We monitored regional cerebral activity with BOLD fMRI during acquisition of a novel semantic category and subsequent categorization of test stimuli by a rule-based strategy or a similarity-based strategy. We observed different patterns of activation in direct comparisons of rule- and similarity-based categorization. During rule-based category acquisition, subjects recruited anterior cingulate, thalamic, and parietal regions to support selective attention to perceptual features, and left inferior frontal cortex to help maintain rules in working memory. Subsequent rule-based categorization revealed anterior cingulate and parietal activation while judging stimuli whose conformity with the rules was readily apparent, and left inferior frontal recruitment during judgments of stimuli whose conformity was less apparent. By comparison, similarity-based category acquisition showed recruitment of anterior prefrontal and posterior cingulate regions, presumably to support successful retrieval of previously encountered exemplars from long-term memory, and bilateral temporal-parietal activation for perceptual feature integration. Subsequent similarity-based categorization revealed temporal-parietal, posterior cingulate, and anterior prefrontal activation. These findings suggest that large-scale networks support relatively distinct categorization processes during the acquisition and judgment of semantic category knowledge.

© 2004 Elsevier Inc. All rights reserved.

Keywords: BOLD fMRI; Rule-based category; Similarity-based category

Introduction

Semantic memory involves semantic content and categorization processes that act on that content (Tulving, 1972). Categorization is essential to organizing semantic content in a meaningful way (Smith, 1995). At least two broad forms of categorization processes have been proposed. In similarity-based categorization, category membership is determined by overall similarity to established exemplars or to a prototype (Medin and Schaffer, 1978; Medin et al., 1993; Smith and Medin, 1981). This process appears to entail exemplar or prototype retrieval, as well as organization of perceptual features contributing to an object’s overall appearance (Ashby et al., 1991). Rule-based categorization is employed when specific features must be considered (Bruner et al., 1956; Smith et al., 1998a,b). This process is resource demanding, since it presumably involves executive processes such as selective attention (to focus on requisite features), inhibitory control (to suppress irrelevant features), and working memory (to retain assessments of various features). Behavioral studies show that healthy subjects are capable of implementing either rule-based or similarity-based semantic categorization processes, even for the same stimulus set (Allen and Brooks, 1991; Grossman et al., 2003b; Koenig et al., 2002; Rips, 1989; Smith and Sloman, 1994). Qualitative distinctions between these two categorization processes suggest that they depend on partially distinct neural substrates. The present study used BOLD fMRI in an event-related design to examine cortical recruitment patterns in healthy young subjects during the acquisition and categorization of novel, meaningful stimuli by rule-based and by similarity-based strategies.

Behavioral studies of patients with central nervous system disease provide some evidence for neurologically separable categorization systems. In two recent studies, patients with Alzheimer’s disease and patients with frontotemporal dementia showed impaired rule-based categorization of object descriptions (Grossman et al., 2003b) and of novel visual items (Koenig et al., 2004, submitted for publication), although they were able to categorize by similarity-based processing in a relatively normal manner. This is consistent with the executive resource limitations observed in both Alzheimer’s disease (Fisher et al., 1990; LaFleche and Albert, 1995; Perry and Hodges, 1999) and frontotemporal dementia (Ellgren et al., 1993; Pasquier et al., 1999).

Functional neuroimaging work also implicates somewhat distinct brain regions for the processes supporting rule-based and similarity-based categorization.
similarity-based categorization. Neuroimaging studies of clinical tasks involving rule-based processes like Wisconsin Card Sorting and Raven’s Progressive Matrices emphasize activation of anterior cingulate and dorsolateral prefrontal cortices (Elliott et al., 1999; Rao et al., 1997; Savage et al., 2001). Some investigators relate executive processes such as selective attention and inhibitory control (i.e., processes characteristic of categorizing by rules) to activation of frontal regions (Petrides et al., 1993; Vendrell et al., 1995; Wagner et al., 2001). Switching processes that govern alternating between attributes of an object, objects in a category, and processes applied to an object or a category have been related to dorsolateral prefrontal cortex (D’Esposito et al., 1995; Garavan et al., 2000; Monchi et al., 2001; Savage et al., 2001; Seger et al., 2000a,b).

By comparison, similarity-based categorization, as purportedly used in classifying dot patterns or learning artificial grammars, appears to implicate parietal as well as frontal-striatal brain regions (Aizenstein et al., 2000; Poldrack and Gabrieli, 2001; Reber et al., 1998a; Seger et al., 2000a,b). The inferior parietal component may support feature integration necessary for global comparisons (Wilkinson et al., 2002). Neuroimaging work also suggests the contribution of posterior cingulate cortex, which may support retrieval of previously encountered exemplars (Rugg et al., 1998; Tulving et al., 1996) and anterior prefrontal cortex (Buckner et al., 1998; Lepage et al., 2002) for the successful retrieval of information from long-term memory.

Very few imaging studies have explicitly compared rule-based and similarity-based processing. A recent fMRI study (Grossman et al., 2002a,b) examined cortical activation during rule-based and similarity-based category assignments of written descriptions of objects that bore equal resemblance to either of two choice categories, but violated the rule-like specifications of one. For example, a “round object 3 in. in diameter” equally resembled a “pizza” or a “quarter,” but, by the rule governing a quarter’s invariable size, could only unequivocally belong to the category “pizza.” A direct contrast of rule-based minus similarity-based categorization decisions showed left dorsolateral prefrontal activation, consistent with support for executive resources during rule-based categorization.

Closer in structure to the present study is a recent PET study examining activation during categorization of novel visual stimuli (cartoon-like animals) by rule- and similarity-based processes (Patalano et al., 2001). Both categorization conditions showed activation in visual cortex, while only the rule-based categorization showed bilateral superior parietal, right dorsolateral prefrontal, and bilateral supplementary motor cortex activation. The authors argued that these regions support selective attention to spatial position, shifting attentional focus, and working memory.

The present study departs from the previous ones in several ways. We used a visual category of novel animals, although unlike the stimuli in the Patalano et al. study, the animals were relatively naturalistic. We imagined category acquisition as well as category use, whereas Patalano et al. imaged only the latter. Hence, the present study assesses cortical activation during learning by rule- and similarity-based processes as well as decision-making by those processes subsequent to learning. We designed the study so that we could directly compare the two categorization processes: Unlike Patalano et al., we employed a single category, structured such that subjects would generally endorse the identical items as members across both the rule-based and similarity-based categorization conditions. Finally, we employed empirical assessments of the relative contribution of particular stimulus features to inter-stimulus resemblance. This allowed us to construct a non-arbitrary category; that is, our category members were subjectively perceived as similar, and our rules of category membership reflected this category coherence. Our intention was to make our category easily learnable, so that patterns of activation would reflect the two categorization processes rather than subjects’ efforts to master a difficult task.

With these materials, we expected to observe distinct neural networks for each semantic categorization process: For rule-based categorization, we anticipated a network centering around inferior frontal and anterior cingulate regions to sustain working memory (e.g., for retaining the rules) and executive resources such as selective attention (e.g., for identifying relevant features). For similarity-based categorization, we anticipated a network including lateral temporal-parietal cortex supporting perceptual feature integration, and posterior cingulate and anterior prefrontal activation supporting successful retrieval of previously encountered exemplars.

