Published in final edited form as:

Neuroimage. 2011 June 1; 56(3): 1791–1802. doi:10.1016/j.neuroimage.2011.02.011.

Cortical and Striatal Contributions to Automaticity in Information-Integration Categorization

Jennifer G. Waldschmidt and F. Gregory Ashby

University of California, Santa Barbara

1. Introduction

Evidence is rapidly accumulating that human categorization is mediated by a number of functionally distinct category-learning systems. The evidence suggests that these different systems are each best suited for learning different types of category structures and are mediated by different neural circuits (e.g., Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Erickson & Kruschke, 1998; Love, Medin, & Gureckis, 2004; Reber, Gitelman, Parrish, & Mesulam, 2003). Agreement is good that at least one of these systems uses some form of explicit reasoning (e.g., hypothesis testing or rule learning) and recruits working memory and perhaps other declarative memory systems as well. There is also broad consensus that another system classifies more on the basis of overall similarity. The evidence suggests that this system recruits procedural memory and depends on synaptic plasticity within the striatum (Ashby et al., 1998; Ashby, Ell, & Waldron, 2003; Ashby & Ennis, 2006; Maddox, Bohil, & Ing, 2004b).

The task that has yielded the best evidence for explicit reasoning processes in category learning is the rule-based task. In rule-based categorization tasks the optimal strategy is easy to verbalize and the categories can be learned via a logical reasoning process (e.g., the Wisconsin Card Sorting test; Heaton, 1981). In contrast, the task that has yielded the best evidence for a procedural-learning categorization system is the information-integration (II) categorization task. In II tasks, accuracy is maximized only if information from two or more stimulus components is integrated at some pre-decisional stage (Ashby & Gott, 1988). Typically, the optimal strategy in II tasks is difficult or impossible to describe verbally (Ashby et al., 1998). An example is shown in Figure 1 in which two categories are constructed from circular sine-wave gratings that vary across trials only in the width and orientation of the dark and light bars. The subject's task is to learn to assign each stimulus to its correct category. On each trial, one stimulus is presented, the subject makes his or her categorization decision, and then feedback is provided about the accuracy of the response. Note that there is no simple verbal description of the diagonal category boundary. Rule-based strategies can be (and frequently are) applied in II tasks, but they produce sub-optimal accuracy.

Much is known about behavioral profiles during II learning. For example, improvement in II tasks is incremental and can continue throughout weeks or months of practice. In addition, even highly accurate subjects are poor at describing their classification strategy (Ashby &

Contact: F. Gregory Ashby, Department of Psychology, University of California, Santa Barbara, CA 93106, Phone: 805-893-2858, Fax: 805-893-4303, ashby@psych.ucsb.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Maddox, 2005). II learning differs qualitatively from rule-based learning in many important ways. First, II learning is of response goals whereas rule-based learning is of category labels (Ashby et al., 2003; Maddox et al., 2004b). Second, II learning is extremely sensitive to the timing and nature of the feedback, whereas rule-based learning is not (Ashby, Maddox, & Bohil, 2002; Maddox, Ashby, & Bohil, 2003). For example, II learning is optimized if the feedback is given immediately after the response, whereas long feedback delays do not interfere with rule-based learning. Third, feedback processing is automatic during II learning but requires attention and effort during rule-based learning (Maddox, Ashby, Ing, & Pickering, 2004a). Fourth, a dual task requiring working memory interferes with rule-based category learning, but not with II learning (Waldron & Ashby, 2001; Zeithamova & Maddox, 2006, 2007).

There is also good evidence that II learning depends critically on the striatum (for reviews, see Ashby & Ennis, 2006, or Seger, 2008). Even so, little is known about the relative contributions of the various striatal regions to this type of learning. The early category-learning literature focused on the caudate nucleus (Ashby et al., 1998), and the few fMRI studies of II categorization have all reported significant caudate activation (Cincotta & Seger, 2007; Nomura et al., 2007; Seger & Cincotta, 2002). There are at least two reasons to question this conclusion however. First, previous fMRI studies have not separated the effects of categorization from feedback processing, so it is difficult to rule out the hypothesis that the caudate activation observed in these studies might be due to feedback processing, rather than to categorization. Second, several recent studies have uncovered a significant motor learning component in II tasks (Ashby et al., 2003; Maddox et al., 2004b). Since motor regions receive input from the posterior putamen rather than the caudate (Matelli & Luppino, 1996), this would suggest that the putamen might play a more prominent role than has previously been thought.

Another question, which has never been addressed empirically, is how the neural processes that mediate II category learning are related to those that mediate automatic II categorization. Before addressing this question however, it is important to examine how neural activation changes in other tasks as automaticity develops. In general, results have been different depending on whether the task recruits declarative or procedural memory systems (e.g., Kelly & Garavan, 2005). In tasks that depend on declarative memory (e.g., working memory), reductions in neural activation with practice are common, especially in regions associated with executive attention that are thought to be critical for early, effortful processing in explicit tasks (Chien & Schneider, 2005). Procedural-learning tasks (e.g., sequence learning), on the other hand, have often reported increases in neural activation with extended practice, primarily in motor regions (Hazeltine, Grafton, & Ivry, 1997; Honda et al., 1998; Karni et al., 1995, 1998).

In the case of II categorization, there are at least two competing theoretical hypotheses about how automaticity might develop (for a review, see Ashby, Turner, & Horvitz, 2010). One prominent hypothesis is that the development of automaticity is mediated by a gradual transfer of control from the associative striatum to the sensorimotor striatum (Belin, Jonkman, Dickinson, Robbins, & Everitt, 2009; Costa, 2007; Yin & Knowlton, 2006). A contrasting view is that automaticity is associated with a transfer of control from the basal ganglia to cortico-cortical projections from the relevant sensory areas directly to the premotor and/or motor areas that initiate the behavior (Ashby, Ennis, & Spiering, 2007). According to this view, the primary role of the basal ganglia is to train these cortico-cortical projections.

To address these questions, we report the results of an experiment in which human participants each practiced the II categorization task shown in Figure 1 for more than 11,000

trials distributed over 20 separate experimental sessions. Four of these sessions were performed inside an MRI scanner. Category learning is thought to follow the power law of practice (Newell & Rosenbloom, 1981; Nosofsky & Palmeri, 1997), and for this reason the four scanner sessions were 2, 4, 10, and 20 (which are nearly equally spaced on a log scale).

A theoretical challenge when studying automaticity is to identify a point in training at which the behavior has become automatic. This is a difficult problem because the development of automaticity occurs gradually over long periods of time. For example, Crossman (1959) reported that factory workers were still improving their cigar-rolling performance after a million trials of practice. Such gradual and extended change makes it problematic to identify a single time point at which the behavior has become automatic.

Another problem is that many different criteria for identifying automaticity have been proposed and none of these are widely accepted as definitive. Schneider and Shiffrin (1977; Shiffrin & Schneider, 1977) proposed some of the most influential criteria for automaticity. Two of their criteria, which are especially relevant in the present study, are that a behavior should be considered automatic if 1) it can be executed successfully while the participant is simultaneously engaged in some other secondary task, and 2) it becomes behaviorally inflexible. For example, by this latter criterion a behavior should be considered automatic if switching the location of the response keys interferes with its expression. These criteria are especially problematic for II categorization. As mentioned earlier, several studies have reported that during the first session of practice, a dual task requiring working memory interferes with rule-based category learning but not with II learning (Waldron, & Ashby, 2001; Zeithamova & Maddox, 2006, 2007), whereas switching the locations of the response keys interferes with initial II but not rule-based performance (Ashby et al., 2003; Maddox et al., 2004b; Maddox, Glass, O'Brien, Filoteo, & Ashby, 2010; Spiering & Ashby, 2008). Therefore, by the Schneider and Shiffrin criteria, II categorization is automatic after the first training session but rule-based categorization is not. Such a conclusion is incompatible with intuitive notions of automaticity, because accuracy in II tasks requires several thousand trials to asymptote (Hélie, Waldschmidt, & Ashby, 2010).

For these reasons, we used a converging operations approach to identifying automaticity. In particular, our strategy was to use a wide variety of criteria, including behavioral inflexibility, resistance to dual-task interference, examining accuracy and response time profiles, and looking for qualitative changes in the imaging results. As we will see, these various criteria suggest that the performance of our participants was automatic during the final scanning session (session 20), but not during the earlier sessions.

2. Materials and Methods

2.1 Participants

Eleven right-handed participants (7 male, 4 female; age range = 19 – 26 years-old) from the University of California, Santa Barbara community were recruited to participate in 23 sessions of training. Participants were healthy and reported no previous brain injury or neurological disorder. All participants received course credit for participation or a monetary compensation of \$10/hour for each behavioral session and \$20/hour for each fMRI session. No participants reported any previous fMRI experience. One participant was excluded from the experiment due to an inability to learn the correct category structures by Session 5.

2.2 Stimuli and Apparatus¹

The stimuli were circular sine-wave gratings of constant contrast and size that varied across trials in spatial frequency and orientation. Each stimulus was defined by a set of points (x_1, x_2) sampled from a 100×100 stimulus space and converted to a disk using the following equations: *bar width* = $(x_1/30+0.25)*5$ cycles/disk (cpd), and *orientation* = $9x_2/10+20$ degrees of counterclockwise rotation from horizontal. This yielded stimuli that varied from 1.25 to 17.9 cpd in bar width and from 20° to 110° in orientation. The stimuli were generated with MATLAB using Brainard's (1997) Psychophysics Toolbox and subtended a visual angle of approximately 5°. Example stimuli, category structures, and the optimal category boundary are shown in Figure 1.

Each category was defined by a bivariate normal distribution in stimulus space and the category exemplars were generated by drawing 240 random samples from each of these distributions (Ashby & Gott, 1988) for a total of 480 stimuli. The mean vectors of these distributions were:

$$\underline{\mu}_A = \begin{bmatrix} 40 \\ 60 \end{bmatrix}; \underline{\mu}_B = \begin{bmatrix} 60 \\ 40 \end{bmatrix}.$$

The variance-covariance matrices for the distributions were equal with a variance of 185 and a covariance of 170. The 240 A and B random samples were linearly transformed so that the sample statistics (means, variances, and covariances) exactly matched these population values. Perfect accuracy was theoretically possible.

2.3 Apparatus and Behavioral Methods

The experiment included 23 separate sessions (each on a different day). Nineteen of the sessions were conducted in the laboratory and four were conducted in an MRI scanner. The laboratory sessions consisted of 12 blocks of 50 stimuli, for a total of 600 stimuli per session. The scanner sessions were composed of 6 blocks of 80 stimuli, for a total of 480 stimuli per session. The scanning sessions were sessions 2 (after 600 trials of practice), 4 (after 1,680 trials of practice), 10 (after 5,160 trials of practice), and 20 (after 11,040 trials of practice). One participant was scanned on the nineteenth session of the study (after 10,440 trials of practice).

Participants were told that they were taking part in a categorization experiment and they were to assign each stimulus to either an A or B category. Stimulus presentation, feedback, response recording, and response time (RT) measurement were controlled and acquired using MATLAB run on a Macintosh computer. In the laboratory, subjects were tested while sitting at a table and looking at a monitor where a stimulus subtended approximately 5 degrees of visual angle. "A" responses were given by the left index finger while "B" responses were given by the right index finger, on the "d" and "k" keys of the keyboard, respectively. Feedback was delivered through headphones where a high-pitched tone indicated a correct response, a low-pitched tone indicated an incorrect response, and a sawtooth tone indicated a wrong button press or a response that occurred after 5 seconds. If the sawtooth tone occurred, it was followed by the words "wrong key" or "too slow", depending on the reason for the tone.

¹This subsection details the materials and procedures used in the scanning sessions. Details concerning the training sessions outside the scanner can be found in Hélie, Waldschmidt, and Ashby (2010).

During the scanning sessions, participants selected category A or B using the Lumina LP-400 Response Pad System (model LU400-Pair). The button box in the left hand indicated an “A” category response and the button box in the right hand indicated a “B” response. Participants were instructed to use their index finger to make responses. Correct responses were indicated by a green check mark displayed for 2,000 ms. An incorrect response was indicated by a red “X” displayed for 2,000 ms. If a response was too slow (i.e., more than 2,000 ms), a black dot was displayed for 2,000 ms.

During scanning sessions, a fixation point (crosshair) appeared for 1,000 ms before the stimulus on an average of 50% of the trials. The irregular presentation of the crosshair effectively decorrelates the regressors representing the crosshair and stimulus events (corresponding to a partial trials design; Serences, 2004). The crosshair carried no information about category membership, so it should not affect response accuracy. In the laboratory sessions, a crosshair never appeared. More details on the timing of a trial (including the jittering parameters) are shown in Figure 2.

Participants used the same fingers to respond A and B in the scanner and in the laboratory, but note that there were many differences between laboratory and scanning sessions (e.g., participant body position, timing on each trial, response devices, feedback). Parker, Ngu, and Cassaday (2001) reported that performance on a procedural-learning task can be affected if even subtle changes in context occur (i.e., background odor), so the procedural differences between laboratory and scanning may have impaired performance on scanning sessions. In fact, mean RT was slower on scanning days, although accuracy appeared unaffected by these differences.

2.4 Neuroimaging

A rapid event-related fMRI procedure was used. The scanning sessions were conducted at the University of California, Santa Barbara Brain Imaging Center using a 3T Siemens TIM Trio MRI scanner with an 8-channel phased array head coil. Cushions were placed around the head to minimize head motion. Functional runs used a T2*-weighted single shot gradient echo, echo-planar sequence sensitive to BOLD contrast (TR: 2,000 ms; TE: 30 ms; FA: 90°; FOV: 192 mm; voxel: 3×3×3 mm) with generalized auto-calibrating partially parallel acquisitions (GRAPPA; Tintera et. al., 2004). Each volume consisted of 33 slices acquired parallel to the AC–PC plane (interleaved acquisition; 3 mm thick with .5 mm gap; 3 mm×3 mm in-plane resolution; 64×64 matrix). Stimuli were viewed through a mirror mounted on the head coil and a back projection screen. A localizer, a GRE field mapping (3 mm thick; FOV: 192mm; voxel: 3×3×3 mm; FA = 60°), and a T1-flash (TR=15 ms; TE=4.2 ms; FA=20°; 192 sagittal slices 3-D acquisition; 0.89 mm thick; FOV: 220 mm; voxel: 0.9×0.9×0.9 mm; 256×256 matrix) were obtained before the EPI scans, and an additional GRE field-mapping scan was acquired at the end of each scanning session. Slice orientation was equal for all GREs and EPIs. Each scanning session lasted about 90 minutes.

2.5 Neuroimaging Analysis

Preprocessing and data analysis were conducted using FEAT (FMRI Expert Analysis Tool) version 5.98, a part of FSL (www.fmrib.ox.ac.uk/fsl). Preprocessing was done separately on each EPI scan to reduce sources of noise and artifact. Preprocessing included motion correction using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002), slice timing correction (via Fourier time-series phase-shifting), BET brain extraction (Smith, 2002), spatial smoothing with a FWHM of 5 mm, grand-mean intensity normalization, and a high pass filter with a cutoff of 100 seconds. The structural scan was registered to the MNI152-T1-2mm standard brain using FLIRT (Jenkinson et al., 2002; Jenkinson & Smith, 2001) and further refined using FNIRT (nonlinear registration; Anderson, Jenkinson, & Smith, 2007).

Each functional scan was registered to the standard brain by first registering to the structural scan and then applying the same parameters used to register the structural scan to the standard brain. Scanning data with excessive head motion correction (i.e., greater than 3 mm) were excluded from the remaining analyses (~ 5% of the data).

First, low-level analyses were performed separately on each EPI scan. Three explanatory variables (EVs) were defined: stimulus, feedback, and baseline (defined as the TRs during which the fixation point crosshairs were shown). Only stimuli that were classified correctly were included in the analysis (i.e., error rates were low, especially in later sessions). Boxcar functions were defined for each EV. For baseline and stimulus the boxcar heights were 1. The baseline boxcar duration was 1,000 ms, whereas the stimulus boxcar duration was set to equal the participant's RT on that trial. Participants had extensive prior practice in all scanning sessions. As a result, they were only uncertain about the accuracy of responses to stimuli that were near the category boundary (Paul, Boomer, Smith, & Ashby, in press). For this reason, it seemed likely that participants would attend more closely to feedback on trials when uncertainty was high. RT in II tasks has a strong negative correlation with distance to the decision bound (Ashby, Boynton, & Lee, 1994). As a result, we set the height of the feedback boxcar function to the normalized RT on that trial (equal to the RT divided by the mean RT in that session). Thus, the feedback boxcar was highest on trials when uncertainty was greatest. The duration of the feedback boxcar was set to 2,000 ms (i.e., the TR). The boxcar events were convolved with a gamma function with a standard deviation of three seconds and a mean lag of six seconds. A temporal derivative and temporal filtering were added to the design matrix. Three contrasts were formed. Stimulus and feedback contrasts were created by subtracting the baseline EV from each of the other EVs. A feedback deactivation contrast was created by subtracting feedback from baseline.

Second, the results of the low-level analyses were input into mid-level analyses to aggregate the block data into session data using a fixed effects model in FLAME (FMRIB's Local Analysis of Mixed Effects; Woolrich, 2008). The mid-level analyses yielded a separate brain map for each participant in each session. Active clusters were identified at all levels of analysis by setting an initial z threshold of 2.3 and then using a threshold on cluster size derived from Gaussian random field theory that yielded an experiment-wise false positive rate of $\alpha = .05$ (Friston, Worsley, Frackowiak, Mazziotta, & Evans, 1994). Third, the results of mid-level analyses were input into a high-level analysis to generate group maps for each session using a mixed-effects FLAME 1+2 model. Active clusters were identified using the same statistical criteria as in the mid-level analysis.

In addition to whole brain analysis, anatomical regions of interest (ROIs) were examined. The ROIs were obtained from the only existing detailed neurobiological theories of category learning (COVIS; Ashby et al., 1998) and automatic categorization in II tasks (SPEED; Ashby et al., 2007). Briefly, COVIS assumes separate rule-based and procedural-learning category-learning systems that compete for access to response production (Ashby & Valentin, 2005; Ashby et al., 1998). The rule-based system selects and tests simple verbalizable hypotheses about category membership while the procedural system gradually associates categorization responses with regions of perceptual space via reinforcement learning. COVIS assumes that rule-based categorization is mediated by a broad neural network that includes the prefrontal cortex [ventral lateral PFC (vLPFC)], anterior cingulate [middle ACC (mACC)], head of the caudate nucleus, globus pallidus (GP), medial dorsal nucleus (MDN) of the thalamus, and hippocampus. The key structures in the COVIS procedural learning system are the putamen and/or the body and tail of the caudate nucleus, the GP, the ventral anterior and ventral lateral thalamic nuclei (VA/VL), and premotor cortex [supplementary motor area (SMA)]. SPEED (Ashby et al., 2007) extends the COVIS procedural system to account for the development of automaticity in II tasks by adding

cortico-cortical projections from sensory cortex directly to the relevant areas of premotor and motor cortex [SMA, dorsal premotor cortex (PMd), ventral premotor cortex (PMv), posterior ACC (pACC), and M1]. This model assumes that a major role of the subcortical path through the striatum is to train these cortico-cortical projections. Thus, SPEED assumes that the development of automaticity is a gradual transfer of control from the striatum to the cortex. Finally, we also examined the preSMA because this area has been implicated as a key structure in a network that might mediate the competition between rule-based and procedural learning systems (e.g., via the hyperdirect pathway through the striatum; Ashby & Crossley, 2010; Hikosaka & Isoda, 2010).

The ROIs were also grouped by network. The cortical motor ROIs consisted of M1, SMA, PMd, PMv, and pACC. The non-motor cortical ROIs were vlPFC, hippocampus, preSMA, and mACC. The basal ganglia regions included the striatum (head of the caudate, body and tail of the caudate, and putamen) and striatal output regions (GP, MDN, and VA/VL).

The anatomical boundaries of each ROI were created using the FSL Harvard-Oxford Cortical Structural Atlas or the Harvard-Oxford Subcortical Structural Atlas. The mACC and pACC were defined by taking the anterior cingulate cortex (as defined by the Harvard-Oxford Cortical Structural atlas) and dividing it based on structural landmarks (Vogt, 2005). PMd and PMv were created using the premotor cortex (as defined by the Harvard-Oxford Cortical Structural Atlas) and dividing the regions as defined by Picard and Strick (2001). The vlPFC was defined by combining the PFC with the inferior and middle frontal gyri (as defined by the Harvard-Oxford Cortical Structural Atlas) and removing all motor and premotor areas. The head of the caudate was drawn according to Nolte (2008; see Figures 19.1–19.2 and 19.4–19.5). To obtain the body and tail of the caudate, the head of the caudate was removed from the caudate mask found in the Harvard-Oxford Cortical Structural Atlas. All other ROIs were directly taken from the Harvard-Oxford structural atlases.

ROI analyses were performed on both the group maps (resulting from the high-level analyses) and on the participant maps (resulting from the mid-level analyses). For each ROI, the median percent signal change (%SC) was retained as a measure of activation. The high-level ROI analyses provided an overall look at the activation for that region, while the mid-level ROI analyses were used to correlate %SC with response accuracy for the stimulus contrast. To ensure that outliers did not drive these correlations, all correlations were recomputed with the largest outlier removed.

3. Results

3.1 Behavioral Performance

The behavioral data from this experiment were analyzed by Hélie, Waldschmidt, and Ashby (2010) as part of a much more extensive study that included many more participants, two other rule-based category structures, and two other purely behavioral experiments. Nearly every participant in these experiments completed 23 sessions of categorization with the same category structure. Sessions 1 through 20 were as described here. Sessions 21 – 23 were purely behavioral. In session 21, the location of the response keys was switched. Participants were informed of this switch and asked to categorize each stimulus as accurately as possible. In session 22 the response keys returned to their original locations and performance recovered to session 20 levels. In session 23, participants completed a dual-task during categorization. The goal of Hélie et al. (2010) was to test whether the qualitative differences between rule-based and II categorization that exist during early learning persist after automaticity develops. Results showed that after session 13, all differences disappeared. For example, as previously mentioned, during early learning switching the response keys interferes with II categorization but not with rule-based

categorization, whereas rule-based but not II categorization is susceptible to dual-task interference. Hélie et al. (2010) found that after 20 sessions of practice, rule-based and II categorization were both impaired when the response keys were switched and neither was susceptible to dual-task interference.

The remainder of this section summarizes analyses for the behavioral data from the II condition. For more details, see Hélie et al. (2010). II categorization accuracy dropped significantly when the response keys were switched [from 94% to 84.9%; $t(7) = 3.69, p < .01$]. In session 23, accuracy on the dual task was high and categorization accuracy was not significantly different than on session 20 [$t(5) = 1.64, p > .05$]. Thus, after 20 sessions of training, II categorization behavior displayed two classic criteria of automaticity – resistance to dual-task interference and behavioral inflexibility. Even so, because II categorization displays these same two features after only one session of training, these results by themselves provide no evidence of automaticity. On the other hand, Hélie et al. (2010) showed that after the same amount of training as in the present study (and on the same stimuli), rule-based categorization becomes behaviorally inflexible and resistant to dual-task interference. Thus, the training experienced by our participants was enough to induce automaticity in rule-based categorization according to popular automaticity criteria.

The individual and mean accuracies for sessions 1 through 20 are shown in Figures 3a and 3b, respectively. Mean accuracy was 79.0% in Session 1 and increased significantly to 94.3% in Session 20 [$F(29,152) = 7.850, p < .001$]. Accuracy in each of the first 3 sessions was significantly less than asymptotic accuracy [session 1: $t(9) = -3.833, p < .01$; session 2: $t(9) = -2.359, p < .05$; session 3: $t(9) = -3.115, p < .05$], but accuracy in session 4 did not differ significantly from asymptotic performance [$t(9) = .09, p = .931$]. The variances in percent correct for sessions 2, 4, 10, and 20 (which were used for later correlation analyses) were 49.71, 4.87, 17.97, and 17.83, respectively. Model fits were also completed for all behavioral data. After session 1, the responses of all participants in every session, including every scanning session, were best accounted for by a model that assumed an II decision strategy.

Figure 4 shows the mean RT for each participant in each session. As mentioned previously, note that mean RT increased on scanning sessions. The thick curve is the best fitting power function $\overline{RT} = 0.373n^{-0.992} + 0.532$, where n is session number. The parameters in this model were estimated using the method of least squares on the mean RTs from all non-scanning sessions. A well-known result that has been replicated for many different tasks is that mean RT decreases as a power function of the amount of practice (e.g., Logan, 1992). Figure 4 shows that this result holds for II categorization. Figure 4 also suggests that RT was still decreasing when the experiment ended. Thus, even if classic automaticity criteria are met, it seems likely that learning-related changes were still occurring after 11,000 trials of practice.

In summary, a variety of evidence suggests that automaticity developed before the last scanning session. First, two popular automaticity criteria (behavioral inflexibility and resistance to dual-task interference) were met immediately after session 20 by both the II learners and participants who had the same amount of practice in a similar rule-based task. Second, accuracy had long since asymptoted and response time had nearly asymptoted by session 20. Third, all qualitative differences between rule-based and II performance disappeared after the 13th session. It is also safe to conclude that automaticity had not yet developed during the first two scanning sessions (sessions 2 and 4) since accuracy had not yet asymptoted for most participants. Classifying performance in the third scanning session (session 10) solely on the basis of the behavioral data is more difficult. A conservative position would be to conclude that categorization behavior was not yet automatic since

performance was still qualitatively different from the performance of control participants who participated in a similar rule-based task (Hélie et al., 2010).

3.2 Group-Maps Whole Brain Analyses

The whole brain group analyses are summarized in Table 1 and Figure 5. Table 1 lists the locations of the highest peaks within each activated cluster in each session. Figure 5 shows significant activation in 8 different axial slices for each group map during each scanning session. The results were generally consistent with other fMRI studies of procedural-learning tasks. For example, visual areas and M1 were active in each session. Peaks occurred in premotor areas on sessions 10 and 20, and the striatum was active in all sessions. Furthermore, there was little activation in the hippocampus or other medial temporal lobe structures in any sessions.

A 1-factor (4 levels) repeated measures ANOVA was used to identify voxels where activation changed significantly across sessions for the *stimulus – baseline* contrast (i.e., where baseline was defined as activation to the crosshair). This analysis produced an F-statistic for each voxel. To correct for multiple comparisons, significant clusters were identified as before by setting an intermediate initial threshold and then using a threshold on cluster size derived from Gaussian random field theory that yielded an experiment-wise false positive rate of $\alpha = .05$ (Friston et al., 1994). This process revealed three significant clusters. Table 2 contains the Z values and coordinates of each peak in the three clusters. The peaks were in the following regions: ACC (medial and posterior), left extrastriate cortex (V2 or Brodmann Area 18), left M1, and postcentral gyrus.

3.3 ROI Analyses

This section describes the mid-level analysis, in which each ROI was examined for every participant and session.

3.3.1. Feedback Analyses—For the *feedback - baseline* contrast, no region had more than 3 participants with active voxels. Therefore, we examined the *baseline - feedback* contrast to identify regions that were significantly deactivated during feedback processing. Table 3 gives the proportion of participants who had significant feedback-related deactivation in each ROI. Note that the only regions in which more than half the participants showed feedback-related deactivations in every session were the vlPFC and the motor regions.

Table 4 lists the median %SC (from the group map) in every ROI for which at least half of the participants showed significant feedback-related deactivation. Note that %SC peaked in every motor ROI on session 4. Within the striatum, the largest deactivations in sessions 2, 4, and 20 were in the putamen.

3.3.2 Stimulus-Related Activation—Table 5 displays the proportion of participants who had active voxels in each ROI across sessions for the *stimulus - baseline* contrast. If more than half of the participants had no active voxels in an ROI, that ROI was removed from further analysis. This removed all four sessions of the following ROIs from further analysis: head of the caudate, body and tail of the caudate, MDN, GP, and VA/VL. In addition, sessions 2, 4, and 10 were removed for the hippocampus. For each remaining ROI, median %SC was examined across sessions (high-level analysis). These values are shown in Table 6. As expected, %SC generally increased across sessions in all active areas, increasing from non-significance on session 2 to significance by session 20 (Kelly & Garavan, 2005).

3.3.3 Correlations between Stimulus-Related Activation and Accuracy—To explore the relationship between activation in an ROI and performance in the task, we computed correlations across participants between median %SC and session accuracy. These values are listed in Table 7 for each ROI and session. Correlations are not reported for the hippocampus for sessions 2, 4, or 10 because fewer than half the participants showed significant hippocampal activation in any of these sessions. Note that for 8 of the 9 ROIs the correlations follow an identical pattern: increasing up to session 10 and then decreasing in session 20. The only exception is the vIPFC, where the correlation begins decreasing a session earlier. With random data, the probability that this pattern would occur in any one ROI is 0.125 (i.e., $.5^3$). Thus, using a sign test we can reject the null hypothesis that the pattern occurred by chance in favor of the alternative that the pattern represents a real trend in the correlations (8 successes in 9 trials, $p < .001$). Since the peak correlation occurs in session 10, a critical question is whether the correlations from this session are significant. Since session 10 correlations from different ROIs are positively correlated (because the same accuracy data were used to compute all these correlations), a Bonferroni correction for multiple comparisons is not appropriate. Instead we corrected for multiple comparisons by controlling the false discovery rate at .05 (Benjamini & Hochberg, 1995). Using this approach, all session 10 correlations are significant. A different question is whether any of the session 20 correlations are significant. The same false discovery rate procedure revealed no significant correlations on session 20. The range of accuracy values across participants in session 20 was about the same as in session 10, so the absence of significant correlations in session 20 was not due to a restricted range (i.e., the across-participant accuracy variance was 17.97 in session 10 and 17.83 in session 20).

For each ROI in sessions 10 and 20, we also removed the largest outlier (i.e., the point furthest from the best-fitting line) and recomputed the correlation. No significance decisions were changed by this analysis. Thus, none of the session 10 correlations were significant because of outliers and outliers did not prevent any session 20 correlations from reaching significance.

Median %SC is a crude measure of activation since it only changes across sessions if there are activation changes in most voxels. Therefore, after 20 sessions of practice, it is perhaps not surprising that such a crude measure of activation failed to correlate with accuracy in any ROI. For this reason, we conducted a finer-grained analysis for the data in session 20. Specifically, for each voxel within the Table 7 ROIs, we computed the across-participant correlation between session 20 accuracy and the mid-level t -statistic from the *stimulus – baseline* contrast. Each resulting correlation was then converted to a z -statistic using Fisher's z -transformation. Under the null hypothesis that the correlations are zero, the resulting z values have a z distribution. Finally, within each ROI, significance decisions were made on these z -values using a false discovery rate of .05 (Benjamini & Hochberg, 1995). Using this finer-grained analysis, only two regions contained voxels where activation was significantly correlated with accuracy on session 20 – SMA (with 10 significant voxels) and preSMA (with 23 significant voxels). All other ROIs had zero significant voxels. Thus, in session 20, activation was correlated with accuracy only in cortical areas.

4. Discussion

The following results stood out. 1) Automaticity developed sometime between sessions 10 and 20. 2) Pre-automatic performance depended heavily on the striatum, and more specifically on the putamen, rather than the body and tail of the caudate nucleus. 3) Automatic performance depended only on cortical regions. Thus, extended practice was associated with a gradual transfer of control from the basal ganglia to cortex. 4) Feedback processing was associated with widespread deactivations, which were especially consistent

across sessions in the vLPFC and premotor and motor regions of cortex. 5) The overall effects of practice were consistent with the existing literature on the development of automaticity. We now consider each of these conclusions in more detail.

4.1 Automaticity Developed Sometime between Sessions 10 and 20

A number of results suggested that our participants began responding automatically sometime between sessions 10 and 20 (and therefore between the third and fourth scanning sessions). First, the participants' performance met a number of behavioral criteria for automaticity during (or immediately after) session 20, whereas during session 10 performance on our II task was still qualitatively different from the performance of control participants in a similar rule-based task (Hélie et al., 2010). In addition, correlations between activation and accuracy increased in many ROIs through session 10 and then decreased essentially to zero by session 20 (e.g., the putamen). In session 20, the only voxels where activation was correlated with accuracy were in preSMA and SMA. