

# Decoding the Brain's Algorithm for Categorization from Its Neural Implementation

Michael L. Mack,<sup>1,\*</sup> Alison R. Preston,<sup>1,3</sup> and Bradley C. Love<sup>2,3</sup>

<sup>1</sup>Center for Learning and Memory and Department of Psychology, The University of Texas at Austin, 1 University Station C7000, Austin, TX 78712-0805, USA

<sup>2</sup>Cognitive, Perceptual, and Brain Sciences, University College London, 26 Bedford Way, Room 235, London WC1H 0AP, UK

## Summary

Acts of cognition can be described at different levels of analysis: what behavior should characterize the act, what algorithms and representations underlie the behavior, and how the algorithms are physically realized in neural activity [1]. Theories that bridge levels of analysis offer more complete explanations by leveraging the constraints present at each level [2–4]. Despite the great potential for theoretical advances, few studies of cognition bridge levels of analysis. For example, formal cognitive models of category decisions accurately predict human decision making [5, 6], but whether model algorithms and representations supporting category decisions are consistent with underlying neural implementation remains unknown. This uncertainty is largely due to the hurdle of forging links between theory and brain [7–9]. Here, we tackle this critical problem by using brain response to characterize the nature of mental computations that support category decisions to evaluate two dominant, and opposing, models of categorization. We found that brain states during category decisions were significantly more consistent with latent model representations from exemplar [5] rather than prototype theory [10, 11]. Representations of individual experiences, not the abstraction of experiences, are critical for category decision making. Holding models accountable for behavior and neural implementation provides a means for advancing more complete descriptions of the algorithms of cognition.

## Results

A fundamental and long-standing debate in category learning is whether knowledge is based on representations of individual instances of category members, known as exemplar theory [5, 12], or an abstracted representation coding a category's prototypical features, known as prototype theory [10, 11]. Over thirty years of debate in behavioral and modeling research has yet to resolve which of these theories best describes how people represent category knowledge [13, 14].

We applied a novel approach to neuroimaging analysis to inform the debate between exemplar and prototype theories and guide neuroscientific study of how categorization occurs in the brain. Participants ( $n = 20$ ) performed a classic task from the exemplar and prototype theory literature [12] that involved learning to categorize objects (Figure 1A). Exemplar

and prototype models were fit to each participant's learning behavior collected prior to scanning (see Supplemental Information available online). Consistent with previous work [13, 14], computational models of both theories provided accurate accounts of individual participants' behavioral responses during a scanned test phase (Figure 1B), with only a single participant better fit by either model (Figure 1C;  $\chi^2 = 5.34$ ,  $p = 0.021$ ).

Despite their equivalent behavioral predictions, the underlying representations driving exemplar and prototype model category decisions are fundamentally opposed. We captured these model state differences with *representational match*, a measure of summed similarity between a test object and a model's stored category representations. We chose representational match as the latent model signature of interest for three reasons. First, representational match summarizes critical computations in the categorization process for the exemplar and prototype models (see Supplemental Information). Second, representational match is strictly tied to the model parameters optimized for categorization and characterizes the attention and decision processes necessary for the categorization decisions proposed by the two models. If the models are accurately characterizing categorization, then evidence of their mechanisms should be found in brain response. Third, representational match is a latent signature that teases apart the two models. Although the same summed similarity calculation is used for both models, the internal representations to which a test object is compared are vastly different, leading to different representational match functions (Figure 2A). The two models might predict the identical behavioral response on any given trial, but the latent representations that support that decision would be very different. By relating brain patterns to these latent model signatures, we can determine which model conception accurately reflects the nature of representations in this task.

We examined the consistency between brain response and representational match from the exemplar and prototype models by conducting whole-brain multivariate pattern analysis (MVPA [15]) to predict representational match for each test object. Representational match predictions for each participant were derived for both models from parameters optimized for behavioral fits. The mutual information (MI) between cross-validated MVPA output and model representational match served as an index of the consistency between brain states and latent model signatures during category decisions. MI measures the amount of shared information between brain and model. In this context, higher MI reflects a richer mapping between the category representations supporting a model and the underlying information in patterns of brain response during category decisions.