Methods

Subjects

Twenty-eight healthy young volunteers participated, including 11 males and 17 females with a mean (±SD) age of 22.2 (±3.3) years and mean education of 15.2 (±1.7) years. All were right-handed native speakers of English. We were specifically interested in examining cortical activation supporting rule-based and similarity-based categorization, rather than participants’ spontaneous choice of categorization strategy. Hence, we excluded data from participants whose behavioral results suggested that they had not complied with our instructions. These included participants who seemed to focus on a single feature (i.e., they apparently applied a simple rule) in the Similarity condition. These also included participants who apparently did not cooperate, that is, made grossly inaccurate or random judgments. Eleven (6 males, 5 females, mean age of 22.0 (±2.6) years, mean education of 15.1 (±1.2) years) participated in the Rule condition: Fourteen participants were originally run in this condition, but three were excluded because their behavioral responses at test were less accurate than the group mean by at least two standard deviations. Nine (2 males, 7 females, mean age of 22.9 (±4.7) years, mean education of 15.3 (±2.5) years) participated in the Similarity condition: Fourteen were originally run, but data from 1 participant were excluded because responses at test did not differ from chance. Four additional participants (2 males, 2 females, mean age of 21.3 (±2.6) years, mean education of 15.3 (±1.9) years) were excluded from the Similarity condition because they appeared to spontaneously employ a consistent rule-like strategy at test. These four exhibited clearly systematic patterns in their judgments, that is, focusing exclusively on a single feature to determine category membership, or granting membership in exact proportion to how many features each item shared with the prototype. We do not report the results of these participants here since the group may be too small to demonstrate a reliable imaging profile, but the results are available from the authors.

Stimuli

Stimuli were taken from a set of 64 realistic images of biologically plausible novel animals. Examples can be seen in
the training stimuli shown in Fig. 1. The stimulus set, and the method of obtaining empirical rankings of the features’ contribution to judged resemblance (briefly described below), are described in detail elsewhere (Koenig et al., 2002). The set represented all possible combinations of six dichotomous features, that is, color (red-brown or yellow-green), legs (long or short), neck (low or raised), snout (long or short), tail (straight or curled), and teeth (fangs or tusks). All instances of a feature were identical, for example, all straight-tailed animals had tails of identical dimensions. All other features (e.g., texture, torso shape, facial details) were consistent throughout the set. Prior to this study, a cohort of healthy young subjects rated all pair-wise combinations in the set for resemblance. We then used multidimensional scaling to obtain a ranking of the six features’ contribution to resemblance judgments.

We identified the four highest-ranking features from the multidimensional scaling procedure, which served as “contributing” features. These were snout, legs, color, and neck type. One animal was then chosen at random to serve as a prototype of the category. The two lowest-ranked features (i.e., teeth and tail) served as “distractor” features that were irrelevant.

A Rule Training Condition

Which has 3 features?

short legs
upright neck
yellow color

distinctive features

B Similarity Training Condition

Which is a crutter?

crutter

Fig. 1. Illustration of screens used during the training session. Panel A: rule-based training, showing the four criterial features in the upper half of the screen and two crutters in the lower half (MEMBER on left). Panel B: Similarity-based training, showing a prototype in the upper half of the screen and two crutters in the lower half (MEMBER on left).
in sequence. Stimuli were video-projected onto a screen, which subjects viewed via a system of mirrors while in the magnet bore. Presentation was controlled by a computer using PsyScope software (Cohen et al., 1993), which recorded responses.

We presented these MEMBER/HIGH DISTORTION pairs in each of two training conditions, described below.

Rule training procedure

Participants were told that an animal called a “crutter” had to have at least three of four particular features, and that they would be shown pairs of animals, one of which would be a crutter and one of which would not. Their task was to decide which animal in each pair was a crutter, based on the “at-least-three-of-four features” rule. The training session in the scanner followed. The 40-item sequence of MEMBER/HIGH DISTORTION training pairs was presented. Each pair was arrayed horizontally in the lower half of the projected image, and was accompanied by brief written descriptions and outlines of the requisite features (e.g., the words “long snout” with a sketch of a long snout) in the upper half. The description of the color feature (i.e., “yellow color”) was illustrated by a yellow rectangle rather than a sketch. The sketches were black outlines of the contours of the features, intended to clarify the verbal descriptions by clearly showing which feature variants they referred to. At the same time, the sketches’ lack of visual detail ensured that subjects could not rely on a perceptual identity match when identifying the features in the stimulus items. The MEMBERS and HIGH DISTORTION items were equally distributed in the left and right positions. The sentence “Which has 3 features?” appeared at the top of the screen. Fig. 1 panel A shows a sample Rule training stimulus. Stimuli appeared every 12 s. Participants’ responses were indicated by a button press on a handheld box, with the buttons corresponding spatially to the location (left or right) of the chosen item. The stimulus remained until the subject responded, after which the screen became blank; response times never exceeded 6 s. Nine seconds after a stimulus item first appeared, the item reappeared for 2 s with a black rectangle surrounding the correct choice (i.e., the category member), as a means of providing feedback. We analyzed only items to which participants responded (five participants each missed one training trial).

Prior to entering the scanner, participants received a practice session at a laptop computer. Although we expected the training tasks to be readily comprehensible, we wished to ensure that subjects’ efforts during the first few trials would be directed towards performing the training task rather than acclimating themselves to the procedure. Hence, the practice stimuli were designed to familiarize subjects with the training procedure without exposing them to the animal stimuli. They consisted of simple abstract geometric figures with three variable features: large or small size, circular or triangular appendage, and striped or dotted pattern. None of these features coincided with features of the animal stimuli. The practice stimuli were presented in a series of images that were configured analogously to the training stimuli; that is, written descriptions and sketches of the three features appeared above a pair of items, one of which contained two of the features and one of which contained one. Participants were told that an object called a “dort” had to have at least two of the features, and they were asked to indicate which item in each pair was a dort. Feedback was analogous to the actual training trials, that is, a black rectangle appearing around the correct choice. Participants were exposed to three practice trials, and all of them indicated understanding of the procedure.

Similarity training procedure

Participants were told that they would see pictures of an animal called a “crutter” along with pairs of pictured animals, one of which would also be a crutter and one of which would not. Their task was to decide which animal in each pair was a crutter, based on which animal was “more similar” to the sample crutter. Subjects were urged to make a quick decision based on their initial impressions. The training session in the scanner followed. A sequence of stimuli was presented consisting of the prototype in the upper half of the projected image and the same horizontally arrayed MEMBER/HIGH DISTORTION pairs as in the Rule training procedure in the lower half. The caption “crutter” appeared directly below the prototype and the sentence “Which is a crutter?” appeared at the top of the screen. The stimulus displays for the Similarity and Rule training conditions thus each contained pictorial and verbal elements (including sentences and captions), and were designed to minimize display discrepancies between these inherently different training processes. Fig. 1 panel B shows a sample Similarity training stimulus. All other aspects (e.g., timing, mode of response, feedback, and data collection) were the same as in the Rule training session. We analyzed only trials to which participants responded (three participants missed a total of seven trials).

