Supplementary Material

Inter-subject similarity for individual versus collapsed episodes

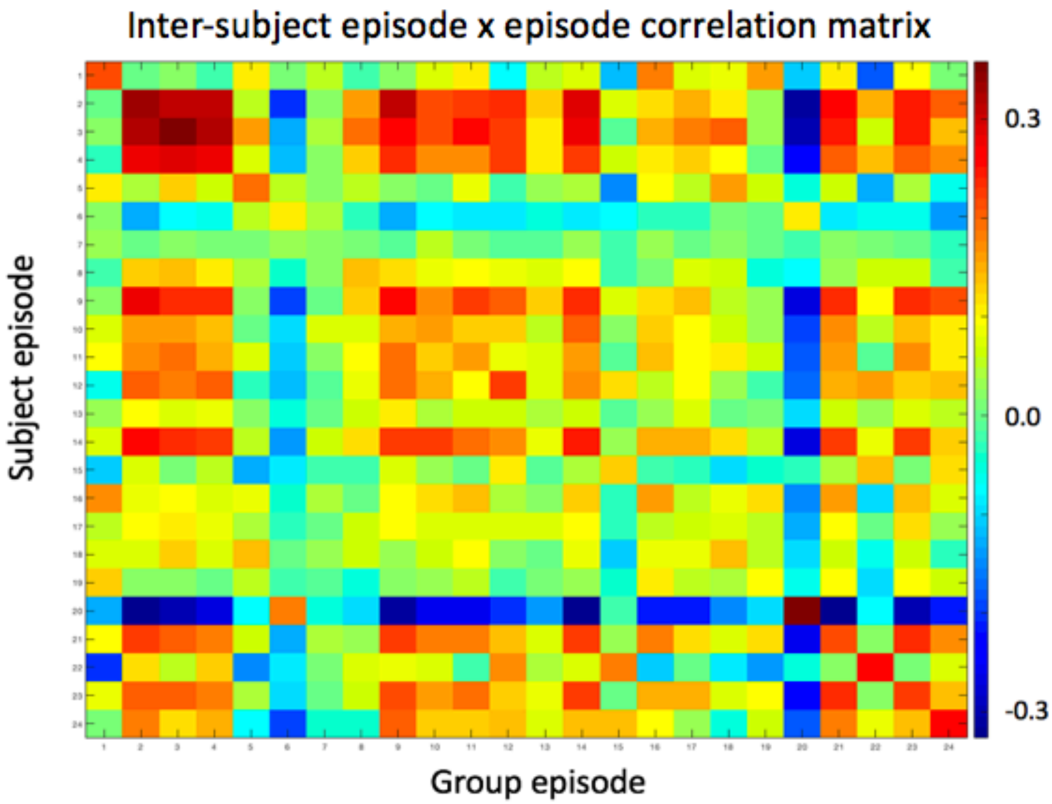
Results in the main manuscript include inter-subject similarity values based on patterns collected as subjects observed the same episode, as well as patterns that were collapsed across episodes. We take a moment here to note that although having higher Fisher-*Z* *r*-values for collapsed-, compared to same-, episode patterns within primary visual areas might be unexpected at first, the collapsing process acts to remove visual idiosyncrasies that are typically present in a small number of time-points. This leaves a reliable and coarse signal for the visual commonalities, which early visual cortex is particularly responsive to. For instance, the background and environment across the four episodes featuring the mudskipper are all relatively similar: a dark muddy and sandy beach below the horizon, and a prominent dark fish in the center. Averaging the patterns across episodes does not eliminate (or scramble) all visual information (only time-point by time-point idiosyncrasies). The figure below illustrates this, which shows an average across all four mudskipper episodes, revealing strong visual commonalities.

These basic commonalities can activate visual cortex in a very predictable and reliable fashion across people due to common coarse-level retinotopy, giving *r*-values that are higher than are observed for more idiosyncratic patterns. This also explains why the collapsed values are particularly high in early visual cortex, which has a stable and predictable organization across people (e.g., voxels that respond to luminance in the top/bottom and right/left of the visual field are in similar locations across people).