To first validate our approach, we performed a model recovery study by simulating voxel activations corresponding to exemplar or prototype-based representations and tested whether or not our technique could identify the cognitive model from the simulated neural activity. We evaluated MVPA output of activation patterns from 5,000 simulated voxels with MI. In each simulation, 5% of the simulated voxels were assigned activation profiles corresponding to the

<sup>3</sup>These authors contributed equally to this work

\*Correspondence: [michael.mack@utexas.edu](mailto:michael.mack@utexas.edu)



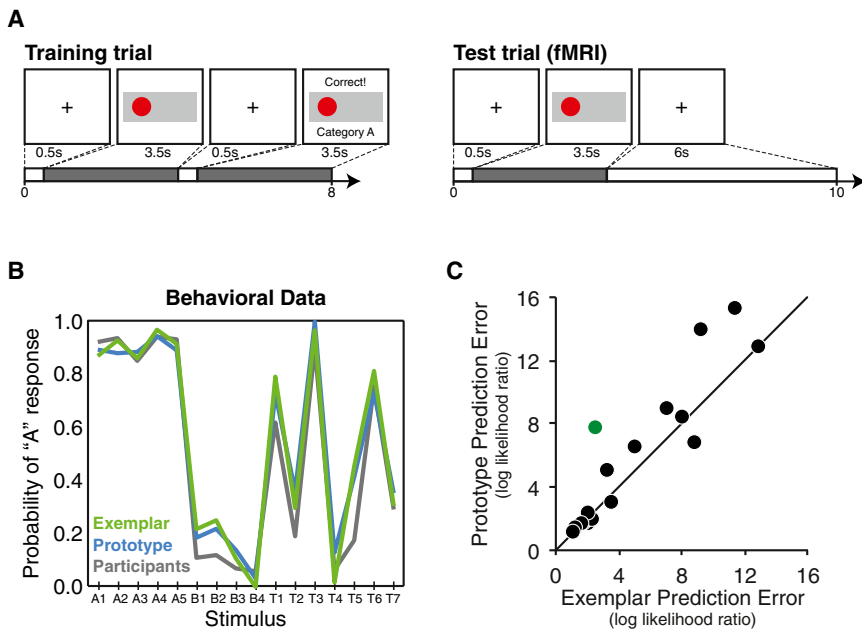


Figure 1. Learning and Testing Trial Schematics and Behavioral Modeling Results

(A) Participants were first trained outside of the scanner to categorize nine training objects (five category A and four category B members; 20 repetitions of each stimulus; see Table S1) with corrective feedback. During a scanned test phase, participants were tested on the nine training objects and seven novel objects (18 repetitions of each stimulus; no feedback).

(B) For each training stimulus (A1–A5, B1–B4) and testing stimulus (T1–T7), mean probability of participants' "A" category responses during the testing phase (gray) and mean predicted responses from exemplar (green) and prototype (blue) models fit to each participant.

(C) Prediction errors from both models for each participant showed that only one participant was better fit by either model (green circle).

representations of either the exemplar or prototype model. The remaining voxels' activation profiles were modeled as Gaussian noise (see Supplemental Information). MVPA was performed on the simulated data sets to recover the embedded representational match signatures of the two models. The simulation results showed successful recovery with higher MI between the exemplar model and simulated exemplar voxels (mean MI = 3.06) than the simulated prototype voxels (mean MI = 0.01). Similar results were found for the prototype model and the simulated prototype (mean MI = 2.76) versus exemplar voxels (mean MI = 0.07). Consistent results were found in simulations that varied the proportion (1%–25%) of simulated voxels with activation profiles from model representations. Notably, despite the relatively higher model complexity and variability in the latent model measure for the exemplar compared to the prototype model, the prototype model was still preferred when it was the model generating the neural activity. These simulation results suggest our approach can identify cognitive models that are reflected in coherent brain states and favors models that share rich relationships, i.e., more MI, with brain response.

