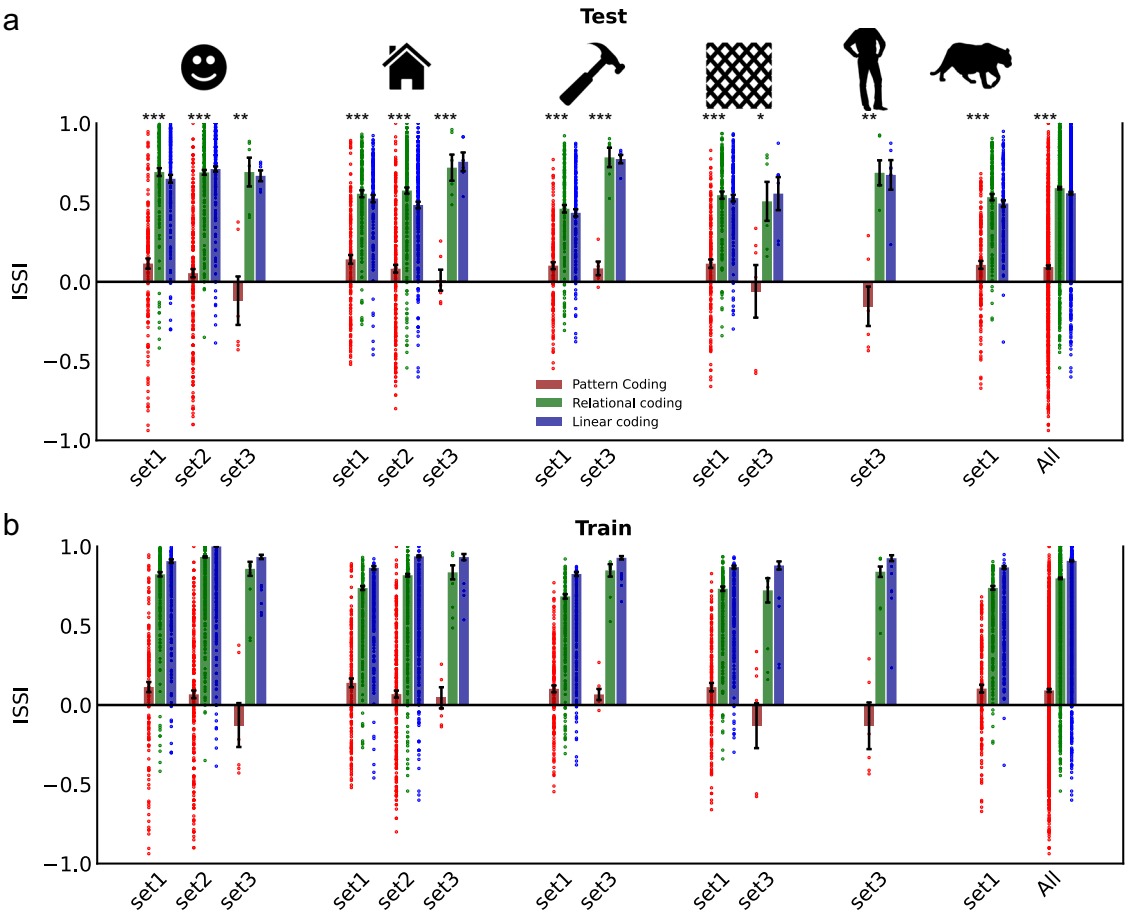

Fig. 7 | Comparison of mean inter-subject similarity index (ISSI) across the three coding schemes for all visual categories. a Independent test set. **b** Training set. Red, green, and blue bars depict the mean ISSI values +/- SEM for the activation pattern coding, relational coding and linear coding, respectively, across all subject pairs and for each image category. Image categories from left to right: faces, houses, tools, patterns, body and animals. Single dot denoting individual pairs are overlaid on bars (set 1, 2 and 3 include 156, 272 and 6 subject pairs, respectively). The last set on the right denotes the averaged mean ISSI +/- SEM across all categories. Note the significant (two-sided Wilcoxon signed-rank test) increase for the inter-subject invariance of the relational coding relative to the activation pattern coding. *, **, *** indicate $p < 0.05$,0,01 and 0.001, respectively. The small but significant (non-parametric two-sided statistics on all signed differences, relational higher than linear in 9/12 samples, $p = 0.05$) reduction in the linear transformation in the validation set suggests that the linear transformation overfits the inter-subject invariance (see Fig. 1), leaving the relational coding scheme the best model for invariance. Source data are provided in Source Data file.

rotation, the linear hypothesis yields invariant representations up to a linear transformation. Such a transformation maintains ratios of distances between three co-linear points, but its rotation does not preserve distances and angles.

To examine this hypothesis, we conducted the same analysis as the one carried out for relational coding, only this time using linear transformations rather than orthogonal transformations. Note, also in this case, that the orthogonal transformation is a particular case of the more general linear transformation model. The optimal rotation of a linear transformation is evaluated by solving (for more details see Methods):

$$\widehat{M} = argmin_M \sum_{i \in train\_set} ||M\mathbf{a}_{\mathcal{P}}^i - \mathbf{a}_{\mathcal{Q}}^i||_2 \quad (9)$$

Where M is a general linear transformation matrix.