Participants practiced at a laptop computer prior to this training procedure as well. The practice items were the same as those used prior to the Rule training session. The stimuli consisted of a prototype geometric configuration in which the three requisite features were represented; this appearing above the same pairs used in the Rule practice. Participants were told that the prototype was a “dort” and were asked to indicate which item in each pair was also a dort based on greater similarity to the sample dort. All participants indicated understanding of this training task as well.

Testing procedure

Both training conditions were followed by the identical test. Participants saw a sequential presentation of individual animals in a fixed semi-random order, and judged whether each was a crutter. They were instructed to judge in accordance with how they had been trained; that is, participants in the Rule condition were to judge category membership by an item’s adherence to the “three-out-of-four features” rule, and participants in the Similarity condition were to judge category membership by an item’s overall resemblance to the training example. This sequence consisted of a 50-item subset of the complete set of 64 animals. We used a reduced set so that the test session would not exceed an imaging run that is a maximum of 10 min in duration. The subset contained 20 MEMBERS, 14 LOW DISTORTION items, and 16 HIGH DISTORTION items. Items of a particular type (e.g., LOW DISTORTION) appeared no more than three times sequentially. Stimuli appeared every 12 s, and remained visible until the subject responded. Participants indicated their choice by button press, with the right and left buttons corresponding, respectively, to “yes” and “no.” No feedback was provided. Responses and reaction times were recorded. Analyses included only trials in which subjects responded (only one Rule participant did not respond to one LOW DISTORTION stimulus).
Imaging data acquisition and statistical analysis

The experiment was carried out at 4T on a GE prototype Echospeed scanner capable of ultrafast imaging. Firm foam padding was used to restrict head motion. Each imaging protocol began with a 10–15 min acquisition of 5 mm thick adjacent slices for determining regional anatomy, including sagittal localizer images (TR = 500 ms, TE = 10 ms, 192 × 256 matrix), T2-weighted axial images (FSE, TR = 2000 ms, TEeff = 85 ms), and T1-weighted axial images of slices used for fMRI anatomic localization (TR = 600 ms, TE = 14 ms, 192 × 256 matrix). Gradient echo echoplanar images were acquired for detection of alterations of blood oxygenation accompanying increased mental activity. All images were acquired with fat saturation, a rectangular FOV of 20 × 15 cm, flip angle of 90°, 5 mm slice thickness, an effective TE of 50 ms, and a 64 × 40 matrix, resulting in a voxel size of 3.75 × 3.75 × 5 mm. The echoplanar acquisitions consisted of 18 contiguous slices covering the entire brain every 2 s. To minimize susceptibility artifact, a separate acquisition lasting 1–2 min was needed for phase maps to correct for distortion in echoplanar images (Alsop, 1995). We also inspected raw data of individual subjects to check for artifacts. Raw data were stored by the MRI computer on DAT tape and then processed off-line.

Initial data processing was carried out with Interactive Data Language (Research Systems) on a Sun (Santa Clara, CA) Ultra 60 workstation. Raw image data were reconstructed using a 2D FFT with a procedure that minimized artifact due to magnetic field inhomogeneities. Individual subject data were then prepared for statistical parametric mapping (SPM 99) developed by the Wellcome Department of Cognitive Neurology (Frackowiak et al., 1997). This system, operating on a Matlab platform, combines raw images from subjects into a statistical t map. Briefly, the images in each subject’s time series were registered to the initial image in the series. The images were then aligned to a standard coordinate system (Talairach and Tournoux, 1988). The data were spatially smoothed with a 12-mm Gaussian kernel to account for small variations in the location of activation and sulcal anatomy across subjects. Low-pass temporal filtering was implemented to control auto-correlation with a first-order auto-regressive method.

We used an event-related approach to data acquisition. Each event was 12 s in length, thus capturing most of the hemodynamic response function associated with each stimulus. Hence, the initial images acquired for each 12-s trial were used in the analyses. The exclusive use of the initial image per trial also insured that activation results would not be affected by between-condition differences in time until the button-press response. Event-related data acquisition allowed us to group and compare images corresponding to specific stimuli. For the training phase, we assessed condition-specific cerebral activation. We also examined changes in activation with increased exposure to stimuli during the course of training by analyzing activation for the first eight and last eight of the 40 training trials (i.e., the first and fifth quintiles). Thus, we used a condition (rule-based, similarity-based) × learning phase (first quintile, last quintile) factorial design. For the test phase, we grouped MEMBERS and HIGH-DISTORTION stimuli. This was because their membership status was relatively transparent, according to similarity-based as well as by rule-based processing. Hence, even though the two categorization conditions elicited different decision-making processes, MEMBERS should generally be endorsed to comparable extents across conditions, and HIGH-DISTORTION should stimuli should generally be rejected to comparable extents across conditions. To help ensure that patterns of activation reflected subjects’ employing the categorization processes of interest, we analyzed activation only for accurate responses to these stimuli. As shown below, this included the vast majority of MEMBER and HIGH-DISTORTION stimuli in both conditions. We included all LOW-DISTORTION stimuli in our analyses, regardless of subjects’ category judgments, since membership or non-membership status was equally plausible by similarity-based criteria. Moreover, as shown below, proportions of member/nonmember judgments for LOW DISTORTION items differed depending on the categorization condition, and we did not want our imaging analyses to be biased by the inclusion of different numbers of a particular type of stimulus across conditions. Test data were analyzed with a condition (rule-based, similarity-based) × stimulus (MEMBER/HIGH-DISTORTION, LOW-DISTORTION) factorial design. The resulting statistical contrasts were converted to z-scores for each compared voxel. We used a statistical threshold of $P < 0.001$ uncorrected for multiple comparisons. The extent criterion was set to a 20 voxel threshold.

Results

Training phase

Behavioral results

Participants were accurate at judgments MEMBERS under both training conditions. Accuracy in the Rule training condition was 96% (±6%), which differed from chance [$t(10) = 28.09; P < 0.001$]; accuracy in the Similarity training condition was 85% (±9%), likewise differing from chance [$t(8) = 11.08; P < 0.001$]. The higher accuracy for identifying MEMBERS in the Rule condition relative to the Similarity condition reached significance [$t(17) = 3.0, P = 0.009$]. Moreover, participants took significantly longer to respond in the Rule condition [3946 ms ± 914] than in the Similarity condition [2886 ms ± 1003; t(18) = 2.5, P = 0.02]. These results are consistent with the qualitative differences between the two categorization processes: Rule-based training, which depends on working memory and executive processes such as selective attention, appears to be more cognitively demanding, and yields judgments with a higher rate of identifying MEMBERS.

We also compared judgment accuracy during the first and fifth quintiles for each condition. For the Rule condition, accuracy was 89% during the first quintile and 100% during the fifth. First quintile accuracy differed from 100% in a one-sample $t$ test, $t(10) = 2.83, P < 0.02$. (We could not use paired $t$ tests to directly compare first versus fifth quintile accuracy in the Rule condition because the 100% accuracy in the fifth quintile had no variance.) For the Similarity condition, accuracy was 81% during the first quintile and 94% during the fifth. These differed significantly by paired $t$ test, $t(8) = 3.62, P = 0.007$. Thus, participants became more successful at correctly identifying the MEMBER by the latter part of the session in both training conditions.

Imaging results

We analyzed training data with a condition (Rule, Similarity) × learning phase (early, late) factorial analysis. Table 1 summarizes the activation peaks using Talairach coordinates, the associated Brodmann areas, and the $z$ scores of the heights for each activation peak during training. We found significant differences for the main effect of condition, as illustrated in Fig. 2. Fig. 2 Panel A shows...
that the Rule condition, relative to the Similarity condition, activated left inferior frontal, left anterior cingulate, left inferior parietal, bilateral parietal/occipital, and left thalamus regions. By comparison, Fig. 2 Panel B shows that the Similarity condition, relative to the Rule condition, activated bilateral temporal–parietal and bilateral anterior prefrontal regions. These findings suggest different patterns of activation depending on the training condition. We also found that activation for each training condition varied depending on the early versus later phase of training. Thus, the factorial analysis showed significant main effects for the early versus late portions of the learning phase, that is, changes over the course of acquisition collapsed across conditions. As summarized in Table 1, the early portion, relative to the late portion of the learning phase, activated right superior parietal and right superior temporal areas. The late portion of training, relative to the early portion, activated left temporal–parietal, left posterior cingulate, right temporal/insula, and left caudate regions. The overall interaction effects in the factorial analysis involved bilateral striatum including right putamen and left caudate ($x = 20, y = 4, z = -4$; $z$ score = 4.02), and left inferior frontal cortex in a manner that approached significance ($x = -48, y = 4, z = -4$; $z$ score = 2.88). More detailed evaluation of the interaction effect is illustrated in Fig. 2. As shown in Fig. 2 Panel C, greater activation later in the Rule training phase, relative to earlier, was seen in left inferior frontal cortex. For the Similarity condition, Fig. 2 Panel D shows relatively greater activation later in training compared to the early in bilateral temporal–parietal, left posterior cingulate, and right striatal regions. Thus, distinct activation patterns appeared to become consolidated over time in part in different brain regions, depending on the specific training condition. These patterns suggest that left inferior frontal cortex supports increased use of executive resources such as working memory as rule-based training progressed. In contrast, later activation of temporal–parietal, posterior cingulate, and striatal regions during similarity-based training suggests an increasing tendency towards feature configuration and integration along with successful recall of previously seen training items.

**Test phase**

**Behavioral results**

We calculated “% yes” scores, that is, how often participants granted membership to each item type (i.e., MEMBER, LOW DISTORTION, and HIGH DISTORTION items) at test. Performance during the test phase is summarized in Fig. 3. As can be seen, patterns of responses differed across conditions and types of

---

**Table 1**

Distribution of activation during the training phase of category acquisition

<table>
<thead>
<tr>
<th>Anatomic area</th>
<th>Brodmann area</th>
<th>Talairach coordinates</th>
<th># Voxels</th>
<th>Z score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>x  y  z</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MAIN EFFECTS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RULE minus SIMILARITY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L inferior frontal</td>
<td>6/44</td>
<td>-56 0 36</td>
<td>17</td>
<td>3.47</td>
</tr>
<tr>
<td>L anterior cingulate</td>
<td>24/32</td>
<td>-8 8 40</td>
<td>69</td>
<td>4.11</td>
</tr>
<tr>
<td>L inferior parietal</td>
<td>40</td>
<td>-44 -48 40</td>
<td>26</td>
<td>3.43</td>
</tr>
<tr>
<td>L parietal/occipital</td>
<td>19/7</td>
<td>-8 -76 32</td>
<td>236*</td>
<td>4.23</td>
</tr>
<tr>
<td>R superior parietal</td>
<td>7</td>
<td>28 -64 52</td>
<td>236*</td>
<td>4.64</td>
</tr>
<tr>
<td>L thalamus</td>
<td>-8 -20 4</td>
<td>39</td>
<td>3.52</td>
<td></td>
</tr>
<tr>
<td>SIMILARITY minus RULE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L temporal–parietal</td>
<td>39</td>
<td>-40 -72 12</td>
<td>25</td>
<td>3.59</td>
</tr>
<tr>
<td>R temporal–parietal</td>
<td>39</td>
<td>56 -60 12</td>
<td>36</td>
<td>3.32</td>
</tr>
<tr>
<td>R anterior prefrontal</td>
<td>8</td>
<td>20 28 44</td>
<td>12</td>
<td>3.30</td>
</tr>
<tr>
<td>L anterior prefrontal</td>
<td>10</td>
<td>16 56 8</td>
<td>23</td>
<td>3.20</td>
</tr>
<tr>
<td>EARLY minus LATE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R superior parietal</td>
<td>7</td>
<td>4 -72 48</td>
<td>30</td>
<td>3.28</td>
</tr>
<tr>
<td>R superior temporal</td>
<td>22</td>
<td>60 -28 4</td>
<td>13</td>
<td>3.13</td>
</tr>
<tr>
<td>LATE minus EARLY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L temporal–parietal</td>
<td>40/22</td>
<td>-44 -20 20</td>
<td>241</td>
<td>4.09</td>
</tr>
<tr>
<td>L posterior cingulate</td>
<td>23/29</td>
<td>-8 -36 20</td>
<td>55</td>
<td>3.43</td>
</tr>
<tr>
<td>R superior temporal/insula</td>
<td>42</td>
<td>28 -20 16</td>
<td>91</td>
<td>3.31</td>
</tr>
<tr>
<td>L thalamus</td>
<td>-20 -24 4</td>
<td>30</td>
<td>3.13</td>
<td></td>
</tr>
<tr>
<td><strong>INTERACTION EFFECTS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RULE: EARLY minus LATE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R superior parietal</td>
<td>7</td>
<td>8 -72 48</td>
<td>40</td>
<td>3.65</td>
</tr>
<tr>
<td>L parietal–occipital</td>
<td>19</td>
<td>-32 -88 16</td>
<td>15</td>
<td>3.48</td>
</tr>
<tr>
<td>L thalamus</td>
<td>-4 -8 8</td>
<td>85</td>
<td>3.27</td>
<td></td>
</tr>
<tr>
<td>RULE: LATE minus EARLY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L inferior frontal</td>
<td>6/44/45</td>
<td>-44 0 4</td>
<td>82</td>
<td>3.06</td>
</tr>
<tr>
<td>SIMILARITY: LATE minus EARLY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L temporal–parietal</td>
<td>40/22</td>
<td>-52 -20 24</td>
<td>279</td>
<td>3.99</td>
</tr>
<tr>
<td>R temporal–parietal</td>
<td>40/22</td>
<td>56 -16 20</td>
<td>29</td>
<td>3.40</td>
</tr>
<tr>
<td>L posterior cingulate</td>
<td>29/23</td>
<td>-8 -36 20</td>
<td>103</td>
<td>4.26</td>
</tr>
<tr>
<td>R striatum</td>
<td>20 0 -4</td>
<td>49</td>
<td>3.83</td>
<td></td>
</tr>
</tbody>
</table>