In applying this approach to our experimental data, we found that the exemplar model was more consistent with brain response than the prototype model across all participants, as measured with MI ( $t_{19} = 4.39$ ,  $p = 0.0003$ ; Figure 2B) and correlation (mean  $r_{GCM}$ : 0.67, mean  $r_{MPM}$ : 0.46,  $t_{19} = 3.24$ ,  $p = 0.004$ ). The greater consistency for the exemplar model was not due to greater complexity of the exemplar model relative to the prototype model, as the exemplar model was also more consistent with brain response than an overparameterized saturated model (see Supplemental Information and Figure S1). A region of interest analysis (Figure S2 and Table S1) showed converging evidence in that activation patterns in lateral occipital and posterior parietal cortex were more consistent with the predictions of exemplar than prototype theory. Critically, using brain response for model selection greatly increased the sensitivity of analysis at the individual participant level beyond that offered from only behavioral measures. Indeed, 13 participants showed greater brain-

model consistency with the exemplar model and one participant with the prototype model (Figure 2C). The remaining six participants had brain states equally consistent with both models.

Using brain response for model selection suggests that category decisions are supported by representations of category exemplars. A critical mechanism of exemplar theory is selective attention to diagnostic information [5]. It is hypothesized that this attention mechanism biases exemplar representations toward information most relevant for category decisions. To identify brain regions consistent with attention-biased category exemplars, we performed a representational similarity analysis (RSA) [16]. We targeted brain regions with neural patterns for category exemplars that matched the attention-weighted pairwise similarity structure predicted by the exemplar model. For each participant, exemplar model dissimilarity matrices (DMs) were derived from pairwise distances between stimuli taking into account the fitted attention weight parameters. Neural DMs were derived for each participant from pairwise correlation distances between neural patterns of the stimuli. The searchlight mapping technique [17] was used to compare exemplar model similarity to neural similarity in localized regions throughout the whole brain. The maps from group analysis (Figure 3A) showed significant correlation between exemplar model and neural similarity in lateral occipital cortex (LO), inferior parietal cortex, inferior frontal gyrus (IFG), and insular cortex. In contrast, significant correlation between neural similarity and DMs derived from the physical similarity of training stimuli, a condition akin to the exemplar model with no selective attention, was restricted to primary visual cortex (Figure 3B).

## Discussion

Recent efforts linking algorithmic models to brain measures have offered a sophisticated advance in analyzing fMRI data. These approaches correlate measures of model processes with fMRI data to identify brain regions whose activation profiles during a cognitive task are predicted by the model's algorithms [18]. These model-based fMRI studies have informed our understanding of the role of individual brain regions in specific computations [19–21], but the vast majority of neuroimaging studies are conducted through the lens of

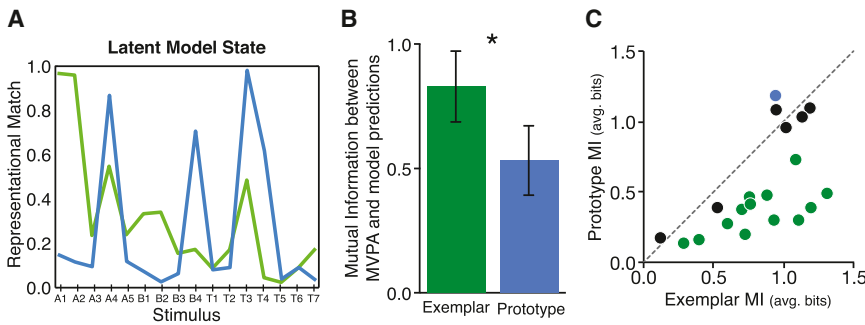


Figure 2. The Consistency between Latent Model States and Brain States

(A) The internal model measure representational match, the extent a test object activates stored category representations, varies between exemplar (green) and prototype (blue) models, offering a discriminatory trial-by-trial measure of the underlying categorization process.