Evaluating the similarity between the two patients after applying the linear transformation is calculated, similarly, as follows:

$$ISSI(\mathcal{P}, \mathcal{Q}) = \frac{1}{|test\_set|} \sum_{i \in test\_set} S\left(\widehat{M}\mathbf{a}_{\mathcal{P}}^i, \mathbf{a}_{\mathcal{Q}}^i\right) \quad (10)$$

The results for this transformation are presented in Fig. 6 in blue curves (see Supplementary Fig. 1 for all additional categories of images). As can be seen, in the training set, this richer transformation produced a higher level of similarities compared to the relational code. This is to be expected since the linear transformation has more degrees of freedom. However, critically, when examining the test set, in most cases the similarities fell below the orthogonal transformation (non-parametric statistics on all categories, 9 out of 12 categories show a preference for relational over linear coding, $p = 0.05$). This is further demonstrated in Fig. 6 and Supplementary Fig. 1. These results indicate that the enhanced power of the linear transformation likely resulted in over-fitting.

Neural coding across image categories

To examine how consistent were these results across patient groups and image categories, Fig. 7 compares the ISSI levels of the three neural coding schemes, calculated separately for the different image categories and for the contacts corresponding to the three independent datasets in high-level visual areas (see Fig. 3c). The lower panel shows the results for the training set while the upper panel shows results for the test data. As can be seen, and similar to Fig. 6 and Supplementary Fig. 1, there was a small and fairly consistent level of ISSI of activation patterns across patients, as revealed by the positive red bars across most (9 out of 12) image categories. Importantly, this inter-subject activation pattern similarity was apparent also in the independently validated set. However, critically, there was a consistent and striking

enhancement in the ISSI level when shifting to the relational coding scheme. The effect was consistent across all image categories (see green colors in Fig. 6, Supplementary Fig. 1, and Fig. 7), attesting to the cross-category superiority of relational coding.

Could the results be affected by the number of contacts or images? To examine this possibility, we simulated the three coding schemes across two simulated representations and varied the number of contacts and images for each of these models. The results of these simulations are depicted in Supplementary Figs. S3 and S4. As can be seen, despite the variation in the number of contacts and images, for each modeled coding scheme the one that appeared superior in the testing phase was the corresponding scheme that underlies the simulation.

Discussion

Inter-subject invariance as a window into cortical perceptual codes

Our strategy in the current work was to search for coding schemes that are preserved across the visual systems of different individuals. As discussed in the introduction, finding such cross-subject invariances is of great interest since our visual world is primarily shared, and individuals most often agree on how things look like. It should be emphasized that, given the shared nature of the visual world, finding neural codes that are not preserved across individuals can be taken, tentatively, as an argument against their essential role in perception. By contrast, those codes that evolution appears to have preserved across individuals' visual systems are likely to play an essential role in visual perception. As we will discuss in detail below, following this logic and comparing three neural schemes of such perceptual coding clearly highlights relational geometries as a likely essential coding scheme of perceptual images in the human brain.

Interestingly, random matching of contacts across individuals had a rather mild impact on the relational and linear-based correlations (see Supplementary Fig. 2) while substantially decreasing the activation pattern correlations, further illustrating the robustness of the relational coding scheme. Note, that orthogonal and linear transformations include permutation transformations, which, in fact, are the transformations performed in the optimal pairing.

Relational coding of visual perceptual content

Our results reveal a weak but consistent inter-subject invariance in activation pattern coding (Fig. 7). However, our central finding is that when the activation pattern for each image is allowed to vary across individuals—while the inter-vector angles are kept invariant, i.e., preserving the relational coding across individuals, the inter-subject similarities were strikingly and significantly enhanced (compare green vs. red lines and bars in Figs. 6, 7 and Supplementary Fig. 1).

It is important to clarify that it is to be expected that relational coding should be better matched across individuals compared to activation pattern coding when examining the training set alone. This may be a trivial consequence of the larger flexibility (more degrees of freedom) we allow in the comparison across patients in the former compared to the latter models. However, critically, the fact that such enhanced inter-subject similarity was transferred to the test set demonstrates that the preservation of relational similarity distances is a genuine cortical feature and not a technical consequence of a better fit due to a more flexible model. Moreover, following the same logic, it is expected that the linear coding scheme should better match across individuals in the training set compared to the relational coding scheme, as orthogonal transformation is a special case of a linear transformation. However, it is apparent that in the test set, the relational coding still surpasses the linear coding, although it has fewer degrees of freedom. Thus, our central finding points to relational similarity distances as a likely functional cortical feature.

The essential representational scheme underlying the content of human shared perception

It should be clarified in what manner relational coding is fundamentally different from the better-known neural tuning curves and activation patterns coding schemes. The fundamental difference is that according to the single unit and activation-vector schemes, the information pertaining to the visual stimuli is coded in the unique patterns of activations to each stimulus. By contrast, in relational coding, the specific activation pattern profile elicited by each stimulus is irrelevant to the perceptual content. Instead, what defines the content of a visual percept is the set of similarities and dissimilarities, i.e., the distances between the neural population profiles elicited by this percept and those of all other percepts. The neural mechanism underlying these relational geometries is embedded in the unique synaptic structure of each cortical region and is likely revealed through recurrent neural mutuals within each neural structure[6].

In recent years, relational geometries have emerged as a central concept in artificial networks, as well as in characterizing the primate visual cortex[24–26]. It may be informative to note here that similar relational geometries are found across different artificial networks[27]. Importantly, a growing number of studies have demonstrated a link between relational coding and behavior. These studies include perceptual similarity measures[7,28], semantic categories[29], emotional representations[30], and even in episodic memory semantization processes[31]. The present study extends these findings by demonstrating that despite the unique nature of individual visual representations, there is a cross-individual preservation of the relational geometries—thus highlighting this coding scheme as essential for human perception.

It is also interesting to note here that a similar invariance has been found in artificial neural networks (ANNs). Empirical evidence suggests that two corresponding layers in two different ANNs, when trained on the same dataset but with different initializations, exhibit similar responses up to orthogonal transformations[27]. This phenomenon corresponds to relational coding.

Although we have emphasized in this work the overall similarity in inter-individual perceptions, it is interesting to consider possible predictions of differences in those cases that show marked inter-individual differences. One interesting prediction concerns the phenomena termed "the other race effect" in which individuals tend to lack face discrimination when viewing unfamiliar races[32]. Here the prediction will be that the correlation of relational coding for faces belonging to one of the races will be correspondingly different between individuals exposed to the same race vs. those exposed to different races.

In the extreme case, where there is a complete dissociation between the perceptual contents of two individuals, e.g., for a specific category, our model will predict no correlation between the relational codes of two individuals exposed to the visual category. Such cases could be when comparing color perception in a totally color-blind individual (rod monochromat) and an individual possessing normal color vision.

Weak inter-subject tuning similarity across anatomically similar homologous loci

Our results show a weak but significant similarity in tuning of individual contacts occupying similar anatomical locations across individuals (Fig. 5). This result is compatible with previous findings of such inter-subject synchronization revealed through fMRI (e.g.[33],). However, this similarity was very small at fine resolution, likely masked by a substantial level of local scatter (see Fig. 5). A likely source for the weak relationship may be individual variabilities in the precise location of homolog cortical regions, which may occupy different anatomical loci relative to the major brain landmarks used for cortical alignments. Finally, it is plausible that the anatomical mosaic of functional

properties at high spatial resolutions varies randomly across individual brains, while at a coarser level, it is preserved. This will result in a lack of precise functional matching between homologous cortical points at the local scales within single cortical regions while showing consistent inter-subject correlations at coarser scales, as was indeed demonstrated by[14].

A major concern related to these findings could be that they reflect the limitation of the data structure, such as a small number of contacts per patient or a small number of images in each visual category, rather than the true cross-subject invariance of coding schemes. Addressing this concern using simulations, we failed to confirm this possibility. Specifically, when modeling two representations with similar activation patterns, reducing or extending the contacts' number failed to produce superiority of the relational coding correlation. Similarly, in the relational coding scheme extending the images' number did not result in the linear coding achieving superiority.

An interesting point concerns the temporal dynamics, i.e., the possibility that the activation patterns and relational coding may fluctuate over time. In the present work, this issue was not explored since the HFA signal was averaged over a 50–350 ms time window. The potential role of fast dynamics was a topic of two previous studies[13,24] where we show that at least within the resolution of the iEEG method, these neural representations appear to remain stable over short time scales (1–2 s).

Potential theoretical Implications of the present findings

It should be noted that relational structures, since their original introduction to neuro-cognitive research[7,26] have become an established research practice in cognitive neuroscience. However, demonstrating the functional role, if any, of relational structures in brain representations is still an on-going work. Some support for the functionality of relational information comes from correlations found between biological and artificial relational structures in previous studies (e.g refs. 24,34,35), which we more recently termed "convergent evolution"[36].

In the present work, we contribute further support for the functionality of relational information by demonstrating that relational information may provide a superior explanation for the inter-individual similarity found in human visual perception compared to population-vector coding. It should be noted that while our results demonstrate what may be considered as stability of relational information over space, i.e., across different brains, a complementary support for the stability of relational information has been demonstrated over time, i.e., within the same brain. In a recent study[37], the authors found superior stability of relational information over activation pattern coding within the same brains of rodents when measuring responses to identical movie stimulation over days. Together, our present results and those of others provide support for a significant functional role for relational information in visual perception.

If indeed relational information plays a central functional role in visual perception, then a central question concerning this function is this: In real natural conditions, given a single visual stimulus, which results in a single activation pattern, how can the relations between different stimuli be derived? To put it differently, how can a single activation pattern reflect a set of comparisons to activation patterns associated with different stimuli that are not presented to the individual at the moment of perception? While the actual mechanism that may implement such relational comparisons in the human brain is still unknown, we have previously proposed that a plausible mechanism underlying such comparisons may be the local synaptic structures within each cortical region[38]. The critical point to emphasize concerning this hypothesis is that the comparisons essential for deriving relational information are conducted prior to the emergence of each activation pattern and, in fact, determine its content. It should be noted that the dynamics of the visual responses offer sufficient time for such pre-activation computations, as the activation patterns generated in high-order visual areas emerge at a delay of about 100 msec following the early visual cortex activation[13]. As was proposed previously[38], the dynamic signature of such neural activations strongly suggests a process of rapid local recurrent activity that takes place prior to the appearance of the observed single activation pattern. It is hypothesized that the relational information is stored in the synaptic weights of the local circuits. The singular activation pattern that is observed following a single stimulus, should be viewed as the outcome of rapid filtering and amplification of incoming inputs by the underlying relational structures stored in the local neurons' synaptic weights. The relational structures that are revealed in detailed experiments by sequentially presenting one stimulus at a time are merely an experimental tool that allows us to expose these underlying local synaptic weights. These synaptic structures store the unique relational structure defining the functional specializations of each cortical region. While speculative at this point, this hypothesis can be tested in future studies, especially in animal models where local synaptic activity can be causally modulated.