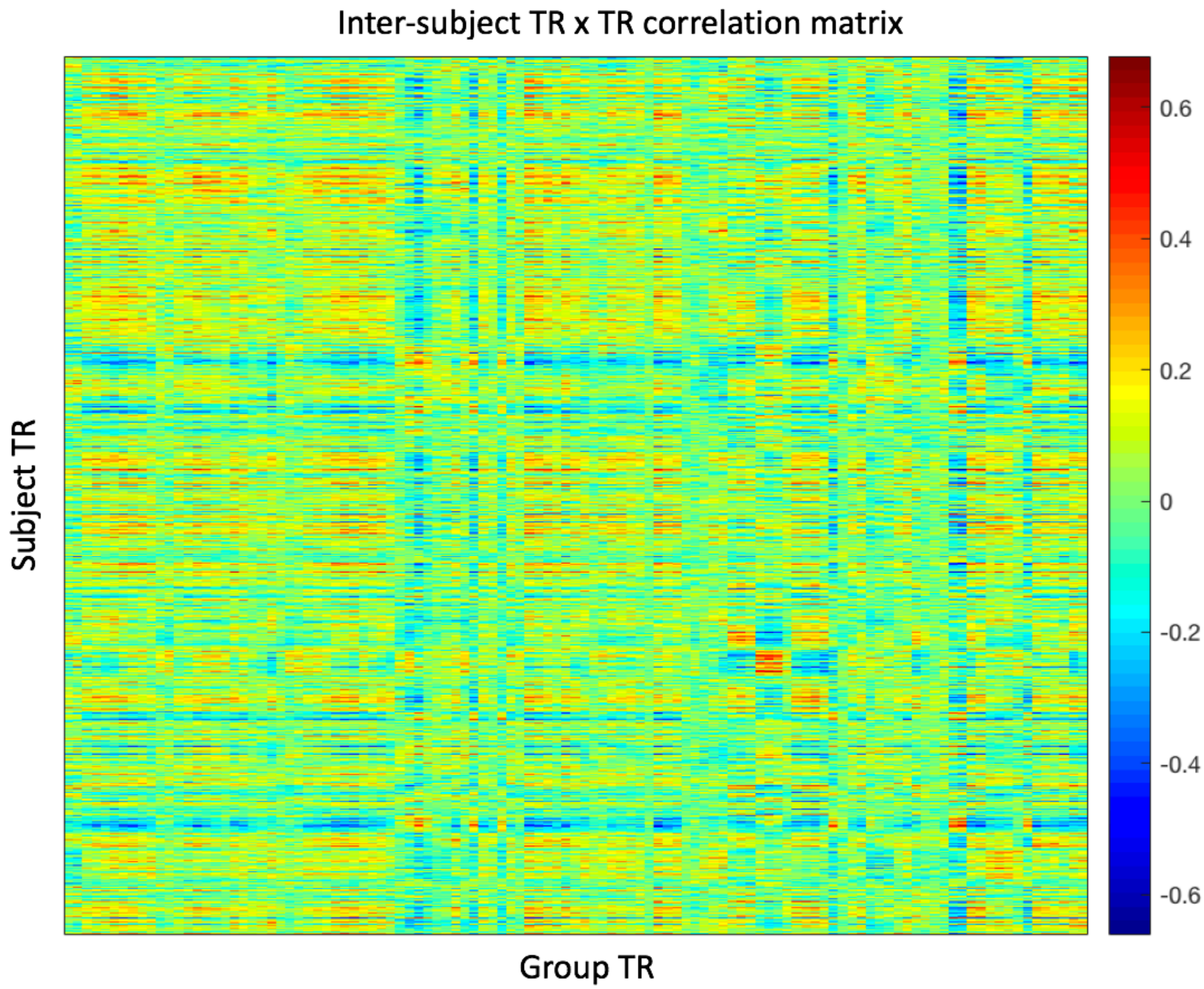
To test this idea further, we predicted that collapsing across all episodes seen (not just separately collapsing across retrieved and then not-retrieved episodes) would give greater *r*-values than observed for the retrieved and not-retrieved episodes separately, because of the even coarser structure that can be extracted from more time-points. We tested this hypothesis with the same procedure used to create Figure 4, for all episodes seen. The resulting inter-subject average Fisher-*Z* *r*-value was indeed higher ($r = 0.470$) within the primary visual ROI, indicating that collapsing across time-points left a coarse and common pattern across participants.