(B) The mutual information (MI) between responses of brain patterns as revealed by multivariate pattern analysis (MVPA) and representational match was higher for the exemplar relative to the prototype model (mean  $\pm$  95% confidence interval). The exemplar model also outperformed an overparameterized saturated

model and a model with no attention mechanism (Figure S1). Similar results were found in ROI-based analyses (Figure S2). (C) Using brain response for model selection greatly increased the sensitivity of analysis beyond that offered from only behavioral measures with thirteen participants better fit by the exemplar model, one better fit by the prototype model, and six fit equally well by both models.

only one cognitive model [22]. This narrow focus assumes the studied model is an accurate description of the underlying cognitive phenomena, thereby limiting theoretical interpretation. In studies that do compare models [18, 20], evaluating the link between model processes and neural response has been limited to correlations with independent voxel activations. These model-based analyses are confirmatory in nature and are not ideally suited for model comparison.

Most domains of inquiry in human cognition involve intense debate and competing models. We have focused on one of these fundamental debates in the current study: are the representations underlying category decisions more concrete or abstract? By bridging levels of analysis with the proposed novel method, we can determine whether the radically different algorithmic-level assumptions of exemplar and prototype models are more consistent with the underlying neural patterns evoked during category decisions, thus providing strong evidence toward resolving this long-standing debate. Algorithmic models should be favored to the extent that continuous measures of the model's state (e.g., activation of internal representations) during categorization is tracked by trial-by-trial measures of brain states revealed with pattern information analysis [15, 23]. By relating patterns of activation over multiple voxels to internal measures of competing algorithmic models, we have used brain activation measures to determine which categorization algorithm the brain uses and how that algorithm is implemented in the brain.

We contend that this direction of analysis, using brain states to infer model states, is the best method to isolate the brain's algorithms. Many different brain states can manifest as a single mental state, but different mental states necessitate different underlying brain states. Multivariate techniques of fMRI analysis that leverage this analysis direction [15] have led to breakthroughs in understanding low-level stimulus [24–26] and semantic [23, 27] representations. Our approach extends beyond these existing techniques by marrying the advance of multivariate methods with model-based predictions of brain response to fill in the critical gap of linking theory, brain, and behavior and quantitatively determine the degree of correspondence across levels of analysis. In applying this approach to category learning, we found that brain response during categorization decisions is most consistent with the specific computations and representations posited by exemplar theory.

Unlike existing methods that attempt to describe isolated voxel activation, favoring models that are correlated with

more voxels [18], our novel method favors those models that are predictable by coherent brain states regardless of whether those states are found in activation patterns that reach across the extent of the brain or are localized to specific regions. Of course, this method shares the limitations of any model selection approach, namely that selection is respective to a particular set of tasks, data, and models [28, 29]. In the current task and category structure, which has been used as a benchmark for models of categorization [6, 12–14], we found evidence strongly in favor of the exemplar model over the prototype model. While this classic debate that has raged for over thirty years cannot be resolved with one study, we provide strong neural evidence for the role of exemplar representations in category decisions. Applying this approach of searching for the model in the categorizing brain to other tasks and category structures will lead to a more complete theory of categorization.

We were able to identify a network of brain regions tracking the representations of the exemplar model with the follow-up RSA. Specifically, we were able to target exemplar theory's mechanism of selective attention, finding attention-biased exemplar representations in LO and posterior parietal cortex. LO is known to be critical for representation of objects [30], notably showing effects of representational changes as a result of category learning [31–33]. That activation patterns in LO correlate with exemplar representations builds on these findings to suggest that LO plays a critical role in representing individual experiences of category members and that these LO-based representations are brought to bear during category decisions.

Processing within posterior parietal cortex has been implicated in goal-directed, top-down selection of relevant information [34], episodic memory retrieval [35, 36], and experience-dependent representational changes during perceptual learning [32, 37]. Although parietal activity related to category learning has been observed previously [38–40], the specific mechanistic role of parietal regions in category decisions has been underexplored. The correlations found here between parietal areas and attention-biased exemplar representations offer a novel step in characterizing the parietal cortex's computational role. Successful learning in the current task depends on learning the appropriate attention weights to stimulus dimensions. Our findings suggest parietal cortex plays a critical role in the attentional processes required for encoding and retrieval of stored exemplar representations.

