ELSEVIER

Research report

# Changing patterns of brain activation during category learning revealed by functional MRI

Deborah M. Little[a],*, Raymond Klein[a,b], Donna M. Shobat[a],
Erik D. McClure[a], Keith R. Thulborn[a]

[a]Center for Magnetic Resonance Research, College of Medicine, University of Illinois, Room 1193, MC 707,
1801 W. Taylor Street, Chicago, IL 60612, United States
[b]Department of Psychology, Dalhousie University, Halifax, Nova Scotia, Canada

## Abstract

Functional magnetic resonance imaging (fMRI) was used to investigate neural changes as a function of category learning in normals ($n=8$). Subjects were trained to classify patterns of dots into four categories over 4 consecutive days. fMRI monitored the changes that occurred during learning prior to training and then following each training session. During fMRI, subjects determined whether two patterns of dots were members of the same category. The behavioral changes that occurred as a result of the training were observed as increases in response accuracy within shortened response times. fMRI illustrated initial increases in volumes of activation distributed across the known visuospatial processing networks. The regions affected by learning were identified as those involved in the planning and execution of eye movements (frontal and supplementary eye fields, FEF and SEF), spatial attention (superior and inferior parietal lobules, SPL and IPL) and visual processing (primary, secondary, and tertiary visual cortices). The volumes of activation then decreased as training progressed further. Of the two proposed mechanisms for learning, that of strengthened connectivity on a given network and that of selection of different networks, our data supports the former.
© 2004 Elsevier B.V. All rights reserved.

## 1. Introduction

The neural changes within the human brain during category learning have yet to be fully documented [3,22]. Category learning has been well-characterized behaviorally [28,41], with a variety of different materials and methods that range in breadth from the more ecologically valid (i.e., classification of birds; Ref. [4]) to more experimentally defined (i.e., random dot patterns; Ref. [36]) materials. Although category learning is an important skill, the neuronal response that underlies it is not well understood.

Learning is the acquired behavioral change associated with improved performance on a given task [12,16,33]. The behavioral changes associated with learning are commonly observed as increases in accuracy within shortened response times [2,25]. The rate or speed of learning can be manipulated by changing task difficulty and by including or excluding feedback [1,27].

Unlike category learning, the neural changes associated with other areas of skill learning (such as learning a motor sequence) have been fairly well explored (for reviews, Refs. [16,33]). The neuronal response to successful skill learning can be described by and mapped onto multiple neural networks. The specific network is dependent upon the skill required to complete the task (motor, visual–spatial or

auditory). The changes that occur in these networks during learning can be described by one of two possible mechanisms, or an interaction between the two [32].

The first possible mechanism proposed by Petersen et al. [32] is that the neuronal network subserving behavioral improvements on a given task becomes more defined or efficient. With this increased efficiency, the associations between the regions of this network also strengthen (e.g., Ref. [47]). The second possible mechanism is that with successful learning, a more task-appropriate network is selected from a number of different networks (e.g., Ref. [17]). This second mechanism would suggest that with practice, a different set of cognitive operations is recruited to more efficiently accomplish the directed goal. The literature supporting the existence of both of these mechanisms is now presented briefly.

For the first mechanism, the establishment of more efficient connections would result in more specialized, or more well-defined, regions of brain activation. These regions of activation would implicate the same network whether early or late in the training program. Such neuronal changes, observed as reductions in the volume of activation as training progresses, have been reported for learning to associate an object with a location [5] and for artificial grammar learning [43]. Although each of these reports shares the finding that the same network is implicated following practice, there is heterogeneity in the time course of these changes [13,19].

The second mechanism, recruitment of a different network, might be realized in two ways. The first is that an entirely new network is utilized as learning progresses. The second outcome is that, rather than the creation of a new network, the newly acquired skill is incorporated as a component of an already existing expert skill and therefore mapped onto an already defined region known for its response to that expert skill. One example of this second outcome is that when subjects were trained to classify novel 3D characters by identifying specific features, similar to the requirements of face recognition, there was an increase in neuronal activity in fusiform gyrus, the same region that responded to faces [17]. Additional support for this mechanism was observed when subjects were trained to read mirror-reversed text. As training on mirror-reading progressed, the same regions active in normal lexical processing and word recognition were activated [34,35].

Regardless of the underlying mechanism, the observation has been made that the neural changes that underlie learning are dynamic and evolve as behavioral performance changes. One example involves changes in the volume of activation with learning. Specifically, as training progressed across a single session, the volumes of activation across all regions have been reported to show overall decreases when associating an object with a location [5], both increases and decreases in activation dependent upon the region involved when learning to associate a verbal cue with a visual stimulus [29], and overall initial increases followed by overall decreases in activation across all regions during motor sequence learning [46]. These results, when taken together, illustrate that the neural response that underlies learning is dynamic and varies depending upon the skill to be learned. For those studies that report individual differences, the variability can be characterized either as differences between subjects in the regions implicated in performing the task or as differences in the magnitude of the changes observed (i.e., Ref. [5]).

The goal of the current investigation is to determine whether the mechanism that underlies category learning involves network specialization (mechanism 1) or recruitment and use of separable networks (mechanism 2), across the different stages of training. By using materials novel to the subject (random dot patterns), we have investigated the learning process from a naïve state to a more skilled state without concern about previous experience or about mapping the trained skill onto an already existing expert process throughout the training process. We have characterized these dynamic changes within the brain as a function of training using functional magnetic resonance imaging (fMRI) [23,30,31], and a paradigm based on using the learned skill. We have also investigated differences between patterns used during training (trained) and patterns that were created from the same prototypes but not presented during training (untrained). The comparison between the trained and untrained patterns enable us to determine whether subjects acquired the general rules governing the classification of these dot patterns or specific information about the individual, trained patterns. If subjects learned the general rules, behavioral improvements and the neuronal response to the untrained patterns will be similar to the trained patterns. However, if subjects simply learned that a specific pattern belonged to a category (information specific to an individual, trained pattern) without learning the classification rules, there will be no behavioral improvement on the untrained patterns and no change in activation.

## 2. Methods

### 2.1. Participants

Eight volunteers (four male, four female; age range, 24–31 years; mean, 27.6 years) participated after giving written, informed consent consistent with and approved by the Institutional Review Board. All were right handed, healthy, without history of neurological or psychiatric illness and had normal or corrected to normal visual acuity. Participants were instructed on the fMRI paradigms and protocol before the study began.

### 2.2. Stimuli

Subjects were trained to classify random dot patterns into categories over 4 consecutive days. The stimuli were created
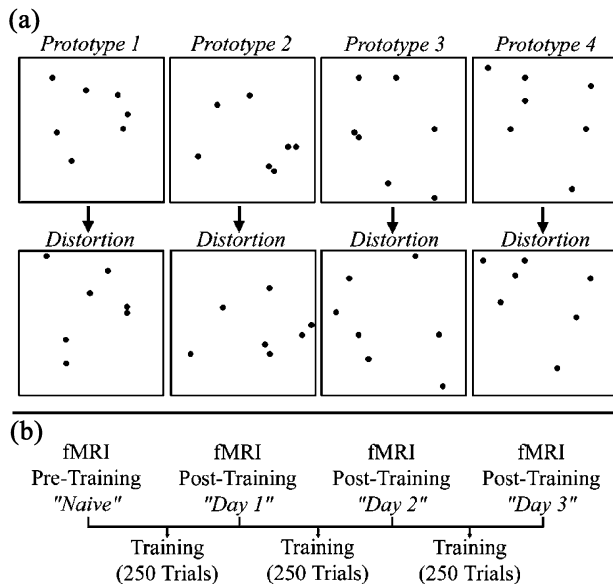
Fig. 1. Experimental protocol: (a) each of the four prototype patterns used in the experiment are presented along with exemplar (distortion) from each prototype. The exemplar patterns were created by displacing randomly each dot within each prototype by four cells in one of four orthogonal directions. Note: although the dots and lines are shown here in black, the actual stimuli used were inverted so that the dots and lines were white on a black background. (b) Protocol for training (below the line) and testing (above the line).

by distorting four "prototype" patterns of dots (see Fig. 1). Each pattern was composed of seven white dots randomly distributed across a grid ($30 \times 30$ cells). The distortions were created by moving each dot across the grid by four cells in any one of four orthogonal directions [36]. This amount of distortion is consistent with previously published work [37] and is based upon behavioral results from our pilot testing ($n=3$). The pilot results suggested relatively small individual differences in rates of learning across subjects and allowed sampling of the entire learning process from naïve to high performance.

The direction of movement for each dot was random, so that each time the distortion was applied, a new pattern of dots or "exemplar" was created. Each of these four prototypes was distorted 30 times to create 120 unique exemplars. Half of the exemplars created by distorting each of the four prototypes were used during the training phases and the remaining exemplars were set aside for use during the fMRI testing sessions. Fig. 1a shows all four prototype dot patterns used to generate the exemplars and one exemplar (or distorted pattern) created from each prototype.

In the present report, the term "prototype" will be used to describe the four original patterns of dots. Subjects were never exposed to these prototypes. The term "exemplar" will be used to describe the patterns of dots that were created by distorting the prototype patterns. Training will be used to refer to the study phase during which feedback was used to aid subjects in learning the rules for classification of the dot

patterns. Testing of the trained skill will refer to the tasks completed while fMRI data were acquired and for which no feedback was provided.

### 2.3. Procedure

Training was completed over 3 consecutive days as illustrated in Fig. 1b. Prior to any training, subjects completed a single fMRI session (Naïve). A training period was presented prior to each successive fMRI session over the remaining 3 days. Each training session lasted approximately 1 h and involved the presentation of 750 trials of training. Training and testing were completed at the same time of day for each subject across all 4 days to remove any effects of circadian cycle.

### 2.4. Training design

A forced, 4-choice, overt classification task with immediate visual feedback was used to train the subjects. At a viewing distance of approximately 30 cm, a single exemplar containing seven white dots on a black background, framed inside a box, was presented centrally. The training was implemented in PsyScope [7] on a Macintosh G4 computer. The entire target subtended approximately 15° of visual angle (head position during the training sessions was not constrained). Subjects were instructed to classify the pattern (exemplar) into one of four possible categories (prototypes) with a key press. Response latency (the time between onset of the exemplar and the key press) and accuracy were recorded. Successful completion of the task required that the subject extract the commonality between dot patterns belonging to each category (prototype). Immediately following selection of a category with a key press, visual feedback was presented centrally, while the exemplar was still present on the computer screen. The feedback not only indicated accuracy of the response but also provided the correct category membership if an incorrect response was made. Study time of this feedback was not limited. Subjects then initiated presentation of the next trial with a separate key press. Subjects were trained to identify category membership on a total of 15 exemplars from each prototype that were presented randomly throughout the training. Over the entire 3 day training program, each exemplar was presented between 65 and 70 times. Chance performance on the classification task was 25%.

### 2.5. fMRI/testing design

During fMRI, two exemplars were projected side by side to the subject in the MRI scanner at a viewing distance of approximately 16 cm. Each exemplar was projected at approximately 15° of visual angle with a 5° separation between the two exemplars. Subjects were instructed to determine if the two exemplars were members of the same category. Responses were recorded with a button press

(same/different). All experimental paradigms were presented as block designs with 30 s of central fixation interspersed with 30 s of the matching task. Trials were presented every 5 s during each 30-s block of the matching task. Each of the paradigms was presented for 6.5 min (the start and end conditions were central fixation).

Two versions of the category matching testing task were presented. The first involved pairs of exemplars that were used during the training (*trained*). The second involved exemplars from the trained prototypes but that had not been used in the training protocol (*untrained*). The order of the tasks was counterbalanced across the days of training and across subjects. Each task (trained/untrained) was composed of 18 trials of exemplar pairs from different categories (non match) and 18 trials of exemplar pairs from the same category (match). Within a single testing session, any given pattern was presented twice. Repetitions within a testing session occurred so that a given pattern was presented once on a matched trial and once on a nonmatched trial (and always on the opposite side of fixation relative to the first trial). Although patterns were repeated within sessions, they were not repeated across sessions, so that any given pattern was only presented during a single fMRI session. Chance performance was 50% for the fMRI task.

## 2.6. fMRI data acquisition

A 3.0-Tesla (T) whole body scanner (Signa VHi, General Electric Medical Systems, Waukesha, WI) using serial gradient echo, echo-planar imaging (epiRT, plane=axial, TR=2999 ms, TE=30.7 ms, flip angle=90°, NEX=1, bandwidth=62 kHz) at high spatial resolution (acquisition matrix=128×128, FOV=20×20 cm$^2$, slice thickness/gap=3/1.5 mm/mm, slices=22) was used for all data acquisition. The paradigms were presented in the scanner and coordinated with behavioral and physiological measurements by a custom designed synchronization control system ([45]; MRIx Technologies, Bannockburn, IL) and visor [44].

Following the completion of each functional imaging session, a single 3D high-resolution anatomical scan was acquired (3D inversion recovery fast spoiled gradient recalled (3D IRfSPGR), plane=axial, TR=9 ms, TE=2.0 ms, flip angle=25°, NEX= 1, bandwidth=15.6 kHz, acquisition matrix=256×256, FOV=22×16.5 cm$^2$, slice thickness/gap=1.5/0 mm/mm, slices=124).

## 2.7. fMRI data analysis

The image processing for each subject and for the group data was accomplished with a combination of complementary statistical software packages FIASCO [11] and AFNI [8]. Voxel-wise statistics for the comparison of the category matching task with the central fixation condition were calculated in FIASCO. To accommodate the shape of the hemodynamic response, data from the first 6 s of each 30-s block were discarded [6]. A false discovery rate of 0.05 was applied to the calculated voxel-wise statistics to reduce the likelihood of Type I errors [18]. Because a false discovery rate was applied to all individual data, no further correction for multiple comparisons was applied. A minimum $t$-threshold ($t>6.8$, $p<.00001$) was applied to the data from each individual subject.

AFNI was used for cluster thresholding, group averaging, normalization of the anatomy, and region of interest (ROI) analysis. Head motion was calculated within AFNI and, although no data were eliminated based upon excessive head motion, the maximum motion observed across subjects and days was less than 1/2 of 1 voxel (<0.78 mm; mean=0.57 mm). A cluster threshold requiring three contiguous voxels was applied to the individual data to further reduce the likelihood of Type I errors [14]. All data were smoothed using a small isotropic Gaussian kernel (full-width half-maximum (FWHM)=1.21). The functional images were aligned over the high-resolution structural data set, resampled to 1 mm$^3$, and transformed into Talairach space using standard anatomical landmarks.

Separate mean images were calculated for the tasks involving trained and untrained exemplars for each subject on each day of testing (Naïve, Day 1, Day 2, Day 3). For the purpose of identifying significant clusters for further analyses, group images were calculated for these eight activation maps for all voxels that exceeded the threshold. Significant voxels in the group map were required to exceed a $t$-threshold of $t>3.5$ ($p<0.01$). Clusters were identified for further analyses if the volume of activation exceeded 200 mm$^3$. Although this method will identify significant clusters across subjects, it may result in areas not being identified but which show small but significant changes within data of individual subjects. This specific limitation in group data is addressed by examining the individual subject data in a second set of analyses described below.

## 2.8. Region-of-interest analysis

All ROI analyses were carried out on data from each individual subject. Two separate sets of analysis were required to test the two alternative hypotheses that describe change in activation as a function of learning. The first set of analyses set out to identify those regions, common across subjects, which were affected by the experimental protocol. The regions were first identified based upon the cluster threshold set in the group map. Only those regions that demonstrated a significant effect of training (defined based upon an analysis of variance (ANOVA) for each region over the days of testing) were included. The second set of analyses set out to determine if the same network or a different network became implicated as a function of training. To identify any new regions that appeared as a function of the protocol, regions of activation, identified based upon the cluster threshold set in the group map, were identified across all days (not limited to those regions that

changed as a function of the protocol). This ROI analysis would allow the identification of any new regions at any stage of training.

ROIs were defined based upon the Talairach and Tournoux stereotaxic atlas [42] and anatomical boundaries consistent with the existing neuroimaging literature as reviewed below. The ROIs that met the above stated criterion of demonstrating changes as a function of experimental protocol included those areas commonly identified as part of the visuospatial networks including left and right frontal eye fields (FEF; along the precentral gyrus to include the adjacent gyrus; Ref. [20]), the supplementary motor areas including the supplementary eye fields (SEF; anterior to the precentral sulcus and posterior to the caudate nucleus; Ref. [26]), left and right superior parietal lobule (SPL; superior and anterior to the intraparietal sulcus and posterior to the postcentral sulcus; Ref. [9]), left and right inferior parietal lobule (IPL; inferior and posterior to the intraparietal sulcus including the supramarginal gyrus; Ref. [15]), primary and secondary visual cortices (along the calcarine fissure to the cuneus and lingual gyrus; Ref. [10])

Table 1
Regions of interest with corresponding Broadmann's areas (BA) and the center (in Talairach coordinates, $x$, $y$, $z$) of each ROI

| Region | BA | | Talairach | | |
|---|---|---|---|---|---|
| | | | $x$ | $y$ | $z$ |
| *Primary motor (M1; Control region)* | | | | | |
| | 4 | | | | |
| | | Right | 44.1 | −23 | 42.4 |
| | | Left | −44.8 | −23 | 42.4 |
| *Supplementary eye fields (SEF)* | | | | | |
| | 6 | | | | |
| | | Right | 5.3 | 8.1 | 52.8 |
| | | Left | −4.8 | 8.2 | 52.9 |
| *Frontal eye fields (FEF)* | | | | | |
| | 8 | | | | |
| | | Right | 41.1 | −6 | 43.9 |
| | | Left | −43.4 | −4.7 | 42.7 |
| *Superior parietal lobule (SPL)* | | | | | |
| | 7 | | | | |
| | | Right | 27.2 | −59 | 52.7 |
| | | Left | −27.2 | −59 | 52.7 |
| *Inferior parietal lobule (IPL)* | | | | | |
| | 40 | | | | |
| | | Right | −48.4 | −41 | 38.7 |
| | | Left | −48.5 | −41 | 38.7 |
| *Visual cortex (V3)* | | | | | |
| | 19 | | | | |
| | | Right | 23.9 | −88 | 16.3 |
| | | Left | −24.9 | −88 | 16.7 |
| *Visual cortex (V1/V2)* | | | | | |
| | 17,18 | | | | |
| | | Right | 10.4 | −88 | −7.7 |
| | | Left | −11.4 | −88 | −7.7 |

and tertiary visual cortex (V3; from the borders of V1/V2 to the middle occipital gyrus; Ref. [10]).

The center of each ROI and its respective Broadmann area are identified in Table 1. The ROIs accounted for on average 73% of all of the total active voxels (Range 67.5–90.5%) across each day of training for both the trained and untrained exemplars.

As with the above $t$-statistics, the percent signal changes between the category matching and fixation conditions were also calculated on a voxel-wise basis for both tasks and for each subject. For those voxels exceeding a $t$-threshold of $p=0.001$, the average percent signal changes were calculated for all subjects across each ROI.

### 2.9. Center of activation analysis

Differences between networks involved over the training and between the trained and untrained tasks were examined by calculating the center of activation (center of mass) within each ROI for each subject, in all three directions ($x$, $y$, $z$), across each day of the protocol.

## 3. Results

The results arising from the training and testing protocol include behavioral performance (accuracy, response times) on the training task, the behavioral results of the testing (accuracy) during fMRI data collection and the activation maps and ROI analysis derived from the fMRI data.

### 3.1. Behavioral training

The group behavioral results from the training are summarized in Fig. 2. As training progressed from the first to the final day, there was an overall increase in accuracy, $F(2,360)=107.39$, $p<0.01$, $\eta^2=0.38$; and an overall decrease in response latency, $F(2,360)=99.14$, $p<0.01$, $\eta^2=0.36$. Post hoc comparisons (Bonferroni, $p<0.05$) confirmed that for both accuracy and latency, subjects showed no additional change in performance between the second and third session of training. A well-established characteristic of skill learning is that behavioral change is best represented by a power function [2,25]. Best-fit lines representing a two-parameter power function, for both latency and accuracy, are superimposed over the experimental data in Fig. 2 (accuracy: $r^2=0.84$, $p<0.001$; latency: $r^2=0.95$, $p<.001$).

### 3.2. Behavioral testing during fMRI

Fig. 3 illustrates that as training progressed, subjects became more accurate in determining whether the two exemplars were members of the same category, $F(4,20)=29.09$, $p<0.01$, $\eta^2=.85$. Although there is a trend for subjects to be more accurate in matching the trained exemplar pairs as compared to the untrained pairs, this
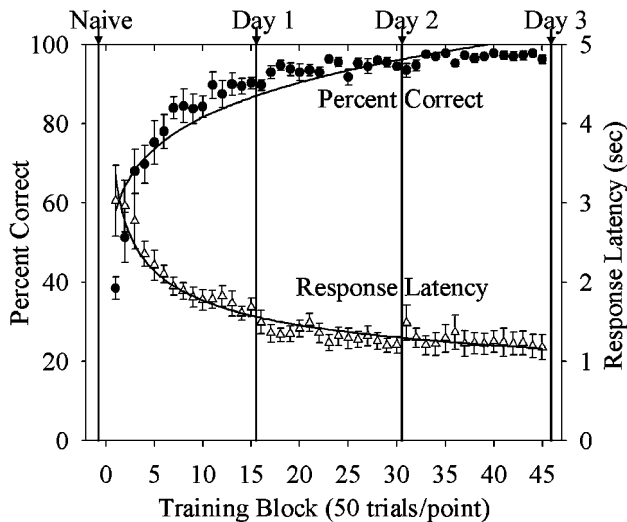
Fig. 2. Response accuracy (●) and latency (△) over the experimental protocol. Each data point represents 50 trials of the classification task. Power functions are shown with the black line drawn through the data points for accuracy and latency. Vertical lines represent position of fMRI sessions relative to training.

observation was not significant, $F(1,5)<1$, ns. Three observations are depicted in Fig. 3. First, performance in the naïve state, prior to any training, was above chance performance, $t(7)=17.24$, $p<0.01$, $\eta^2=0.63$, indicating that although subjects had not received any training, the task could be completed on the basis of visual similarity on a small percentage of the trials. Second, although there was no additional increase in accuracy for the behavioral training between Days 2 and 3, there were still improvements in performance on the matching task, $t(7)=3.464$, $p<0.01$. Finally, although performance was more accurate on the trained than the untrained patterns, this difference was only significant in the last testing session.

## 3.3. Cortical areas involved in category learning and volumetric change

The patterns of activation for the category matching task (for the trained exemplars) as compared to central fixation across the imaging sessions are presented in Fig. 4a (note that a group activation map is presented although all ROI analyses were completed on individual subjects). The largest clusters of activation associated with the category matching task were identified to include the supplementary eye fields, left and right frontal eye fields, left and right superior and inferior parietal lobules and left and right primary, secondary, and tertiary visual cortex. As presented in Fig. 4b, significant overall effects as a function of day were observed for the supplementary eye fields (SEF; trained exemplars only, $F(3,21)=3.3$, $p<0.05$), frontal eye fields (FEF; both trained, $F(3,21)=4.5$, $p<0.05$, and untrained, $F(3,21)=3.2$, $p<0.05$, exemplars), both superior and inferior parietal lobules (SPL for both trained, $F(3,21)=12.7$, $p<0.001$, and untrained, $F(3,21)=5.8$,

$p<0.005$, exemplars, IPL for both trained, $F(3,21)=7.4$, $p<0.001$, and untrained, $F(3,21)=23.2$, $p<0.001$, exemplars), and visual cortex (V3; both trained, $F(3,21)=9.3$, $p<0.001$, and untrained exemplars, $F(3,21)=3.5$, $p<0.05$; V1/V2; trained exemplars only, $F(3,21)=3.3$, $p<0.05$). No significant effects of training over time were observed in the primary motor area which served as a control. These data have been collapsed across hemispheres where no differences between left and right hemispheres in volume of activation were observed.

Three trends can be observed in the data presented in Fig. 4b. First, there was an increase in activation from the Naïve state (pretraining) to the testing session immediately following the first training session (Day 1). Second, there was a decrease in volume of activation as training progressed for the trained exemplars in all areas (excluding the motor control) as training progressed to either Day 2 or Day 3 (demonstrated by positive $t$-values for the comparisons between Day 1 and either Day 2 or Day 3).

Third, there were differences in the percentage of active voxels within the ROIs as a function of the task materials (trained exemplars or untrained exemplars) but only on the last day of testing. As can be seen in the comparison between trained and untrained exemplars in Fig. 4b, there were significantly fewer active voxels in the trained as compared to the untrained state on the last day of testing, Day 3 (mean difference between trained and untrained, 1.05%, $t(7)=2.65$, $p=0.03$). This pattern was maintained for each of the individual ROIs (V1, V2, V3: $t(7)=2.722$, $p=.030$, FEF: $t(7)=3.329$, $p=0.013$, SEF, $t(7)=2.606$, $p=0.035$), with the exception of IPL, SPL, which although there was a trend, did not reach significance.
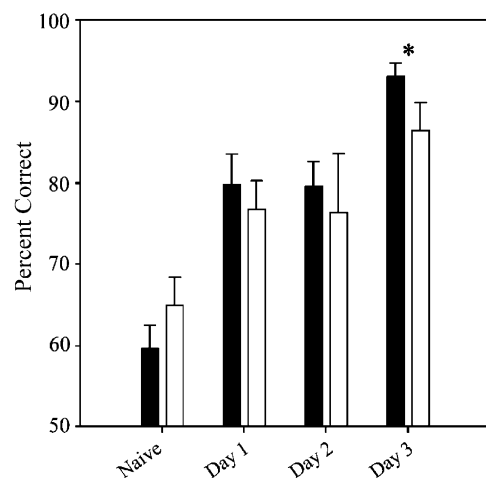


Fig. 3. Response accuracy for performing the category matching task while fMRI data were collected before training (Naïve) and following each training session (Day 1, 2, 3). Solid black bars represent exemplars used in the training program. Solid white bars represent untrained exemplars from trained prototypes. Chance performance is 50%. Error bars represent one standard error. (*) Significant comparisons, $p<.05$, between the trained and untrained.
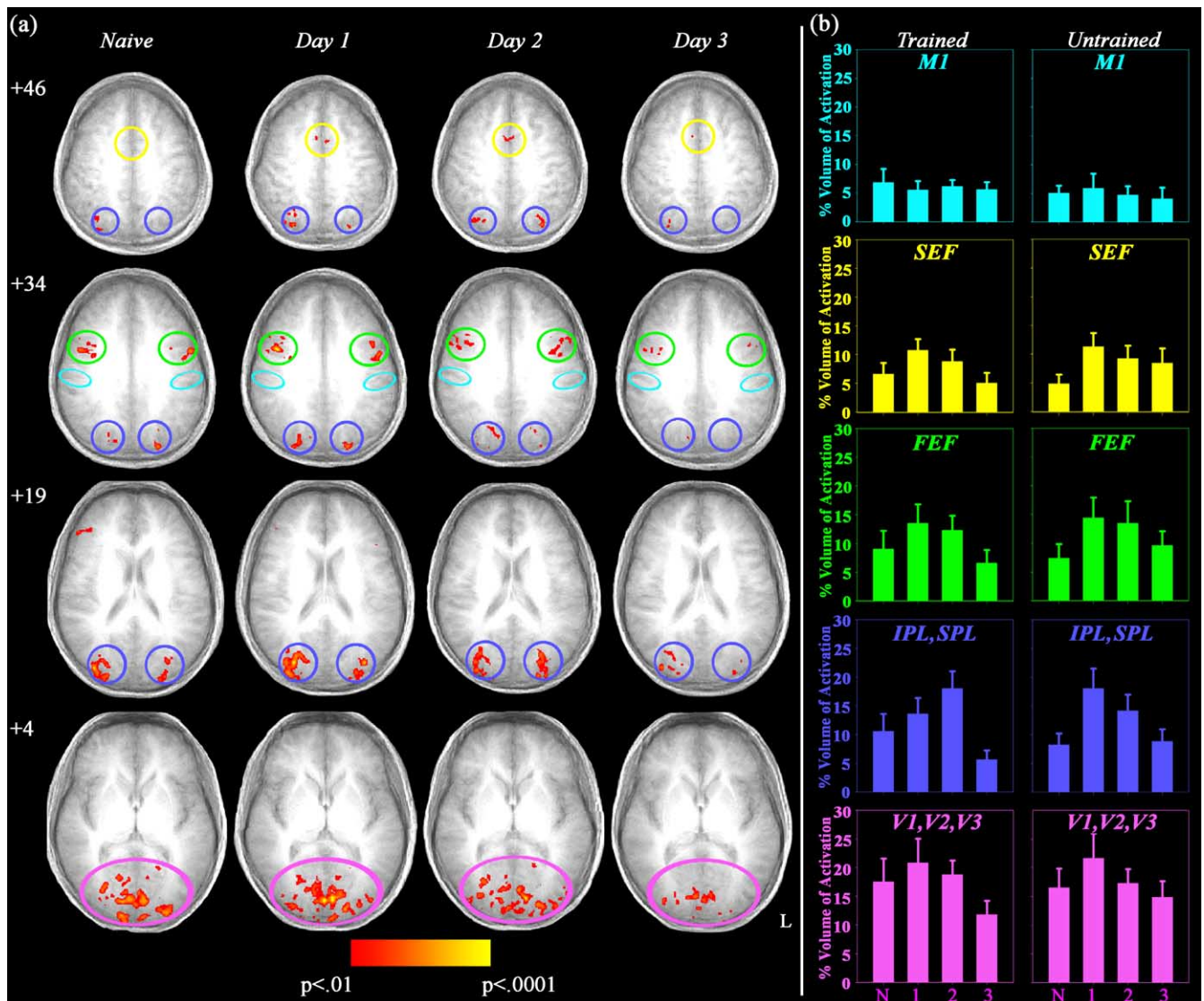
Fig. 4. (a) Representative axial slices of the whole brain activation maps associated with the Category Matching task for trained exemplars as compared to central fixation for the Naïve state prior to training and then post training on Days 1, 2 and 3. A pictorial representation of the areas affected by training are indicated by the yellow (supplementary eye fields), green (frontal eye field), dark blue (superior and inferior parietal lobules), red (primary, secondary, and tertiary visual cortex), and light blue (primary motor areas) circles. Note that the circles are only representations to connect the activation to the graphs not the actual regions of interest. Criteria for identifying specific regions of interest are indicated in the Methods. Images are presented according to radiological convention (right side of image presented on the left). (b) Specific volumetric changes (presented as a percent of each ROI) are presented for those five areas of interest for trained and untrained patterns.

## 3.4. Identification of new regions involved in category learning

A final set of analyses was carried out to determine whether there were any new regions implicated in completing the category matching task. As previously described (see Section 2.8), this set of analyses involved the identification of significant clusters for each day without the additional requirement that those clusters be active across all days. One additional region was identified with this set of analyses. Three of the eight subjects showed activation in a region identified as prefrontal cortex ($x=39\pm2$, $y=36\pm1.5$, $z=14\pm1$) in the Naïve state. Only 1

of the 3 subjects showed activation in this region on Day 1 and never thereafter. No new areas were detected across Days 1, 2 and 3.

## 3.5. Locus of activation changes as a function of learning

Although the only statistically significant difference in the volume of activation between the two sets of task materials (trained exemplars versus untrained exemplars) occurred on Day 3, the question as to whether there are subtle differences in the locus of activation within each ROI remained. As previously discussed (see Section 2.9), the center of activation within each ROI was identified for each

subject across training. To directly compare the center of activation for each task, the differences in all three directions (*x*, *y*, *z*) for each subject were submitted to a repeated measures ANOVA. There were no differences observed as a function of training across the ROIs (*p*>0.1).

### 3.6. Cortical areas involved in category learning and signal intensity change

It is important to note that although there were significant decreases in the overall volume of activation within the ROIs, there were no significant changes in either the distribution ($\chi^2$>0.06) or the magnitude of the BOLD signal change as a function of training (F-tests with *p*>0.08). The signal intensity remained stable over training at 5.3%±2.1% across all subjects for each ROI.

## 4. Discussion

The current investigation has established a baseline pattern of brain activity for learning to categorize random dot patterns. That baseline pattern includes regions involved in visuospatial processing (visual and association cortex), spatial attention (superior and inferior parietal lobules) and the control of eye movements (frontal and supplementary eye fields) [21,24,39] (Table 1, Fig. 4a). Following the initial behavioral training session, increases in the volumes of activation were observed across this network. These initial increases in activation parallel the large improvements in accuracy on the fMRI testing task from Naïve to Day 1 (Fig. 3).

Further training on the classification of the random dot patterns (Days 2 and 3) continued to produce additional improvements in accuracy on both the training and testing tasks (Figs. 2 and 3). These improvements in behavioral performance paralleled decreases in the volumes of activation across this network (visuospatial processing, spatial attention and eye movement control) during the testing program. The largest decreases in activation occurred only after behavioral training performance had stabilized at a high level of accuracy from Day 2 to Day 3 (Fig. 2), although continued improvements in accuracy were still observed on the testing task (Fig. 3).

Our goal was to distinguish between two possible mechanisms (or some combination of both) that have been proposed in the literature to explain the neural changes with learning. The first mechanism is a strengthening of the task-dependent neuronal relationships. This strengthening, which is thought to occur as a function of practice on a task, would produce increased efficiency of the same cognitive operations in the same regions of the brain. The second mechanism is the recruitment of a new cognitive operation on a separate neural network or in other brain regions.

The results suggest that the successful learning of the classification rules of random dot patterns involves a single network throughout training. This conclusion is based upon the evidence for a specialization of activation (decrease in volume of activation, Fig. 4) without a change in the locus of activation (Section 3.5) within each ROI as training progressed. Furthermore, although the volume of activation was reduced as training progressed, the intensity of the signal change was not altered as indicated by the stable signal changes over time (Section 3.6). We interpret this combination of findings to reflect increased efficiency in processing. This increased efficiency is observed as a decrease in the volume of activation without increasing the magnitude of the activation while still completing the task.

Although the current analyses eliminate the possibility that an entirely separate network is recruited, the possibility remains that there are relatively subtle differences in the distribution of activity across the broad network. There is evidence to suggest that regions can begin to serve a different role while remaining active during learning. One such example is the role of the thalamus for gating activity within the striatum in response to a motor task. Rauch et al. [38] demonstrated that early in learning, activation within the thalamus was reduced whereas activity within the striatum was increased. However, later in the learning process, the activation within the thalamus was actually enhanced.

The question remains as to whether the effects demonstrated in the current study are reflective of learning or whether they can be explained by other, related, cognitive processes such as cognitive duty cycle. The cognitive duty cycle is observed as a reduction in response latency as exposure to a task increases. The behavioral and possibly neuronal changes associated with learning are coupled with these practice effects.

The data presented show that an overall reduction in the volume of activation over time is associated with learning mediated by experience with classifying the stimulus materials. The present results also demonstrate this expected reduction in response latency as subjects learn the task. Other investigators have also reported a similar finding, with reduction in activation over time associated with reductions in the response time required to carry out a task. Given the consistency between our results and the literature across very different tasks, it is necessary to address concerns of duty cycle directly.

Neuronal activation is known to decrease when the time required to complete the cognitive operation is reduced [33,40]. If, during a block-designed experiment, the time to complete each trial within the block decreases, a reduction in the magnitude of BOLD contrast throughout the ROIs would be predicted. However, we observed that the BOLD contrast remained stable across the protocol and report only a reduction in the volume of activation. This observation is more consistent with effects of learning than those of duty cycle. The second observation is that the largest reduction in latency is observed between the Naïve and Day 1 sessions.

However, this decrease in latency was observed as an increase in the volume of activation rather than a decrease as predicted by a duty cycle explanation.

Our findings suggest that category learning of novel materials implicates a network consisting of visual cortices, both superior and inferior parietal lobules, and those areas involved in the planning and execution of eye movements. The data support the notion that there is a specialization of this network as subjects become more accurate across the learning protocol. The reduction in activation occurs across the same network implicated early in training and no new regions are recruited.

## Acknowledgements

## References

[1] L.A. Alfonso-Reese, F.G. Ashby, D.H. Brainard, What makes a categorization task difficult? Percept. Psychophys. 64 (2002) 570–583.

[2] J.R. Anderson, Acquisition of a cognitive skill, Psychol. Rev. 89 (1982) 369–406.

[3] F.G. Ashby, S.W. Ell, The neurobiology of human category learning, Trends Cogn. Sci. 5 (2001) 204–210.

[4] J.N. Bailenson, M.S. Shum, S. Atran, D.L. Medin, J.D. Coley, A bird's eye view: biological categorization and reasoning within and across cultures, Cognition 84 (2002) 1–53.

[5] C. Buchel, J.T. Coull, K.J. Friston, The predictive value of changes in effective connectivity for human learning, Science 283 (1999) 1538–1541.

[6] M.S. Cohen, Parametric analysis of fMRI data using linear systems methods, NeuroImage 6 (1997) 93–103.

[7] J.D. Cohen, B. MacWhinney, M. Flatt, J. Provost, PsyScope: a new graphic interactive environment for designing psychology experiments, Behav. Res. Meth. Instrum. C 25 (1993) 257–271.

[8] R.W. Cox, AFNI: software for analysis and visualization of functional magnetic resonance neuroimages, Comput. Biomed. Res. 29 (1996) 162–173.

[9] P. Dassonville, X.-H. Zhu, K. Ugurbil, S.-G. Kim, J. Ashe, Functional activation in motor cortex reflects the direction and degree of handedness, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 14015–14018.

[10] E.A. DeYoe, G.J. Carman, P. Bandettini, S. Glickman, J. Wiser, R. Cox, D. Miller, J. Neitz, Mapping striate and extrastriate visual areas in human cerebral cortex, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 2382–2386.

[11] W.F. Eddy, M. Fitzgerald, C.R. Genovese, A. Mockus, D.C. Noll, Functional image analysis software—Computational olio, in: A. Prat (Ed.), Proceedings in Computational Statistics, Physica-Verlag, Heidelberg, 1996.

[12] S. Edelman, N. Intrator, Models of perceptual learning, in: M. Fahle, T. Poggio (Eds.), Perceptual Learning, MIT Press, Cambridge, Mass, 2002, pp. 337–354.

[13] P. Fletcher, C. Buchel, O. Josephs, K. Friston, R. Dolan, Learning-related neuronal responses in prefrontal cortex studied with functional neuroimaging, Cereb. Cortex 9 (1999) 169–178.

[14] S.D. Forman, J.D. Cohen, M. Fitzgerald, W.F. Eddy, M.A. Mintun, D.C. Noll, Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold, Magn. Reson. Med. 33 (1995) 636–647.

[15] M.E. Frederikse, A. Lu, E. Aylward, P. Barta, G. Pearlson, Sex differences in the inferior parietal lobule, Cereb. Cortex 9 (1999) 896–901.

[16] J.D.E. Gabrieli, Cognitive neuroscience of human memory, Annu. Rev. Psychol. 49 (1998) 87–115.

[17] I. Gauthier, M.J. Tarr, A.W. Anderson, P. Skudlarski, J.C. Gore, Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects, Nat. Neurosci. 2 (1999) 568–573.

[18] C.R. Genovese, N.A. Lazar, T. Nichols, Thresholding of statistical maps in functional neuroimaging using the false discovery rate, NeuroImage 15 (2002) 870–878.

[19] D. Gonzalo, T. Shallice, R. Dolan, Time-dependent changes in learning audiovisual associations: a single-trial fMRI study, Neuro-Image 11 (2000) 243–255.

[20] M.-H. Grosbras, U. Leonards, E. Lobel, J.-B. Poline, D. LeBihan, A. Berthoz, Human cortical networks for new and familiar sequences of saccades, Cereb. Cortex 11 (2001) 936–945.

[21] N. Kanwisher, M.M. Chun, J. McDermott, P.J. Ledden, Functional imaging of human visual recognition, Cogn. Brain Res. 5 (1996) 55–67.

[22] B.J. Knowlton, What can neuropsychology tell us about category learning? Trends Cogn. Sci. 3 (1999) 123–124.

[23] K.K. Kwong, J.W. Belliveau, D.A. Chesler, I.E. Goldberg, R.M. Weisskoff, B.P. Poncelet, D.N. Kennedy, B.E. Hoppel, M.S. Cohen, R. Turner, H.-M. Cheng, T.J. Brady, B.R. Rosen, Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation, Proc. Natl. Acad. Sci. U. S. A. 89 (1992) 5675–5679.

[24] K.S. LaBar, D.R. Gitelman, T.B. Parrish, M. Mesulam, Neuro-anatomic overlap of working memory and spatial attention networks: a functional MRI comparison within subjects, NeuroImage 10 (1999) 695–704.

[25] G.D. Logan, Toward an instance theory of automatization, Psychol. Rev. 95 (1998) 492–527.

[26] B. Luna, K.R. Thulborn, M.H. Strojwas, B.J. McCurtain, R.A. Berman, C.R. Genovese, JA. Sweeney, Dorsal cortical regions subserving visually guided saccades in humans: an fMRI study, Cereb. Cortex 8 (1998) 40–47.

[27] W.T. Maddox, F.G. Ashby, C.J. Bohil, Delayed feedback effects on rule-based and information–integration category learning, J. Exp. Psychol. Learn. 29 (2003) 650–662.

[28] A.B. Markman, B.H. Ross, Category use and category learning, Psychol. Bull. 129 (2003) 592–613.

[29] A.R. McIntosh, R.E. Cabeza, N.J. Lobaugh, Analysis of neural interactions explains the activation of occipital cortex by an auditory stimulus, J. Neurophysiol. 80 (1998) 2790–2796.

[30] S. Ogawa, T.M. Lee, A.S. Nayak, P. Glynn, Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high fields, Magn. Reson. Med. 14 (1990) 68–78.

[31] S. Ogawa, D.W. Tank, R.S. Menon, J.M. Ellermann, S.-G. Kim, H. Merkle, K. Ugurbil, Intrinsic signal changes accompanying sensory stimulation: functional brain mapping using MRI. Proc. Natl. Acad. Sci. U. S. A., 89, 5951–5955.

[32] S.E. Petersen, H. Van Mier, J. Fiez, M. Raichle, The effects of practice on the functional anatomy of task performance, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 853–860.

[33] R.A. Poldrack, Imaging brain plasticity: conceptual and methodo-logical issues-A theoretical review, NeuroImage 12 (2000) 1–13.

[34] R.A. Poldrack, J.D.E. Gabrieli, Characterizing the neural mechanisms of skill learning and repetition priming: evidence from mirror reading, Brain 124 (2001) 67–82.

[35] R.A. Poldrack, J.E. Desmond, G.H. Glover, J.D.E. Gabrieli, The neural basis of visual skill learning: an fMRI study of mirror reading, Cereb. Cortex 8 (1998) 1047–1056.

[36] M.I. Posner, S.W. Keele, On the genesis of abstract ideas, J. Exp. Psychol. 77 (1968) 353–363.

[37] M.I. Posner, R. Goldsmith, K.E. Welton, Perceived distance and the classification of distorted patterns, J. Exp. Psychol. 73 (1967) 28–38.

[38] S.L. Rauch, P.J. Whalen, T. Curran, S. McInerney, S. Heckers, C.R. Savage, Thalamic deactivation during early implicit sequence learning: a functional MRI study, NeuroReport 9 (1998) 865–870.

[39] C. Rosano, C.M. Krisky, J.S. Welling, W.F. Eddy, B. Luna, K.R. Thulborn, J.A. Sweeney, Pursuit and saccadic eye movement subregions in human frontal eye field: a high-resolution fMRI investigation, Cereb. Cortex 12 (2002) 107–115.

[40] N. Sadato, V. Ibanez, G. Campbell, M.P. Deiber, D. Le Bihan, M. Hallett, Frequency-dependant changes of regional cerebral blood flow during finger movements: functional MRI compared to PET, J. Cereb. Blood Flow Metab. 17 (1997) 670–679.

[41] E.E. Smith, A.L. Patalano, J. Jonidas, Alternative strategies of categorization, Cognition 65 (1998) 167–196.

[42] J. Talairach, P. Tournoux, Co-Planer Stereotaxic Atlas of the Human Brain, Thieme Medical Publishers, New York, 1988.

[43] C.M. Thiel, D.R. Shanks, R.N. Henson, R.J. Dolan, Neuronal correlates of familiarity-driven decisions in artificial grammar learning, NeuroReport 14 (2003) 131–136.

[44] K.R. Thulborn, Visual feedback to stabilize head position for fMRI, Magn. Reson. Med. 41 (1999) 1039–1043.

[45] K.R. Thulborn, S.Y. Chang, G. Shen, J.T. Voyvodic, High-resolution echo-planar fMRI of human visual cortex at 3.0 Tesla, NMR Biomed. 10 (1997) 183–190.

[46] I. Toni, M. Krams, R. Turner, R.E. Passingham, The time course of changes during motor sequence learning: a whole-brain fMRI study, NeuroImage 8 (1998) 50–61.

[47] I. Toni, J. Rowe, K.E. Stephan, R.E. Passingham, Changes of cortico-striatal effective connectivity during visuomotor learning, Cereb. Cortex 10 (2002) 1040–1047.