Episode- and TR-based similarity matrices

Patterns for each episode were correlated between each subject and the group's average pattern (consisting of all *other* subjects). The resulting Fisher-*Z* *r*-values are displayed below. Note that each value reflects a correlation between each subject and the rest of the group *without that subject* (so all values reflect inter-subject similarity). The diagonal values (same episode between subjects and the rest of the group) are significantly greater than the off-diagonal values (in a two-sample t-test: $t(574) = 5.81$, $p < .001$; same-episode: $M = 0.19$, $SD = 0.09$; different-episode: $M = 0.05$, $SD = 0.12$). None of the 24 same-episode Fisher-*Z* *r*-values were less than 0, whereas 27% of the 552 different-episode values were less than 0.

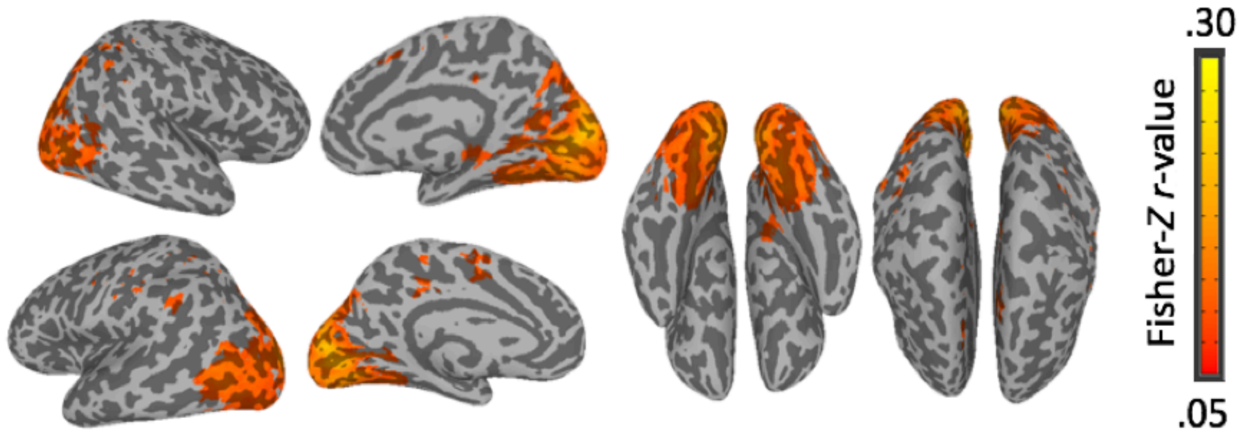


We extended this by correlating the individual TRs that comprise the 24 episodes. The figure below shows the result of this analysis, where the pattern of each subject's set of included TRs are correlated with the average pattern for TRs from the rest of the group without this subject. As in the episode-by-episode correlation matrix, the TR-by-TR correlation matrix depicts significantly stronger similarity for the same-TRs compared to the different-TRs across-subjects ($t(12319) = 13.96$, $p < .001$; same-episode: $M = 0.12$, $SD = 0.07$; different-episode: $M = 0.03$, $SD = 0.07$). Only 2% of the 111 same-TR Fisher-Z r-values were less than 0, whereas 30% of the 12210 different-TR values were less than 0.



Searchlight analysis of inter-subject perceptual similarity

To compare our data to prior results, we conducted a 3-voxel radius searchlight to obtain a whole-brain map of inter-subject pattern similarity, to mirror the inter-subject correlations of representational geometry presented by Guntupalli and colleagues (2016, *Figure 3*). Below, we present our averaged Fisher-*Z* *r*-values for the same-episode inter-subject correlations (without considering memory outcomes). Similar to Guntupalli and colleagues' results, we find that the peak inter-subject correlations arise within visual areas, and are less prominent throughout other cortical regions.



The magnitude of our correlations were slightly lower than the values reported by Guntupalli et al., (2016), though we believe this is likely because of differences in the movie durations – we used episodes within 45 second video clips, compared to the hour-long movie segments used by Guntupalli et al. (2016). Additionally, the Guntupalli et al. (2016) data benefits from having a more consistent theme and narrative structure from the movie, compared to our inter-leaved video clips of six different animals in numerous locations and environments.

Supplementary Material-Only Reference

Guntupalli, J. S., Hanke, M., Halchenko, Y. O., Connolly, A. C., Ramadge, P. J., & Haxby, J. V. (2016). A Model of Representational Spaces in Human Cortex. *Cerebral Cortex (New York, NY)*, *26*(6), 2919–2934. https://doi.org/10.1093/cercor/bhw068