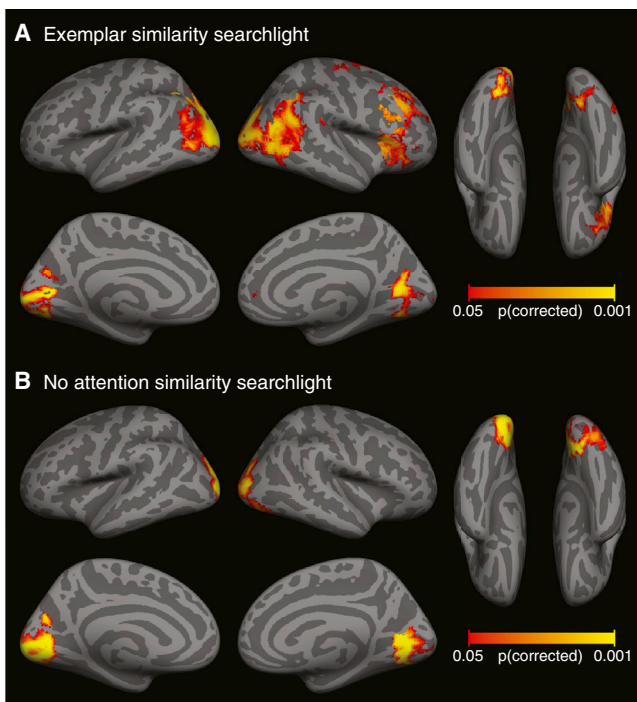


Figure 3. Representational Similarity Analysis of Object Similarity

Searchlight analyses of representational similarity compared the similarity between neural representations of the objects with the pairwise similarities as predicted by the exemplar model. Correspondence between the attention-biased exemplar representations and neural representations extended into lateral occipital, posterior parietal, and right lateral prefrontal regions ( $p < 0.05$ , familywise error rate corrected). In contrast, the correspondence between neural representational similarity and the representational similarity as predicted by a model with no selective attention was restricted to early visual areas (similar results were also found with the prototype model; see Figure S3).

Lateral prefrontal regions, including insular cortex and IFG, were also linked to attention-biased exemplar representations. Previous work has highlighted the role of these regions in categorization and concept learning [41–43], specifically in representing evidence for perceptual decision making [44]. Most notably, processing in anterior insula and right IFG has been associated with active maintenance of information for comparative processing [45] that is sensitive to category boundaries [46]. Our results extend beyond these previous findings by showing that localized activation within lateral PFC tracks the specific similarity computations of attention-weighted exemplar representations during categorization.

The current findings make clear that multiple brain regions are engaged during category decision making, a conclusion consistent with previous model-based fMRI studies of categorization [21, 43]. We extend beyond this work by presenting the first results to match computational predictions from formal theories of category learning to the information in patterns of brain response. Critically, we use the consistency between brain and model for theory selection, finding greater consistency with exemplar theory.

Yet, important questions remain. Our results suggest that concrete experiences during learning are stored in similarity-based cortical representations and that these representations are activated during later categorization. Understanding how activation of exemplar memory traces gives rise to coherent

brain states that predict exemplar representational match will further characterize the mechanisms of categorization. It is worth noting that prototype model predictions were not completely inconsistent with brain response. This finding may suggest that prototype mechanisms do explain something about underlying brain states, but the exemplar model explains decidedly more in this categorization task. A limitation of the current study is the focus on the postlearning processes and representations of categorization. Using the proposed model-based fMRI method to investigate the development of category representations and tuning of attention weights over the course of learning [3, 21, 39] will shed light on the algorithms of category learning and how these algorithms are implemented in the brain.

Bridging levels of analysis represents a fundamental shift in testing algorithmic models of cognitive theories. The novel approach detailed here has broad implications for cognitive neuroscience research beyond the exemplar versus prototype debate we explore in the current paper [47]. Holding models accountable for not only behavior but also neural implementation provides a means for forging links between theory and brain and advancing more complete descriptions of the algorithms of cognition.

#### Supplemental Information

Supplemental Information includes three figures, two tables, and Supplemental Experimental Procedures and can be found with this article online at <http://dx.doi.org/10.1016/j.cub.2013.08.035>.