We demonstrate here, using three iEEG data sets, that the neural perceptual coding scheme that is most preserved across individuals is relational coding. In this coding scheme, the similarity between activation patterns is maintained across individuals rather than the tuning profiles of individual contacts or the activation patterns themselves. The possibility that relational coding may play a fundamental role in human visual perception opens up interesting future possibilities of clinical significance, such as useful tools for a wide range of neural diagnosis tools as well as for brain machine interfaces. These findings offer a robust explanation for the ability of different individuals to perceive the world in a similar manner and point to relational information as an essential cortical code underlying visual perception.

Methods

Participants

Data were retrieved from a study consisting of three tasks and reported in two previous studies[24,39]. Sixty-eight participants were monitored for pre-surgical evaluation of epileptic foci. Among these participants, 19 had a minimum of 4 visually responsive selective contacts located in high-order visual areas (14 of which took part in two out of the three tasks, as detailed in Supplementary Table 1). All the participants gave their fully informed consent, including consent to publish, in compliance with NIH guidelines. This was done under the supervision of the institutional review board at the Feinstein Institute for Medical Research, in line with the principles of the Declaration of Helsinki.

Experimental design

The three tasks performed were 1-back visual tasks in which different images are presented sequentially with rare (7–12%) image repeats. Patients were instructed to keep their gaze on the central fixation point throughout the task and to click the mouse button whenever they observed consecutive repetitions of the exact same image. The set up did not include eye tracking recordings. For two patients in set 1, instructions were pressing on every second image. For two patients included in set 2, instructions were passive viewing while maintaining fixation, with no task. For one patient included in set 3, instructions were to press after each image.

Each task consisted of a different set of face images, houses/places, objects, patterns (sets 1 and 3 only), animals (set 1 only) and bodies (set 3 only). The number of exemplars in each visual category ranged between 6 and 10 (see Fig. 3a for a schematic illustration of the three tasks). Each patient performed either one or two versions of the three (for individual specifications, see Supplementary Table S1).

The first task (set 1) consisted of a total of 60 stimuli: faces (public figures collected from the internet), words, animals, tools, patterns and

places, 10 instances for each class. Throughout the task, each image exemplar was presented a total of six times. During the task, images were displayed for 250 milliseconds, followed by a jittered inter-stimulus interval ranging from 750 to milliseconds. The task consisted of 360 trials, including 24 1-back repetitions. The second task consisted of a total of 56 gray-scaled stimuli: faces, tools, patterns, houses, and body parts. During the task, each image exemplar was shown 3-4 times. The images were displayed for a duration of 250 milliseconds, at a constant frequency of 1 Hz. The task consisted of 205 trials, with 25 of them being 1-back repeats. The third task consisted of a total of 50 stimuli: 10 faces (open-source database[40]), tools (BOSS database[41]), patterns, houses, and body parts. Each image exemplar was presented 4–5 times throughout the task. The task was designed in a block form. Each block consisted of 10 images belonging to the same category, presented in a pseudo-random sequence (with each example displayed once per block, except for 1-back repeats). Images were displayed for 500 milliseconds, followed by a jittered inter-stimulus interval ranging from 500 to 750 milliseconds. The blocks were separated by either 4 or 8 s. The task comprised 260 trials (26 blocks) and 18 1-back repeats.

The patients were seated on a bed in front of an LCD monitor for all three versions of the task. During the tasks, stimuli were centrally presented, with a visual angle of ~13° in task 1 and ~11° in tasks 2 and 3.

Patients' accuracy was 85 ± 12%, 78 ± 20% and 94 ± 7% (ratio of hit responses to all 1-back repetitions in %, mean ± SD) in task 1, 2 and 3, respectively. Note that 1-back repetitions as well as false alarms were discarded from all analyses (see HFA Estimation and Signal Preprocessing below).

iEEG recordings

Recordings were conducted at North Shore University Hospital, Manhasset, NY, USA. Electrodes were either subdural strips or grids placed directly on the cortical surface (Ad-Tech Medical Instrument, Racine, Wisconsin) and/or depth electrodes (PMT Corporation, Chanhassen, Minnesota). For Ad-Tech, the subdural contacts were 3 mm in diameter and were spaced at 1 cm distances. For PMT, the depth contacts were 2 or 1 mm in diameter and spaced at 2.5 or 5 mm apart. The maximum penetration of depth electrodes was 70 mm, corresponding to a maximum of 13 implanted contacts. The signals were referenced to a vertex screw or a subdermal electrode, filtered electronically (analog bandpass filter with half-power boundaries at 0.07 and 40% of sampling rate), sampled at 512 or 500 Hz, and stored for offline analysis using XLTEK EMU128FS or NeuroLink IP 256 systems (Natus Medical Inc., San Carlos, CA). Electrical pulses were sent upon stimuli onsets and recorded along with the iEEG data for precise alignment of task protocol to neural activity.

Anatomical localization of iEEG contacts

Prior to the electrodes implant, patients were scanned with a T1-weighted 0.8 mm isometric anatomical MRI on a 3 Tesla Signa HDx scanner (GE Healthcare, Chicago, Illinois). Following the implant, a computed tomography (CT) and a T1-weighted anatomical MRI scan on a 1.5 Tesla Signa Excite scanner (GE Healthcare) were collected. The CT taken after implantation was overlaid on top of the post-surgery MRI scan, which was then adjusted to the MRI scan taken prior to the implantation surgery using FSL's Flirt. This alignment was then used to visualize the post-implant CT scans as overlayed on the pre-surgical MRI scans. The identification of each contact was then performed manually based on inspection of the CT aligned to the pre-surgery MRI (that is, in the patient's native space) using the BioImage Suite[42].

The localization of iEEG contacts onto the cortical surface was conducted following previous routines[43]. The cortical surface of each patient was reconstructed and segmented based on the pre-surgical MRI scan using the recon-all Freesurfer function (FreeSurfer 5.3). Each contact was then localized to the cortical vertex that was nearest to its estimated position. To project all patients' electrodes contacts onto a single cortical template, the cortical surface was unfolded and for each patient the unfolded spherical mesh of each patient was resampled and projected onto a standard mesh (FSaverage mesh template using SUMA[44]).

To compute pairwise anatomical distances between contacts from different patients (Fig. 5), we first extracted the 3D coordinate of the vertex to which each contact was assigned after projected onto the standard pial surface (FSaverage standardized pial surface of SUMA). We then computed the Euclidean distances between coordinate pairs.

HFA estimation and signal preprocessing

In this study, we analyzed the high-frequency amplitude (HFA, 48–154 Hz) signal, which has been found to be functionally selective[45,46] and to be a reliable indicator of aggregate firing rate in humans[8,10]. All signals were preprocessed to be 500 Hz, signals originally recorded at a sampling rate of 512 Hz were down-sampled. Raw time-courses and power spectra of all channels were manually inspected, and irregular channels were excluded from further analysis. Next, the signal was divided into nine sub-ranges of 10 Hz width to estimate high-frequency amplitude (HFA) modulations, ranging from 48–154 Hz. We excluded 59–61 and 117–121 Hz sub-ranges to discard line noise. We performed band-pass filtering on the signal at each frequency sub-range and the instantaneous amplitude in each sub-range was estimated by taking the absolute value of the filtered signal's Hilbert transform[47]. Each sub-range was divided by its mean value, and the normalized values were then averaged across all nine sub-ranges. This normalization was performed since the 1/f profile of the signal's power spectrum leads to a higher impact of lower frequencies on the overall estimation of high-frequency activity (HFA). Data preprocessing was carried out using Matlab (R2017a). For the frequency sub-range filtering, we used the original EEGLAB's Hamming windowed FIR filter (pop_eegfiltnew function).

For all subsequent analyses, visual responses were defined as the average HFA response at 50–350 msec following stimulus onset. These amplitudes were averaged across repetitions of the image. 1-back repetitions (hit and miss trials) and false alarm trials were discarded from all analyses.

Contacts' inclusion criteria

We defined visually responsive contacts as we reported in previous studies[24,39]. Briefly, the mean response of each contact to all available stimuli from the task versions the patient participated in was compared to baseline (paired t-test on mean exemplar responses versus baseline, at 50–500 ms and −200 to 0 ms relative to image onset, respectively). Hit, miss, and false alarm trials were excluded from all analyses. FDR correction was then applied to the pooled p-values from all 68 patients. Contacts with pFDR < 0.05 and a considerable effect size (Glass' Δ) of >1 were defined as visually responsive. Among all contacts implanted in 68 patients, 852 contacts were found to be visually responsive. For our analyses, we next considered only visual contacts that were localized in high-order visual areas, defined anatomically as the conjunction of labels comprising the ventro-temporal cortex (VTC). Anatomical labels were taken from the FreeSurfer parcellation of individual brains in their native space (prior to projection onto the common surface), based on the Destrieux anatomical atlas[48]. Finally, patients with less than 4 high-order visual contacts were excluded from the analyses. This resulted in a total of 19 patients and 140 contacts. Note that due to the within-task nature of our analyses, contacts recorded in two task versions were considered as separate in the analysis, each time with a different functional tunning curve, corresponding to the images of the relevant task. We therefore count 244 contacts from 19 patients across all the 3 sets (see Table S1).

Data analysis

Definitions. We denote by $\mathbf{a}_j^i$ the activation value of contact j when the participant was exposed to image i. Accordingly, the term $\mathbf{a}_j$* denotes

the contact values across all images, thus $\mathbf{a}_j^*$ represents the *tunning curve* of contact j whose dimensionality equals the number of images (see Fig. 1). Having a group of contacts, $\mathcal{A}$, from the same brain (and usually from a similar cortical area), we rearrange the values of these contacts when exposed to image i into an ***activation pattern*** $\mathbf{a}_{\mathcal{A}}^i$ where the dimensionality of $\mathbf{a}_{\mathcal{A}}^i$ equals the number of contacts in the group $\mathcal{A}$. When analyzing the invariant characteristics of activations from multiple brains, we compare all possible pairs of contacts, each of each from a different brain.

Permutation inter-subject matching. As typically is the case with intracranial recordings, the number and location of contacts is different among subjects. The first step, referred to as permutation, is thus intended to find a mapping between the contacts of each pair of subjects. Due to the limitations of anatomical distances (see Fig. 5), we opt for a cross-subject mapping that optimizes activation correlation.

Assume subjects A and B with contact $\mathcal{A}$ and $\mathcal{B}$, respectively, where $|\mathcal{A}| = n_a$ and $|\mathcal{B}| = n_b$. Without loss of generality, we assume $n_a \geq n_b$. In order to find the optimal contact match between the subjects, we construct a bipartite graph $G = (\underline{\mathcal{A}}, \underline{\mathcal{B}}, E)$. Each node $i \in \underline{\mathcal{A}}$ represents a contact in subject A, and each node $j \in \mathcal{B}$ represents a contact in subject B.

Each edge $e_{i,j} \in E$ from node $i \in \mathcal{A}$ to node $j \in \mathcal{B}$ is weighted by the Spearman correlation between the tunning curves $\mathbf{a}_i^*$ and $\mathbf{a}_j^*$: $e_{i,j} = S\left(\mathbf{a}_i^*, \mathbf{a}_j^*\right)$. In order to find the best matching between the contacts we use the *max-weight matching* algorithm in a bipartite graph[49] which finds the optimal match between the contacts, such that the total Spearman correlations are maximized. Since $n_a \geq n_b$, the $n_b - n_a$ non-matched contacts in $\mathcal{B}$ are discarded. Having an optimal matching between contacts in $\mathcal{A}$ and $\mathcal{B}$, we rearrange the n_a contacts in $\mathcal{A}$ and in $\mathcal{B}$ in the same order. This means, that for each image k, the *activation patterns*, $\mathbf{a}_{\mathcal{A}}^k$ and $\mathbf{a}_{\mathcal{B}}^k$ are both vectors in $\mathbb{R}^{n_a}$ (see Fig. 1) and where $\forall i$, $\mathbf{a}_{\mathcal{A}}^k[i]$ was matched to $\mathbf{a}_{\mathcal{B}}^k[i]$. This procedure is applied to each pair of subjects.

Relational coding: Inter-subject Rotational Invariance. The second hypothesis we consider is that of *relational coding*. Formally, a *relational coding scheme* implies that the activation codes are invariant across individuals up to orthogonal transformations, namely, for each pair of subjects, $\mathcal{A}$ and $\mathcal{B}$, there is an orthogonal transformation R, such that, for each image i, $R\mathbf{a}_{\mathcal{A}}^i \approx \mathbf{a}_{\mathcal{B}}^i$.

First, for each category of visual stimuli and each pair of patients we applied the optimal-match procedure as discussed above. Next, we computed the optimal orthogonal transformation for each pair of patients on the training set, and then applied the evaluated rotation to the test set. The optimal rotation, $\hat{R}$, was evaluated by solving:

$$\hat{R} = argmin_R \sum_{i \in train\_set} ||R\mathbf{a}_{\mathcal{A}}^i - \mathbf{a}_{\mathcal{B}}^i||_2 \tag{11}$$

This optimization is known as *Orthogonal Procrustes Problem*[50,51] where the optimization is performed in the following manner: we construct two matrices from the activation patterns:

$$\mathbf{a}_{\mathcal{A}}^* = \left[\mathbf{a}_{\mathcal{A}}^1, \mathbf{a}_{\mathcal{A}}^2, \cdots, \mathbf{a}_{\mathcal{A}}^n\right] \text{and } \mathbf{a}_{\mathcal{B}}^* = \left[\mathbf{a}_{\mathcal{B}}^1, \mathbf{a}_{\mathcal{B}}^2, \cdots, \mathbf{a}_{\mathcal{B}}^n\right].$$

It can be shown that the problem is equivalent to finding the nearest orthogonal matrix to the matrix $E = \left(\mathbf{a}_{\mathcal{B}}^*\right)\left(\mathbf{a}_{\mathcal{A}}^*\right)^T$ which means solving the following problem:

$$\hat{R} = \min_R ||R - E||_F \; s.t. \; R^T R = I \tag{12}$$

The solution can be calculated using the SVD decomposition of $E: E = U\Sigma V^T$ and $\hat{R} = UV^T$

Linear coding: Inter-subject linear Invariance

Finally, in order to test a more general coding hypothesis, we find the general linear transformation. As in the Relational Coding, we computed the optimal linear transformation for each pair of patients on the training set, and then applied the evaluated transformation to the test set. The optimal linear transformation, $\hat{M}$, was evaluated by minimizing:

$$\hat{M} = argmin_M \sum_{i \in train\_set} ||M\mathbf{a}_{\mathcal{A}}^i - \mathbf{a}_{\mathcal{B}}^i||_2 \tag{13}$$

