

# Cortical areas supporting category learning identified using functional MRI

(nondeclarative memory/occipital cortex/frontal cortex)

P. J. REBER<sup>†</sup>, C. E. L. STARK<sup>†</sup>, AND L. R. SQUIRE<sup>†‡§</sup>

Departments of <sup>†</sup>Psychiatry and <sup>‡</sup>Neurosciences, University of California, San Diego, La Jolla, CA 92093; and <sup>§</sup>Veterans Affairs Medical Center, San Diego, CA 92161

Contributed by L. R. Squire, November 14, 1997

**ABSTRACT** Functional MRI was used to identify cortical areas involved in category learning by prototype abstraction. Participants studied 40 dot patterns that were distortions of an underlying prototype and then, while functional MRI data were collected, made yes-no category judgments about new dot patterns. The dot patterns alternated between ones mostly requiring a “yes” response and ones mostly requiring a “no” response. Activity in four cortical areas correlated with the category judgment task. A sizeable posterior occipital cortical area (BA 17/18) exhibited significantly less activity during processing of the categorical patterns than during processing of noncategorical patterns. Significant increases in activity during processing the categorical patterns were observed in left and right anterior frontal cortex (BA 10) and right inferior lateral frontal cortex (BA 44/47). Decreases in activation of visual cortex when categorical patterns were being evaluated suggest that these patterns could be processed in a more rapid or less effortful manner after the prototype had been learned. Increases in prefrontal activity associated with processing categorical patterns could be related to any of several processes involved in retrieving information about the learned exemplars.

An encounter with a series of items is sufficient to acquire knowledge about the category to which the items belong. Learning about the category requires gaining some information about the exemplars that define the category, but this learning appears to be independent of the brain system that supports the ordinary learning of items and events. Amnesic patients, who have bilateral damage to medial temporal lobe or midline diencephalic structures and who have severely impaired declarative memory, acquired categorical information from exemplars as well as normal subjects (1, 2). This finding suggests that category learning is nondeclarative; and that, like other nonconscious memory abilities (e.g., priming and the learning of skills and habits), category learning depends on brain structures other than the medial temporal lobe-diencephalic brain system that supports declarative memory.

The task used to study category learning in amnesic patients was a modification of one introduced by Posner and Keele (3). Subjects studied a series of dot patterns that were distortions of an underlying prototypic dot pattern. Subsequently, they were told that the study items belonged to a single category and that they should now try to classify new dot patterns according to whether they did or did not belong to the same category. We have used functional neuroimaging with this task to identify areas of the brain important for acquiring and using categorical knowledge. Accordingly, we collected functional MRI data

while normal subjects performed the dot pattern categorization task.

## METHODS

Six healthy adult, right-handed volunteers (two women, four men; mean age, 26 years) gave written informed consent prior to participating in the study. Both the study and test portions of the behavioral paradigm were performed in the MRI scanner. A mirror was placed so that stimuli could be back-projected onto a viewing screen placed at the foot of the scanner. The screen was  $\approx 3.5$  m from the subject's head. Before any stimuli were presented, it was verified for each participant that the full viewing area of the screen was visible. The material on the screen subtended a visual angle of  $5^\circ$ – $7^\circ$ .

**Category Learning.** Participants studied 40 dot patterns one at a time (Fig. 1). Each pattern was composed of nine dots, constructed as described previously (1). Specifically, each dot pattern was a “high distortion” of an underlying prototype dot pattern. Each pattern was on the viewing screen for 5 s, and participants imagined pointing to the center dot in the pattern to guarantee attention. Actual pointing occurred in previous behavioral studies, but here no pointing was done to reduce movement in the scanner. No scanning occurred during the study phase, and participants were not informed of the existence of a prototype.

After a short delay ( $\approx 2$  min), participants were told that the patterns had all belonged to a single category of patterns in the same sense that, if a series of dogs had been presented, they would all belong to the category “dog.” Scanning then occurred while 84 new dot patterns were presented. The 84 patterns consisted of four presentations of the prototype dot pattern, 20 low-distortions of the prototype dot pattern, 20 high-distortion dot patterns, and 40 random pattern (Fig. 1). For each pattern, participants judged whether it came from the same category as the training patterns (a “yes” response was correct for the prototype, low, and high dot patterns; a “no” response was correct for the random patterns). Each pattern was presented for 3.5 s with a 500-ms interval between patterns. The 84 test patterns were presented in 12 blocks of seven patterns each, which alternated between blocks containing predominantly categorical patterns and blocks containing predominantly random patterns. Specifically, each block contained either six categorical patterns that required a “yes” response and one random pattern that required a “no” response or vice versa. Because there were only 40 random patterns to be distributed among the 12 blocks, two blocks that contained mainly random patterns were assigned five random patterns and two categorical patterns. Participants made their yes-no categorical judgments using a fiber-optic button box. The entire study-test sequence was then repeated with different study items and test items (with  $\approx 10$ -min delay between runs). In summary, each of the two scanning runs consisted of

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/95747-4\$2.00/0  
PNAS is available online at <http://www.pnas.org>.