#### Acknowledgments

The authors thank K. Crabtree, A. Wattenberger, and M. Schlichting for assistance with data collection and K. Norman for helpful comments and discussion. This project was supported by National Institute of Mental Health (NIMH) grant R21-MH091523 awarded to B.C.L. and A.R.P., AFOSR grant FA9550-10-1-0268 awarded to B.C.L., and an NSF Career Award awarded to A.R.P.

Received: February 15, 2013

Revised: July 23, 2013

Accepted: August 9, 2013

Published: October 3, 2013

#### References

- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (San Francisco: Freeman).
- Sejnowski, T.J., Koch, C., and Churchland, P.S. (1988). Computational neuroscience. *Science* 241, 1299–1306.
- Love, B.C., and Gureckis, T.M. (2007). Models in search of a brain. *Cogn. Affect. Behav. Neurosci.* 7, 90–108.
- Purcell, B.A., Heitz, R.P., Cohen, J.Y., Schall, J.D., Logan, G.D., and Palmeri, T.J. (2010). Neurally constrained modeling of perceptual decision making. *Psychol. Rev.* 117, 1113–1143.
- Nosofsky, R.M. (1986). Attention, similarity, and the identification-categorization relationship. *J. Exp. Psychol. Gen.* 115, 39–61.
- Love, B.C., Medin, D.L., and Gureckis, T.M. (2004). SUSTAIN: a network model of category learning. *Psychol. Rev.* 111, 309–332.
- Forstmann, B.U., Wagenmakers, E.-J., Eichele, T., Brown, S., and Serences, J.T. (2011). Reciprocal relations between cognitive neuroscience and formal cognitive models: opposites attract? *Trends Cogn. Sci.* 15, 272–279.
- White, C.N., and Poldrack, R.A. (2013). Using fMRI to constrain theories of cognition. *Perspect. Psychol. Sci.* 8, 79–83.
- Rugg, M.D., and Thompson-Schill, S.L. (2013). Moving forward with fMRI data. *Perspect. Psychol. Sci.* 8, 84–87.
- Reed, K. (1972). Pattern recognition and categorization. *Cognit. Psychol.* 407, 382–407.