*These regions were part of the same cluster.*
items in a manner consistent with the two qualitatively distinct categorization processes. The Rule condition thus showed a sharp category boundary, reflecting the requirement that only items meeting the three-out-four feature criteria be endorsed as members; the Similarity condition showed graded responses across types of items with declining membership endorsement as the number contributing of features shared with the prototype declined. A repeated measures ANOVA, with a condition (Rule, Similarity) × stimulus type (MEMBER, LOW DISTORTION, HIGH DISTORTION) design, revealed a main effect of stimulus type \( F(2,36) = 371.5, P < 0.001 \) and a significant condition × stimulus type interaction \( F(2,36) = 22.0, P < 0.001 \). There was no main effect of condition \( F(1,18) = 2.26, \text{ns} \). The \( t \) tests revealed that judgments differ significantly across conditions for all three stimulus types: participants in the Rule condition responded “yes” more often to MEMBERS \( t(18) = 4.0, P = 0.001 \) and less often to LOW DISTORTION items \( t(18) = 3.1, P = 0.006 \) and HIGH DISTORTION items \( t(18) = 2.5, P = 0.02 \) than participants in the Similarity condition. Linear trend analyses confirmed the different shapes of the performance curves: We saw “linearity” in both the Rule condition \( F(1,30) = 777.8, P < 0.01 \) and the Similarity condition \( F(1,24) = 78.2, P < 0.01 \). However, “deviation from linearity” was seen in the Rule condition \( F(1,30) = 157.3, P < 0.01 \) but not the Similarity condition \( F(1,24) = 3.2, \text{ns} \). These findings confirmed that participants respond differently across conditions for the different types of items, reflecting the sharp categorical effect of rule-based categorization and the graded responses of similarity-based categorization.

There was a trend towards longer reaction times in the Rule condition (2453 ms) relative to the similarity condition (2144 ms) that did not reach significance, \( t(18) = 0.79, \text{ns} \). Reaction times did not differ among item types in the Similarity condition \( F(2,8) = 0.86, \text{ns} \). However, reaction times did differ among different item types in the Rule condition \( F(2,10) = 10.1, P = 0.001 \). While reaction times for MEMBERS and HIGH DISTORTION items were equivalent \( t(10) = 1.71, \text{ns} \), participants responded more slowly to LOW DISTORTION items than to either MEMBERS \( t(10) = 2.31, P < 0.05 \) or to HIGH DISTORTION items \( t(10) = 7.09, P < 0.001 \). This finding suggests that judging MEMBERS
and HIGH DISTORTION items may have been facilitated by prior exposure to such items during training.

We compared judgment accuracy for old versus new items in both test conditions; that is, MEMBERS that had been seen five times during training with MEMBERS seen for the first time at test, and similarly, HIGH DISTORTION items seen during training with those newly seen at test. To ensure that the old and new stimulus items were comparable, we included as new items only those with exactly three contributing features appropriately matching or not matching the prototype; that is, we did not include test MEMBERS matching the prototype on all four features or test HIGH DISTORTION items matching it on none. Accuracy for old versus new items was equivalent for each item type in each condition: Rule MEMBERS, \( t(10) = 1.00, P = 0.34 \); HIGH DISTORTION, \( t(10) = 1.18, P = 0.86 \); Similarity MEMBERS, \( t(8) = 0.67, P = 0.52 \); HIGH DISTORTION, \( t(8) = 0.33, P = 0.75 \).

This suggests that participants were using categorization processes at test for all items, rather than relying on recognition of training items.

**Imaging results**

We analyzed categorization data with a condition (rule-based, similarity-based) \times stimulus type (MEMBERS+HIGH DISTORTION, LOW DISTORTION) factorial analysis. Table 2 summarizes the activation peaks using Talairach coordinates, the associated Brodmann areas, and the \( z \) scores of the heights for each activation peak during the test phase. The activations associated with the main effects are illustrated in Fig. 4. Resembling the activation patterns seen during training, the two conditions differed in activated areas during the test phase of the study. Panel A reveals activation during the Rule condition, relative to the Similarity condition, in left inferior frontal and left anterior cingulate regions. Panel B shows activation for the Similarity condition, relative to the Rule condition, in bilateral temporal–parietal regions that is more prominent on the left, in bilateral anterior prefrontal regions, and in right posterior cingulate cortex.

We also found that these condition-specific activations differed somewhat, depending on the types of items being judged. Thus, the main effect for item type revealed activation for LOW DISTORTION items, compared to MEMBERS+HIGH DISTORTION items, in left dorsolateral prefrontal, bilateral anterior cingulate, and bilateral striatal areas, as well as bilateral inferior frontal areas that is marginal on the left. The overall interaction effect for the factorial analysis showed left inferior frontal (\( x = -40, y = 28, z = 0 \); \( z \) score = 2.82) and right dorsolateral prefrontal (\( x = 32, y = 40, z = 20 \); \( z \) score = 2.90) recruitment. We examined interaction effects more closely, as summarized in Table 2 and illustrated in Fig. 4. For MEMBERS+HIGH DISTORTION items, Panel C shows that the Rule condition, relative to the