Where M is a general linear transformation matrix. Using the above notations for $\mathbf{a}_{\mathcal{A}}^*$ and $\mathbf{a}_{\mathcal{B}}^*$, $\boldsymbol{M}$ can be estimated using Moore–Penrose pseudo inverse:

$$\hat{\boldsymbol{M}} = \mathbf{a}_{\mathcal{B}}^*\left(\mathbf{a}_{\mathcal{A}}^*\right)^{\boldsymbol{T}}\left[\mathbf{a}_{\mathcal{A}}^*\left(\mathbf{a}_{\mathcal{A}}^*\right)^{\boldsymbol{T}}\right]^{-\mathbf{1}} \tag{14}$$

Reporting summary

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

Data availability

Source data for all figures are provided with this paper. iEEG preprocessed data files can be found here: https://github.com/oferlip/Invariant-Inter-Subject-Relational-Structures-in-the-Human-Visual/tree/main/data. Source data are provided with this paper.

Code availability

The code used for data analysis is available in https://github.com/oferlip/Invariant-Inter-Subject-Relational-Structures-in-the-Human-Visual.

References

1. Schlaffke, L. et al. The brain's dress code: how the dress allows to decode the neuronal pathway of an optical illusion. *Cortex* **73**, 271–275 (2015).
2. Grill-Spector, K. et al. Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* **24**, 187–203 (1999).
3. Khuvis, S. et al. Face-selective units in human ventral temporal cortex reactivate during free recall. *J. Neurosci.* **41**, 3386–3399 (2021).
4. Haxby, J. V. et al. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science (1979)* **293**, 2425–2430 (2001).
5. Edelman, S. Representation is representation of similarities. *Behav. Brain Sci.* **21**, 449–467 (1998).
6. Malach, R. Local neuronal relational structures underlying the contents of human conscious experience. *Neurosci. Conscious.* **2021**, niab028 (2021).
7. Edelman, S., Grill-Spector, K., Kushnir, T. & Malach, R. Toward direct visualization of the internal shape representation space by fMRI. *Psychobiology* **26**, 309–321 (1998).
8. Nir, Y. et al. Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI Is related to interneuronal correlations. *Curr. Biol.* **17**, 1275–1285 (2007).
9. Mukamel, R. et al. Neuroscience: Coupling between neuronal firing, field potentials, and fMRI in human auditory cortex. *Science (1979)* **309**, 951–954 (2005).
10. Manning, J. R., Jacobs, J., Fried, I. & Kahana, M. J. Broadband shifts in local field potential power spectra are correlated with single-neuron spiking in humans. *J. Neurosci.* **29**, 13613–13620 (2009).