11. Posner, M.I., and Keele, S.W. (1968). On the genesis of abstract ideas. *J. Exp. Psychol.* 77, 353–363.
12. Medin, D.L., and Schaffer, M.M. (1978). Context theory of classification learning. *Psychol. Rev.* 85, 207–238.
13. Zaki, S.R., Nosofsky, R.M., Stanton, R.D., and Cohen, A.L. (2003). Prototype and exemplar accounts of category learning and attentional allocation: a reassessment. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 1160–1173.
14. Minda, J.P., and Smith, J.D. (2002). Comparing prototype-based and exemplar-based accounts of category learning and attentional allocation. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 275–292.
15. Norman, K.A., Polyn, S.M., Detre, G.J., and Haxby, J.V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* 10, 424–430.
16. Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Front Syst Neurosci* 2, 4.
17. Kriegeskorte, N., Goebel, R., and Bandettini, P. (2006). Information-based functional brain mapping. *Proc. Natl. Acad. Sci. USA* 103, 3863–3868.
18. O'Doherty, J.P., Hampton, A., and Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Ann. N Y Acad. Sci.* 1104, 35–53.
19. Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., and Dolan, R.J. (2006). Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.
20. Hampton, A.N., Bossaerts, P., and O'Doherty, J.P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc. Natl. Acad. Sci. USA* 105, 6741–6746.
21. Davis, T., Love, B.C., and Preston, A.R. (2012). Learning the exception to the rule: model-based fMRI reveals specialized representations for surprising category members. *Cereb. Cortex* 22, 260–273.
22. O'Doherty, J.P., Dayan, P., Friston, K., Critchley, H., and Dolan, R.J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron* 38, 329–337.
23. Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430.
24. Kamitani, Y., and Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nat. Neurosci.* 8, 679–685.
25. Naselaris, T., Prenger, R.J., Kay, K.N., Oliver, M., and Gallant, J.L. (2009). Bayesian reconstruction of natural images from human brain activity. *Neuron* 63, 902–915.
26. Miyawaki, Y., Uchida, H., Yamashita, O., Sato, M.A., Morito, Y., Tanabe, H.C., Sadato, N., and Kamitani, Y. (2008). Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. *Neuron* 60, 915–929.
27. Connolly, A.C., Guntupalli, J.S., Gors, J., Hanke, M., Halchenko, Y.O., Wu, Y.-C., Abdi, H., and Haxby, J.V. (2012). The representation of biological classes in the human brain. *J. Neurosci.* 32, 2608–2618.
28. Hintzman, D.L. (1991). Why are formal models useful in psychology? *Relating Theory and Data: Essays on Human Memory in Honor of Bennet B. Murdock* (Hillsdale: Erlbaum), pp. 39–56.
29. Estes, W.K. (1975). Human behavior in mathematical perspective. *Am. Sci.* 63, 649–655.
30. Grill-Spector, K., Kourtzi, Z., and Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Res.* 41, 1409–1422.
31. Palmeri, T.J., and Gauthier, I. (2004). Visual object understanding. *Nat. Rev. Neurosci.* 5, 291–303.
32. Jiang, X., Bradley, E., Rini, R.A., Zeffiro, T., Vanmeter, J., and Riesenhuber, M. (2007). Categorization training results in shape- and category-selective human neural plasticity. *Neuron* 53, 891–903.
33. Folstein, J.R., Palmeri, T.J., and Gauthier, I. (2013). Category learning increases discriminability of relevant object dimensions in visual cortex. *Cereb. Cortex* 23, 814–823.
34. Behrmann, M., Geng, J.J., and Shomstein, S. (2004). Parietal cortex and attention. *Curr. Opin. Neurobiol.* 14, 212–217.
35. Wagner, A.D., Shannon, B.J., Kahn, I., and Buckner, R.L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends Cogn. Sci.* 9, 445–453.
36. Cabeza, R., Ciaramelli, E., Olson, I.R., and Moscovitch, M. (2008). The parietal cortex and episodic memory: an attentional account. *Nat. Rev. Neurosci.* 9, 613–625.
37. Freedman, D.J., and Assad, J.A. (2006). Experience-dependent representation of visual categories in parietal cortex. *Nature* 443, 85–88.
38. Aizenstein, H.J., MacDonald, A.W., Stenger, V.A., Nebes, R.D., Larson, J.K., Ursu, S., and Carter, C.S. (2000). Complementary category learning systems identified using event-related functional MRI. *J. Cogn. Neurosci.* 12, 977–987.
39. Little, D.M., Klein, R., Shobat, D.M., McClure, E.D., and Thulborn, K.R. (2004). Changing patterns of brain activation during category learning revealed by functional MRI. *Brain Res. Cogn. Brain Res.* 22, 84–93.
40. Zeithamova, D., Maddox, W.T., and Schyns, P. (2008). Dissociable prototype learning systems: evidence from brain imaging and behavior. *J. Neurosci.* 28, 13194–13201.
41. Seger, C.A., Poldrack, R.A., Prabhakaran, V., Zhao, M., Glover, G.H., and Gabrieli, J.D. (2000). Hemispheric asymmetries and individual differences in visual concept learning as measured by functional MRI. *Neuropsychologia* 38, 1316–1324.
42. Li, S., Mayhew, S.D., and Kourtzi, Z. (2009). Learning shapes the representation of behavioral choice in the human brain. *Neuron* 62, 441–452.
43. Nosofsky, R.M., Little, D.R., and James, T.W. (2012). Activation in the neural network responsible for categorization and recognition reflects parameter changes. *Proc. Natl. Acad. Sci. USA* 109, 333–338.
44. Ho, T.C., Brown, S., and Serences, J.T. (2009). Domain general mechanisms of perceptual decision making in human cortex. *J. Neurosci.* 29, 8675–8687.
45. Philiastides, M.G., and Sajda, P. (2007). EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. *J. Neurosci.* 27, 13082–13091.
46. Grinband, J., Hirsch, J., and Ferrera, V.P. (2006). A neural representation of categorization uncertainty in the human brain. *Neuron* 49, 757–763.
47. Mather, M., Cacioppo, J.T., and Kanwisher, N. (2013). Introduction to the special section: 20 years of fMRI—What has it done for understanding cognition? *Perspect. Psychol. Sci.* 8, 41–43.