<table>
<thead>
<tr>
<th>Anatomic area</th>
<th>Brodmann area</th>
<th>Talairach coordinates</th>
<th># Voxels</th>
<th>Z score</th>
</tr>
</thead>
<tbody>
<tr>
<td>RULE minus SIMILARITY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L inferior frontal</td>
<td>6/44</td>
<td>-60 4 28</td>
<td>59</td>
<td>3.84</td>
</tr>
<tr>
<td>L anterior cingulate</td>
<td>24/32</td>
<td>-4 8 44</td>
<td>61</td>
<td>3.50</td>
</tr>
<tr>
<td>SIMILARITY minus RULE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L temporal–parietal</td>
<td>39/22</td>
<td>-32 -64 8</td>
<td>261</td>
<td>5.08</td>
</tr>
<tr>
<td>R posterior temporal</td>
<td>21</td>
<td>40 -44 4</td>
<td>44</td>
<td>3.08</td>
</tr>
<tr>
<td>L anterior prefrontal</td>
<td>10</td>
<td>-24 56 20</td>
<td>20</td>
<td>3.39</td>
</tr>
<tr>
<td>R anterior prefrontal</td>
<td>10</td>
<td>16 44 -4</td>
<td>112</td>
<td>3.08</td>
</tr>
<tr>
<td>R posterior cingulate</td>
<td>23/29/30</td>
<td>8 -48 8</td>
<td>24</td>
<td>3.33</td>
</tr>
<tr>
<td>MEMBER + HIGH DISTORTION minus LOW DISTORTION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No significant contrasts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOW DISTORTION minus MEMBER + HIGH DISTORTION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L inferior frontal</td>
<td>6</td>
<td>-32 -12 48</td>
<td>25</td>
<td>2.90</td>
</tr>
<tr>
<td>R inferior frontal</td>
<td>6</td>
<td>40 -16 40</td>
<td>186</td>
<td>3.34</td>
</tr>
<tr>
<td>B anterior cingulate</td>
<td>24/32</td>
<td>0 4 48</td>
<td>186</td>
<td>3.28</td>
</tr>
<tr>
<td>L dorsolateral prefrontal</td>
<td>46</td>
<td>-52 24 24</td>
<td>54</td>
<td>3.20</td>
</tr>
<tr>
<td>L striatum</td>
<td>-16</td>
<td>0 0</td>
<td>45</td>
<td>3.32</td>
</tr>
<tr>
<td>R striatum</td>
<td>28</td>
<td>20 0</td>
<td>39</td>
<td>3.24</td>
</tr>
<tr>
<td>INTERACTION EFFECTS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RULE minus SIMILARITY: MEMBER + HIGH DISTORTION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L inferior parietal</td>
<td>40</td>
<td>-52 -36 40</td>
<td>49</td>
<td>3.14</td>
</tr>
<tr>
<td>L parietal–occipital</td>
<td>19</td>
<td>-32 -76 40</td>
<td>23</td>
<td>3.40</td>
</tr>
<tr>
<td>L anterior cingulate</td>
<td>24/32</td>
<td>-4 8 40</td>
<td>24</td>
<td>2.99</td>
</tr>
<tr>
<td>SIMILARITY minus RULE: MEMBER + HIGH DISTORTION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L temporal–parietal</td>
<td>39/22</td>
<td>-32 -64 8</td>
<td>1411</td>
<td>4.75</td>
</tr>
<tr>
<td>R temporal–parietal</td>
<td>39/22</td>
<td>32 -60 8</td>
<td>1411</td>
<td>4.46</td>
</tr>
<tr>
<td>RULE minus SIMILARITY: LOW DISTORTION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L inferior frontal</td>
<td>6/44</td>
<td>-64 0 20</td>
<td>91</td>
<td>3.19</td>
</tr>
<tr>
<td>SIMILARITY minus RULE: LOW DISTORTION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( ^* \) These regions were part of the same cluster.
Similarity condition, activated left inferior parietal, left parietal–occipital, and left anterior cingulate regions. For LOW DISTORTION items, Panel D shows activation for the Rule condition minus the Similarity condition in left inferior frontal cortex. Panel E shows that the Similarity condition, compared to the Rule condition, activated bilateral temporal–parietal cortex for MEMBERS+HIGH DISTORTION items. Panel F shows activation in left temporal–parietal cortex for LOW DISTORTION items in the Similarity condition minus the Rule condition. These findings suggest that relatively distinct patterns of activation were associated with each categorization condition, consistent with support for executive resources for rule-based categorization and perceptual feature integration for similarity-based categorization. In addition, the condition-specific patterns vary somewhat depending on the type of stimulus item being judged, suggesting that, for each condition, different strategic shifts are required for
judging items whose membership status is readily apparent versus judging equivocal items. We expected our participants to make category judgments at test in accordance with how they had been trained. Overlap in areas of activation within condition from training to test would reflect this continuity. We looked for such overlap by comparing the Talairach coordinates of peak activation between Rule minus Similarity, and between Similarity minus Rule, for training and for test. In both cases, two areas were observed for which all three respective coordinates differed by less than 12 mm, our unit of spatial smoothing, between the corresponding training and test phases. As can be seen in Tables 1 and 2, Rule-based processing, relative to Similarity-based, recruited left inferior frontal (BA 6/44) and left anterior cingulate (BA 24/32) regions, suggesting that executive processing and monitoring of competing responses were common to both rule-based learning and subsequent test. The high levels of accuracy in category membership judgments in this resource-demanding task are consistent with a constant level of response monitoring. Similarity-based processing, relative to Rule-based, recruited left temporal–parietal (BA 39) and left anterior prefrontal (BA 10) regions, suggesting that overall feature configuration and recall of previously encountered exemplars were common to both similarity-based learning and subsequent test. Our experimental design was not conducive to subtractions between the training and test phases. The stimulus displays differed; for instance, subjects saw multiple items along with written material during training, but only single items at test. The tasks also differed; for instance, subjects chose between two repeatedly presented items during training but evaluated unique individual items at test. In addition, training trials were consistent, while test trials involved evaluating different item types in a random pattern. These confounds would make it impossible to reliably interpret the results of such subtractions, and hence we cannot draw conclusions concerning the non-overlapping areas of activation between training and test.

Discussion

We taught two groups of participants a novel semantic category using two different categorization strategies. Participants’ behavioral responses were consistent with the qualitative distinctions between rule-based and similarity-based categorization processes that have been demonstrated elsewhere (Allen and Brooks, 1991; Grossman et al., 2003b; Koenig et al., 2002; Rips, 1989; Smith and Sloman, 1994): Categorization judgments following rule-based training thus resulted in a clearly defined all-or-none category, while categorization judgments following similarity-based training resulted in graded membership that reflected an item’s resemblance to the prototype. We monitored regional cerebral activity with fMRI while participants first learned a novel semantic category under rule-based or similarity-based conditions, and while they subsequently categorized these items in accordance with their training. We found distinct condition-specific activation profiles during training, and unique patterns of activation during the subsequent test phase that overlapped with the corresponding training activation profiles. These findings are consistent with the hypothesis that semantic categorization can be performed by at least two qualitatively different processes, and that these processes rely on distinct large-scale neural networks. We discuss each of these processes and the associated brain activation patterns below.

Activation patterns during rule-based semantic categorization

Rule-based acquisition

There is considerable evidence that brain regions associated with working memory and executive processes such as selective attention are recruited during rule-based categorization. The rule condition thus activated dorsal portions of left inferior frontal cortex (BA 6 and 44). This region shows activation during many tasks involving verbal working memory (Smith and Jonides, 1999; Smith et al., 1998a,b), and appears to be associated in particular with maintaining information in an active state in working memory (Awh et al., 1996; Smith and Jonides, 1999). Recruitment may serve to maintain in verbal working memory the features needed to diagnose category membership in accordance with rules. Involvement of the prefrontal cortex has been found in single-cell recordings of nonhuman primates (i.e., monkeys), apparently specific to the category membership status of items in recently learned all-or-none categories (Miller et al., 2002). The important contribution of working memory is emphasized by the significantly increasing left inferior frontal activation during the course of rule-based training. This is in keeping with the increased accuracy of judgments observed in the latter part of the training phase, and suggests that the need for working memory to employ the rules continued despite subjects’ increasing familiarity with the visual appearance of the stimuli. Brain regions associated with working memory are often lateralized depending on the verbal or visual perceptual–spatial nature of the material, with right-sided dominance for spatial working memory and left-sided activation for verbal working memory (Smith, 1997; Smith et al., 1996). An example of such lateralization is apparent in the left-sided activation seen during rule-based decision-making with verbally presented stimuli (Grossman et al., 2002a,b), and in the right-sided activation during rule-based decision-making with images (Patalano et al., 2001). In the present study, although the novel category was visual, left-lateralized activation may reflect the presentation of the features relevant to the rule as verbal descriptions.

We also observed left anterior cingulate activation during rule-based training. Some studies associate this brain region with components of attention important for selecting and maintaining sensitivity to specific features (Benedict et al., 1998; Corbetta et al., 1993; Coull et al., 1996), compatible with Posner’s executive component of attention (Posner and Petersen, 1990). Thus, anterior cingulate may be important for selectively attending to specific, critical features that determine membership in the new category. Another line of work emphasizes the importance of the anterior cingulate for governing the response to competing candidate choices during task performance (Braver et al., 2001; Carter et al., 1998; MacDonald et al., 2000), although competition between response alternatives was also present during similarity-based training. Regardless of the specific basis for anterior cingulate recruitment, activation of this area and inferior frontal cortex emphasizes the important role of executive processes during rule-based categorization.

We also found left thalamic activation during rule-based training. This region also appears to play an important role in selective (Frith and Friston, 1996; LaBerge, 1995; Shulman et al., 1997), a necessary component in the feature-selection process inherent in rule-based, but not similarity-based, categorization. The connectivity pattern of the thalamus is consistent with this structure serving an attentional gating role, since it appears to modulate the
projection of sensory-motor information to the cerebral cortex and cortical-cortical connectivity.

Parietal activation in BA 7 was observed during rule-based learning, particularly during the early phase of training. Recruitment of this parietal area has been associated with perception of the spatial relationships among the feature elements of the stimuli (Grady et al., 1994; Horwitz et al., 1992; Wilkinson et al., 2002), and with visual selective attention (Bushnell et al., 1981). This activation may diminish over time as subjects accumulate familiarity with the spatial configuration of the stimuli, similar to observations of diminishing cortical activation during learning in studies of repetition priming (Buckner and Koutsaa, 1998; Buckner et al., 1998; Raichle et al., 1994). It is noteworthy that diminution during the course of learning was also found in the thalamus. This finding suggests that the attention-gating role played by the thalamus may be linked to diminishing cortical activation during learning. The link between diminished activation in these areas and learning over the course of training is consistent with the increase in judgment accuracy that participants showed during the final training trials.

**Rule-based categorization**

The pattern of activation in the test phase for rule-based categorization judgments appears to represent recruitment of a large-scale neural network that partially incorporates that which was recruited during training. We observed recruitment of dorsal portions of left inferior frontal cortex during the rule condition. Inferior frontal activation suggests that participants maintained the rules guiding categorization decisions in verbal working memory (Awh et al., 1996; Smith and Jonides, 1999). Inferior frontal activation was more evident during categorization of LOW DISTORTION than MEMBER + HIGH DISTORTION items. These items, which had two of the four requisite features, may depend more on an exhaustive analysis of the rules maintained in working memory because endorsing them as members requires inspection of all four relevant features, in contrast with items which can be accepted when three defining features have been identified. It is also possible that assessing MEMBER and HIGH DISTORTION items was facilitated because of their resemblance to items that were repeatedly seen during training, in contrast to LOW DISTORTION items. Hence, subjects may have had to rely more exclusively on the application of rules in classifying the unfamiliar LOW DISTORTION items. The greater reaction times for these items are consistent both of these interpretations.

Anterior cingulate, also recruited during rule-based acquisition, was again recruited in the left hemisphere during rule-based categorization relative to similarity-based categorization. Previous studies involving rule-based categorization also demonstrated anterior cingulate activation (Elliot et al., 1999; Rao et al., 1997; Savage et al., 2001). Anterior cingulate activation was more strongly associated with rule-based categorization of MEMBER + HIGH DISTORTION than with LOW DISTORTION items. Together with the left parietal activation that was also seen during rule-based decisions about MEMBER + HIGH DISTORTION items, these observations begin to emphasize the contribution of selective attention to the diagnostic value of specific perceptual features during rule-based processing of objects.

Our findings thus suggest a pivotal role for executive resources during rule-based categorization. Evidence for this first emerges during acquisition. An anterior cingulate-thalamic-parietal network appears to support selective attention to specific perceptual features, and inferior frontal cortex is recruited increasingly during rule-based learning to maintain rules in an activated state in working memory. During subsequent categorization, anterior cingulate and parietal cortex continue to play a particularly important role in selectively attending to perceptual features of stimuli, and inferior frontal cortex to maintain rules in working memory.

**Activation patterns during similarity-based semantic categorization**

**Similarity-based acquisition**

Similarity-based training was associated with bilateral anterior prefrontal activation. This region is recruited in retrieval from long-term memory (Buckner and Koutsaa, 1998; Buckner et al., 1998; Lepage et al., 2002). We found activation in a posterior cingulate distribution during similarity-based training as well. As with anterior prefrontal cortex, activation of this area has been associated with retrieval from episodic memory, particularly when retrieval has been successful (Kapur et al., 1995; Rugg et al., 1998; Tulving et al., 1996). It is noteworthy that posterior cingulate recruitment increased over the course of learning, as did the accuracy of participants’ judgments. Together with anterior prefrontal activation, posterior cingulate recruitment during similarity-based training, but not rule-based training, is consistent with subjects’ increasingly depending on the recall of specific exemplars during similarity-based training; that is, unlike rule-based training in which each trial imposes the same rules on successive training items, similarity-based training involves the gradual formation of the category with exposure to a variety of members.

We also observed bilateral temporal-parietal activation. Some studies have emphasized the role of this area in the overall configuration of features that may be important for similarity-based judgments (Wilkinson et al., 2002). Relatedly, other work associates activation of this region with object meaning (Grossman et al., 2003a; Martin et al., 1995; Owen et al., 1996a,b). Temporal-parietal cortex is a heteromodal association region with a connectivity pattern that involves reciprocal projections with modality-specific association cortices (Mesulam et al., 1977; Pandya and Kuypers, 1969; Seltzer and Pandya, 1978). Activation in this brain region thus appears to be well suited to integrating multiple features of a complex stimulus such as a novel category for the purpose of deriving object meaning. The increasing activation of these temporal-parietal regions over the course of similarity-based training appears to emphasize the importance of this integrative function in similarity-based acquisition strategies.

**Similarity-based categorization of a semantic category**

The brain regions associated with similarity-based categorization during the test phase overlapped the areas activated by subjects during similarity-based training. Posterior cingulate and anterior prefrontal activation thus were also evident during categorization. This may continue to reflect dependence on successful retrieval of previously encountered exemplars, that is, the training items and the presented prototype, from memory (Braver et al., 2001; Buckner and Koutsa, 1998; Lepage et al., 2002; Rugg et al., 1998; Tulving et al., 1996). The similarity condition was also associated with bilateral temporal-parietal
recruitment. This region appears to play a continuing role in integrating multiple features crucial for overall similarity-based comparisons (Grossman et al., 2002a,b, 2003a).

Our observations thus suggest that similarity-based semantic categorization activates brain regions important for feature integration and comparisons with previously encountered exemplars. During both similarity-based acquisition and subsequent categorization, this network was seen to include anterior prefrontal, posterior cingulate, and temporal–parietal regions.

**Activation patterns common to rule- and similarity-based categorization**

Several other brain regions associated with executive resources were recruited during judgments of LOW DISTORTION items relative to MEMBER + HIGH DISTORTION items across both the rule-based and similarity-based conditions at test. These include left dorsolateral prefrontal cortex and bilateral striatum. MacDonald et al. (2000) have suggested that left dorsolateral prefrontal cortex selectively attends to information relevant to task demands, and that such activation can be increased by monitoring of response conflict by anterior cingulate cortex. We suggested above that, for rule-based categorization, LOW DISTORTION items may put greater demands on maintenance of the specific rules in working memory because all four requisite features must be inspected. In the similarity-based condition, subjects may have become sensitive to the conflicting status of LOW DISTORTION test items that equally resemble MEMBER and HIGH DISTORTION items, and subsequently resolved the category membership status of the LOW DISTORTION items in a manner mediated in part by dorsolateral prefrontal cortex. These “equivocal” stimulus items, then, may have placed greater demands on executive processes in both categorization conditions, albeit for different decision-making strategies.

Some researchers have associated striatal activation with rule-based categorization (Ashby and Ell, 2001; Ashby et al., 1991). For example, caudate activation was seen in a study of rule-based card sorting (Rao et al., 1997). To the extent that judging stimuli typical of neither members nor nonmembers is inherently resource-demanding, LOW DISTORTION stimuli may require activation of the striatum regardless of the rule-based or similarity-based nature of categorization that we elicited.

**Comparison with previous rule and similarity imaging studies**

There are few neuroimaging studies directly comparing rule- and similarity-based categorization. Most studies that examine one or both of these processes differ critically from the present study, making informative between-study comparisons difficult. For instance, Tracy et al. (2003) compared categorization by either family resemblance (i.e., similarity) or criterion attribute (i.e., an exhaustive rule). They observed medial parietal, inferior frontal, and anterior temporal activation in their rule-based task, and extrastriate and cerebellar activation in their similarity-based task. However, their stimuli were letter strings; the categories were structured differently for each condition; subjects had to deduce the category structure; and finally, the same subjects participated in both conditions, in alternate blocks. In another example, Reber et al. (1998b) found decreased posterior occipital activity associated with classifying members in a resemblance-based categorization task. However, this was in comparison with a recognition memory task, rather than a rule-based task. In addition, the task used semantically meaningless dot patterns which subjects learned to categorize by an implicit, rather than an explicit, process. Hence, studies such as these seem unlikely to draw on the same neural mechanisms as the present study.

The fMRI study by Grossman et al. (2002a,b), described in Introduction, compared rule-based and similarity-based categorization of minimally described real world objects. Thus, that study differs from the current study in many ways, including its use of verbal stimuli, familiar categories (and hence no acquisition phase), and subject-derived rather than experimenter-imposed rules (e.g., participants must rely on their prior knowledge to realize that “quarters must be of a certain size,” and then reason accordingly). These differences would seem to account for activation discrepancies across the two studies. For instance, left dorsolateral prefrontal, left anterior cingulate, and left inferior parietal activation observed in both processing conditions in the Grossman et al. (2002a,b) study could reflect processing requirements that are common to both conditions in that study, but not in the present one. These could include resource-demanding reasoning, responses that must be monitored in the similarity condition because they are equally plausible and in the rule condition because they are potentially error-prone, and feature configuration necessary to make judgments about a verbally presented category. Nonetheless, the Rule condition, relative to the Similarity condition, activated left frontal areas, albeit more dorsal than those activated in the present study.

The PET imaging study by Patalano et al. (2001) seems closer in design to the present study, and also shows left inferior frontal activation during rule-based categorization that is not observed during similarity-based categorization. The discrepancies between other observed patterns of activation across the two studies could reflect the many methodological differences: The current study is event-related, and therefore activation patterns reported for one condition are relative to those in the other. The Patalano et al. study, which was PET rather than fMRI, contrasts activation with within-condition baseline tasks. That study found visual association cortex activity in both conditions; in the present study, any such activation common to both conditions would have been subtracted out. The Patalano et al. study’s similarity-based condition employed exemplars that were highly similar to training items (which consisted of two categories), with a large number of features in common, and no equivocal items. This may account for the lack of temporal–parietal activation in that study, which we suggest reflects the processes of feature configuration necessary to forming the new category in the current study. The strong item similarity, coupled with the several day delay between training and test in the Patalano et al. study, may account for the lack of posterior cingulate activity, which, in the present study, we attribute to the successful recall of previous exemplars: Memory for specific exemplars in the Patalano et al. study may not have been preserved over the lengthy delay, while such memories were presumably available in the present study. The Patalano et al. study’s rule condition consisted of items that either closely resembled or looked greatly unlike category members; these latter items presumably required that participants inhibit perceptual dissimilarity in order to follow the rules. In addition, participants had 3 s, in contrast to the current study’s 12 s, to make a decision at test. Rule-condition participants’ relatively low accuracy rates (78%, compared with 98% in the current study) suggest that they
were unsuccessful in applying rules at least 20% percent of the time, and it presumably cannot be known whether they attempted to apply the rules during those times, responded to perceptual similarity, or simply became inattentive. The current study includes only successful trials. This may account for the anterior cingulate activation observed in the current study, which we attribute to response monitoring.

Conclusion

The observations of the present study extend previous findings (Grossman et al., 2002a,b; Patalano et al., 2001) of distinct cerebral activation patterns for rule-based and similarity-based categorization to the acquisition and utilization of a new semantic category. Our training stimuli were reasonably matched across conditions for verbal and visual content, and contained the identical training items; our test stimulus displays were completely identical across conditions. Hence, qualitatively different responses across conditions, particularly in the test phase, argue in favor of multiple categorization processes. Our overall observation of distinct patterns of decision-making paired with distinct patterns of cortical activation thus is difficult to reconcile with approaches to semantic categorization advocating a single mechanism for all forms of categorization, with differing responses being attributed to varying difficulty of task demands (Nosofsky and Johansen, 2000; Nosofsky and Palmeri, 1998).

We found neural networks unique to each categorization process. It appears that the acquisition of a new category by a rule-based strategy is associated with an anterior cingulate–thalamic–parietal network important for selectively attending to specific perceptual features, and inferior frontal cortex that maintains rules in working memory during categorization. During subsequent categorization of stimuli following rule-based learning, selective attention to perceptual features, supported by anterior cingulate and parietal cortex, apparently plays a particularly important role in the categorization of items with readily apparent category membership. By comparison, categorization of LOW DISTORTION stimuli is associated more closely with left inferior frontal activation, presumably to maintain the rule-based criteria and evaluated features in an active state in working memory. Similarity-based categorization seems to consistently activate temporal–parietal regions, supporting the integration of features necessary for making overall comparisons, and an anterior prefrontal–posterior cingulate network to support successful retrieval from memory of previously encountered exemplars which determine the structure of the category.

References

Horwitz, B., Grady, C.L., Haxby, J.V., Schapiro, M.B., Ungerleider, L.G.,