11. Georgopoulos, A. P., Schwartz, A. B. & Kettner, R. E. Neuronal population coding of movement direction. *Science (1979)* **233**, 1416–1419 (1986).
12. Pouget, A., Dayan, P. & Zemel, R. Information processing with population codes. *Nat. Rev. Neurosci.* **1**, 125–132 (2000).
13. Broday-Dvir, R., Norman, Y., Harel, M., Mehta, A. D. & Malach, R. Perceptual stability reflected in neuronal pattern similarities in human visual cortex. *Cell Rep.* **42**, 112614 (2023).
14. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *Science (1979)* **303**, 1634–1640 (2004).
15. Golland, Y. et al. Extrinsic and intrinsic systems in the posterior cortex of the human brain revealed during natural sensory stimulation. *Cerebral Cortex* **17**, 766–777 (2007).
16. Parvizi, J. & Kastner, S. Promises and limitations of human intracranial electroencephalography. *Nat. Neurosci.* **21**, 474–483 (2018).
17. Norman, Y. & Malach, R. What Can iEEG inform us about mechanisms of spontaneous behavior?. In *Intracranial EEG. Studies in Neuroscience, Psychology and Behavioral Economics?*. (ed. Axmacher, N.) (Springer, Cham, 2023).
18. Axmacher, N. *Intracranial EEG: A Guide for Cognitive Neuroscientists*. (Springer, 2023).
19. Vishne, G., Gerber, E. M., Knight, R. T. & Deouell, L. Y. Representation of sustained visual experience by time-invariant distributed neural patterns. *bioRxiv* **42**, 112752 (2023).
20. Ingram, D. Ferdinand de Saussure and linguistic structuralism. In *Critical Theory to Structuralism* (ed. Ingram, D.) 300 (Routledge, 2020).
21. Haxby, J. V., Guntupalli, J. S., Nastase, S. A. & Feilong, M. Hyperalignment: Modeling shared information encoded in idiosyncratic cortical topographies. *eLife* **9**, e56601 (2020).
22. Haxby, J. V. et al. A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron* **72**, 404–416 (2011).
23. Tanimoto, S. L., Itai, A. & Rodeh, M. Some matching problems for bipartite graphs. *J. ACM (JACM)* **25**, 517–525 (1978).
24. Grossman, S. et al. Convergent evolution of face spaces across human face-selective neuronal groups and deep convolutional networks. *Nat. Commun.* **10**, 4934 (2019).
25. Yamins, D. L. K. & DiCarlo, J. J. Using goal-driven deep learning models to understand sensory cortex. *Nat. Neurosci.* **19**, 356–365 (2016).
26. Kriegeskorte, N. Representational similarity analysis - connecting the branches of systems neuroscience. *Front Syst. Neurosci.* **2**, 4 (2008).
27. Kornblith, S., Norouzi, M., Lee, H. & Hinton, G. Similarity of neural network representations revisited. *arXiv* https://doi.org/10.48550/arXiv.1905.00414 (2019).
28. Davidesco, I. et al. Exemplar selectivity reflects perceptual similarities in the human fusiform cortex. *Cerebral Cortex* **24**, 1879–1893 (2014).
29. Mormann, F. et al. Neurons in the human amygdala encode face identity, but not gaze direction. *Nat. Neurosci.* **18**, 1568–1570 (2015).
30. Lau, H., Michel, M., LeDoux, J. E. & Fleming, S. M. The mnemonic basis of subjective experience. *Nat. Rev. Psychol.* **1**, 479–488 (2022).
31. Norman, Y., Raccah, O., Liu, S., Parvizi, J. & Malach, R. Hippocampal ripples and their coordinated dialogue with the default mode network during recent and remote recollection. *Neuron* **109**, 2767–2780.e5 (2021).
32. Bothwell, R. K., Brigham, J. C. & Malpass, R. S. Cross-racial identification. *Pers. Soc. Psychol. Bull.* **15**, 19–25 (1989).
33. Hasson, U., Nir, Y., Levy, I. & Fuhrmann, G. Intersubject synchronization of cortical activity during natural vision. *Science (1979)* **303**, 1634–1640 (2004).
34. Khaligh-Razavi, S. M. & Kriegeskorte, N. Deep supervised, but not unsupervised, models may explain IT cortical representation. *PLoS Comput. Biol.* **10**, e1003915 (2014).
35. Yamins, D. L. K. et al. Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc. Natl. Acad. Sci. USA* **111**, 8619–8624 (2014).
36. Simony, E., Grossman, S. & Malach, R. Brain-machine convergent evolution: Why finding parallels between brain and artificial systems is informative. *Proc. Natl. Acad. Sci. USA* **121**, e2319709121 (2024).
37. Deitch, D., Rubin, A. & Ziv, Y. Representational drift in the mouse visual cortex. *Curr. Biol.* **31**, 4327–4339.e6 (2021).
38. Malach, R. Special issue: consciousness science and its theories Local neuronal relational structures underlying the contents of human conscious experience. *Neurosci. Conscious.* **2021**, 1–13 (2021).
39. Grossman, S. et al. The noisy brain: power of resting-state fluctuations predicts individual recognition performance. *Cell Rep.* **29**, 3775–3784.e4 (2019).
40. Minear, M. & Park, D. C. A lifespan database of adult facial stimuli. *Behav. Res. Meth. Instrum. Comput.* **36**, 630–633 (2004).
41. Brodeur, M. B., Guérard, K. & Bouras, M. Bank of Standardized Stimuli (BOSS) phase ii: 930 new normative photos. *PLoS One* **9**, e106953 (2014).
42. Papademetris, X. et al. Bioimage suite: an integrated medical image analysis suite: an update. *Insight J.* **2006**, 209 (2006).
43. Golan, T. et al. Human intracranial recordings link suppressed transients rather than 'filling-in' to perceptual continuity across blinks. *Elife* **5**, e17243 (2016).
44. Argall, B. D., Saad, Z. S. & Beauchamp, M. S. Simplified intersubject averaging on the cortical surface using SUMA. *Hum. Brain Mapp.* **27**, 14–27 (2006).
45. Jacques, C. et al. Corresponding ECoG and fMRI category-selective signals in human ventral temporal cortex. *Neuropsychologia* **83**, 14–28 (2016).
46. Privman, E. et al. Enhanced category tuning revealed by intracranial electroencephalograms in high-order human visual areas. *J. Neurosci.* **27**, 6234–6242 (2007).
47. Fisch, L. et al. Neural 'ignition': enhanced activation linked to perceptual awareness in human ventral stream visual cortex. *Neuron* **64**, 562–574 (2009).
48. 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).
49. Schrijver, A. & others. *Combinatorial Optimization: Polyhedra and Efficiency*. Vol. 24 (Springer, 2003).
50. Arun, K. S., Huang, T. S. & Blostein, S. D. Least-squares fitting of two 3-D point sets. *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-9**, 698–700 (1987).
51. Ten Berge, J. M. F. Procrustes problems. New York: Oxford University Press. *Psychometrika* **70**, 799–801 (2005).

Acknowledgements

This work was supported by the Israeli Ministry of Science and Technology under The National Foundation for Applied Science (MIA) (Y.H.); The Horizon 344 2020 FET Proactive project GuestXR (101017884) (D.F.); The Federal Ministry of Education and Research (BMBF) and the Free and Hanseatic City of Hamburg under the Excellence Strategy of the Federal Government and the Länder (S.G.); The CIFAR BMC grant (R.M.).

Author contributions

O.L. contributed to writing, data analyses design, and software development S.G. contributed to writing, data collection preprocessing and analyses. D.F. contributed to writing, overall experimental questions, and discussions. Y.H. contributed to modeling, experimental question overseeing data analyses, and writing. R.M. contributed to the overall design of the experiment, data interpretations, and writing.

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/s41467-025-62551-x.

Correspondence and requests for materials should be addressed to Rafael Malach.

Peer review information *Nature Communications* thanks Gabriel Kreiman, and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.