Abbreviation: ROI, region of interest.

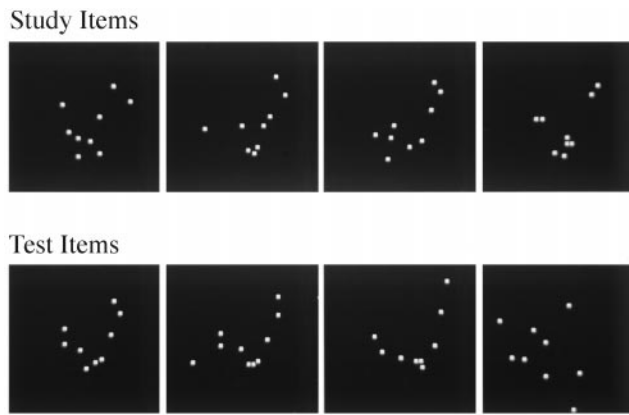


FIG. 1. Examples of study items and test items used to assess classification learning of dot patterns. The study items were high distortions of a prototype dot pattern. The test items, illustrated left to right, were presentations of the training prototype, low and high distortions of the training prototype, and random dot patterns.

12 alternating blocks of mostly categorical stimuli (six blocks) and mostly random stimuli (six blocks).

**Functional Imaging.** Imaging was performed on a GE 1.5T SIGNA clinical MRI scanner fitted with a high-performance local head gradient and RF coils (4, 5). Functional T2\*-weighted images were acquired by using an echoplanar single-shot pulse sequence with a matrix size of  $64 \times 64$ , echo time of 40 ms, flip angle of  $90^\circ$ , and in-plane resolution of  $3.75 \times 3.75$  mm. For each scanning run, a total of 98 images were acquired for 15–16 adjacent 7-mm axial slices in an interleaved fashion with a repetition time of 3.5 s. The first two images from each slice were discarded to assure that the magnetic resonance signal had reached equilibrium on each slice. For anatomical localization, a standard whole-brain, T1-weighted, three-dimensional MP-RAGE sequence was acquired (flip angle =  $10^\circ$ , FOV = 24 cm,  $256 \times 256 \times 128$  acquisition matrix, sagittal slices, thickness = 1.1–1.3 mm).

Images were first corrected for distortion due to field nonhomogeneity (6) and were coregistered through time by using a two-dimensional registration algorithm (AFNI analysis software, ref. 7). Each slice was spatially smoothed by using a two-dimensional (in-plane) Gaussian kernel, FWHM = 7.5 mm. Linear drift in the overall magnitude of the magnetic resonance signal in each voxel over the course of the entire scan was eliminated (linear drift was estimated by computing the change in signal across the blocks that contained noncategorical patterns). Within each run, voxels containing a magnitude change  $>10\%$  in one repetition (3.5 s) were identified as likely contaminated by motion and eliminated from further analysis. Finally, the 12 runs were transformed to conform to the atlas of Talairach and Tournoux (8) (with a final voxel size of  $2.5 \text{ mm}^3$ ) and averaged together to identify areas of common activation across the six participants.

Areas exhibiting activity selective for either the categorical or noncategorical dot patterns were identified by correlating the observed time course of each voxel against an idealized reference function derived from the 12 alternating blocks of dot patterns and adjusted to reflect the lag between neural activity change and hemodynamic response (signal rise was assumed to occur linearly over a 6 s delay; fall time was assumed to be linear over 9 s). The resultant statistical map was then thresholded to eliminate voxels for which the correlation with the reference function was  $<0.40$  ( $P < 6 \times 10^{-5}$ , uncorrected for multiple comparisons). Finally, significant areas of activation were required to comprise a cluster of correlated voxels with a total volume  $>350 \text{ mm}^3$ . A volume  $>350 \text{ mm}^3$  would include at least 22 contiguous voxels in the transformed data.

Simulations with random Gaussian-distributed noise (matched to each participant's brain with respect to size, shape, and the mean and SD of the measured signal for each voxel) indicated that this threshold has an extremely low false positive rate for the imaging parameters and analysis techniques used. In six full simulations of the entire dataset (12 runs each), no clusters of voxels exceeding the correlation threshold (0.4) were observed that were greater than three voxels in extent. The data analysis threshold was 22 voxels. By using a correlation threshold of 0.35, six simulations yielded 58 voxel clusters ranging from two to 10 voxels in extent. At a correlation of 0.30, six simulations yielded two clusters  $>22$  voxels in extent and a large number of smaller ones. A correlation threshold of  $r \geq 0.4$  appears to be a conservative threshold that is unlikely to yield false positive findings.

## RESULTS

Participants were able to endorse the dot patterns according to how closely they resembled the prototype of the training patterns. Fig. 2 shows categorization performance by item type for five of the six participants (for one participant, the response box failed during testing). Performance resembled the performance of subjects tested previously in behavioral studies outside the scanner (1, 2). Categorization judgments were 63.6% correct, significantly better than chance (50%),  $t(4) = 3.14$ ,  $P < 0.04$ , and there was a significant effect of item type on performance,  $F[1,5] = 14.37$ ,  $P < .02$ .

The mean reaction time for correctly endorsing categorical patterns was  $1,498 \text{ ms} \pm 209$ , somewhat faster than the reaction time for correctly rejecting noncategorical patterns  $1,691 \text{ ms} \pm 255$ . This difference was not significant ( $P > 0.20$ , though four of the five subjects who provided behavioral data responded faster to categorical patterns).

Areas of significant change in activation are shown in Fig. 3. All the changes were small, representing  $<1\%$  average signal change, but they were robust as indicated by the fact that they occurred as a cluster of voxels ( $\geq 22$ ) correlating strongly with the reference function ( $r \geq 0.4$ ). A sizeable area of decreased activity associated with processing categorical stimuli (shown in blue on Fig. 3) was observed in posterior occipital cortex, including cortex in areas V1 and V2, over three subclusters

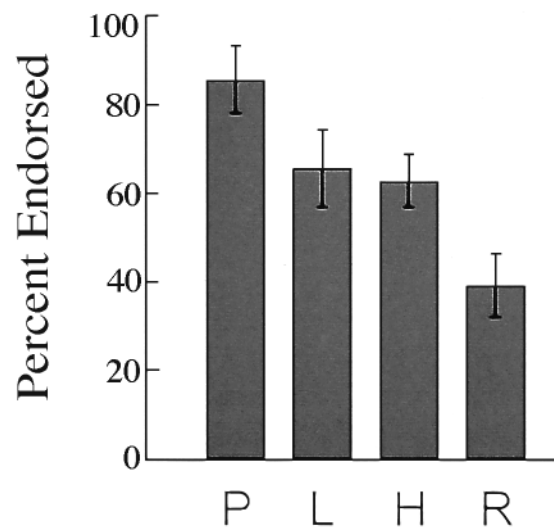


FIG. 2. Classification of new dot patterns after studying 40 exemplar patterns. The endorsement rate varied in accordance with how closely the test items resembled the study items. P, instances of the prototype (four patterns per test); L, low distortions of the prototype (20 patterns per test); H, high distortions of the prototype (20 patterns per test); R, random patterns (40 patterns per test). Error bars show the SEM.

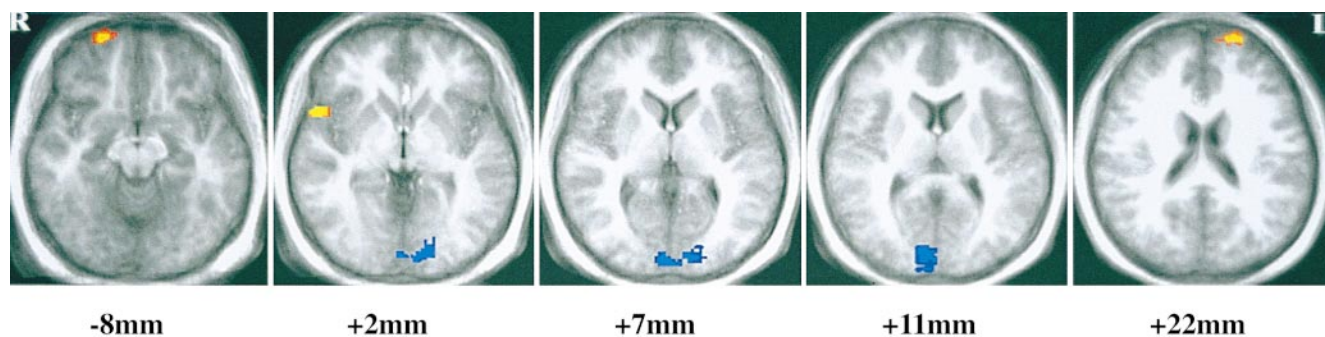


FIG. 3. Brain areas activated. Areas of significant signal change are shown as color overlays on the averaged axial structural images (transformed to the atlas of Talairach and Tournoux, ref. 8). Images are oriented according to radiologic convention (the right side of the brain is on the left side of the image, anterior-posterior is shown top to bottom). The distance of the displayed slice above the AC-PC line is shown below each image. Areas where activation increased when subjects processed categorical patterns are shown in red and yellow. Areas where activation decreased during the processing of categorical patterns are indicated in blue.

centered at Talairach coordinates ( $x, y, z$ ):  $(-14, -86, +1)$ ,  $(+1, -92, +7)$ , and  $(+9, -87, +11)$ . Significant increases in activity during processing the categorical stimuli (shown in red and yellow) were observed in left superior frontal gyrus, Talairach coordinates  $(-16, +62, +23)$  (BA 10), right superior frontal gyrus  $(+20, +62, -9)$  (BA 10), and right inferior frontal gyrus  $(+52, +12, +1)$  (BA 44/47).

The four areas of reliable change in activity were inspected as regions of interest (ROIs) in each scanning run of the individual data from all six participants. There were no systematic differences in the results for the two runs. For three of the brain areas (all except right frontal area 10), all six participants exhibited changes in activity consistent with the averaged data. For the ROIs in posterior occipital cortex and left superior frontal gyrus (area 10), five of the six participants exhibited significant changes in their individual data ( $P < 0.05$ , uncorrected). For the other two ROIs (right area 44/47 and right area 10), only three of the participants exhibited significant changes ( $P < 0.05$ ). Moreover, for the ROI in the right superior frontal gyrus (area 10), no signal was observed in five of the twelve runs, possible due to susceptibility-related signal dropout. It is possible that the findings in these two areas (right 44/47 and right 10) may not be as reliable as in the others.

## DISCUSSION

This study contrasted the processing of categorical and non-categorical patterns. Some areas important for the categorization task may have been active both when categorical and noncategorical patterns were presented. Activity in such areas would not have been detected in this experiment. In contrast, activity that was selective for the processing of categorical or noncategorical patterns was detected.

The decreased activation in the posterior occipital cortex during processing of the categorical patterns occurred in areas involved in relatively early visual processing. The decrease in activity for the categorical patterns compared with the non-categorical patterns suggests that the categorical patterns were visually processed more easily or quickly (leading to less overall activity). Decreases in activity in posterior cortex for recently presented stimuli, in comparison to new stimuli, have been reported in studies of repetition priming for words (9–11). The effect observed here may be similar. That is, decreased activity may be observed not only when a stimulus is repeated a second time but also when a stimulus is encountered that is similar to one presented recently. Thus, the present findings suggest that processing of unfamiliar visual stimuli that are similar to a learned prototype results in less overall activity in visual cortex than the processing of stimuli that are entirely new.

Increased activity in the frontal cortical areas reflect additional processing that occurs as a result of evaluating categor-

ical patterns. Several processes could be involved in identifying a pattern as categorical. For example, judging that a successful match has occurred between the categorical patterns and the prototype (or between the categorical patterns and the exemplars that were studied) may increase spatial working memory demands. Calculating a successful match may well require evaluating more of the dots in the pattern than calculating a nonmatch. Activity in right inferior frontal gyrus in the vicinity of area 47 has been associated with spatial working memory tasks (12, 13). Increased activity in this region could be related to processing needed to judge that a successful match has occurred between a categorical pattern and the prototype (or the exemplars).

Another process that could be associated with evaluating categorical patterns is that the participants may attempt consciously attempt to retrieve information about the previously seen patterns in order to respond positively. Previous studies of memory function have implicated frontal cortical area 10 in the intentional retrieval of nonverbal material (14–17). Frontal cortical area 44 has also been implicated in the retrieval of nonverbal material (14, 17, 18), although the activity observed in these studies was more superior (i.e., 16–24 mm above the AC-PC line) than the activity observed here (+1 mm). The increased activity in area 10 may be related to attempts to retrieve information about the studied exemplars. Although the categorization task depends on nondeclarative memory, it is possible that the participants in this study attempted to explicitly retrieve information about the study patterns. Previous findings that amnesic patients perform as well as controls on this task indicate that the availability of declarative memory does not give the controls any advantage. Thus, attempting to retrieve the exemplars consciously would not be expected to affect categorization performance.

Desimone (19) proposed that prefrontal cortex may index and manage information stored elsewhere in the brain during tasks that involve searching for a match. In the current experiment, this idea suggests that prefrontal activity in area 10 could be related to the executive function involved in successful retrieval of the prototype or the studied exemplars for the purpose of matching to the categorical patterns. If the prefrontal cortex is indexing storage elsewhere in the brain, one might have expected some other more posterior cortical area to exhibit a parallel increase in activity for categorical patterns. Yet increased activity outside the frontal lobe was not observed. If the sites with which the frontal cortex interacts are distributed (resulting in small changes in activity in each sampled region during functional MRI), or if they were not localized to the same areas across participants, such sites would not have been detected in the current experiment.

In summary, four brain areas exhibited activity selective for processing categorical or noncategorical dot patterns. Poste-



rior occipital cortex exhibited decreased activity when participants evaluated categorical patterns for the purpose of making category judgments. This reduction in activity suggests that more rapid (or less effortful) visual processing of the categorical patterns is possible after the category has been learned. Three prefrontal cortical areas exhibited more activity when participants evaluated the categorical patterns than when they evaluated the noncategorical patterns. This activity might arise from the process needed to retrieve prototype information or from the spatial working memory demands needed to achieve a successful match to the prototype.

We thank Richard Buxton and Eric Wong for assistance with the functional imaging protocol and James Moore for assistance in data collection. This research was supported by the Medical Research Service of the Department of Veterans Affairs and by National Institute of Mental Health Grants MH24600 (L.R.S.) and F32 MH11150 (P.J.R.).

1. Knowlton, B. J. & Squire, L. R. (1993) *Science* **262**, 1747–1749.
2. Squire, L. R. & Knowlton, B. J. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 12470–12474.
3. Posner, M. I. & Keele, S. W. (1968) *J. Exp. Psychol.* **77**, 353–363.
4. Wong, E. C., Bandettini, P. A. & Hyde, J. S. (1992) *Proc. Soc. Magnetic Resonance Med. 11th Ann. Meeting*, 105.
5. Wong, E. C., Boskamp, E. & Hyde, J. S. (1992) *Proc. Soc. Magnetic Resonance Med. 11th Ann. Meeting*, 4015.
6. Reber, P. J., Wong, E. C., Buxton, R. B. & Frank, L. R. (1998) *Magnetic Resonance in Medicine*, in press.
7. Cox, R. W. (1996) *Comput. Biomed. Res.* **29**, 162–173.
8. Talairach, J. & Tournoux, P. (1988) *A Co-Planar Stereotaxic Atlas of the Human Brain* (Thieme, Stuttgart, Germany).
9. Squire, L. R., Ojemann, J. G., Miezin, F. M., Petersen, S. E., Videen, T. O. & Raichle, M. E. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 1837–1841.
10. Buckner, R. L., Petersen, S. E., Ojemann, J. G., Miezin, F. M., Squire, L. R. & Raichle, M. E. (1995) *J. Neurosci.* **15**, 12–29.
11. Schacter, D. L., Alpert, N., Savage, C. R., Rauch, S. L. & Albert, M. S. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 321–325.
12. Smith, E. E., Jonides, J., Keoppe, R. A., Awh, E., Schumacher, E. H. & Minoshima, S. (1995) *J. Cognit. Neurosci.* **7**, 337–356.
13. Courtney, S. M., Ungerleider, L. G., Kell, K. & Haxby, J. V. (1997) *Nature (London)* **386**, 608.
14. Tulving, E., Markowitsch, H. J., Craik, F. I. M., Habib, R. & Houle, S. (1994) *NeuroReport* **5**, 2525–2528.
15. Roland, P. E. & Gulyas, B. (1995) *Cereb. Cortex* **1**, 79–93.
16. Haxby, J. V., Ungerleider, L. G., Horwitz, B., Maisog, J. M., Rapoport, S. I. & Grady, C. L. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 922–927.
17. Schacter, D. L., Reiman, E., Uecker, A., Polster, M. R., Yung, L. S. & Cooper, L. A. (1995) *Nature (London)* **376**, 537–539.
18. Moscovitch, M., Kapur, S., Kohler, S. & Houle, S. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 3721–3725.
19. Desimone, R. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 13494–13499.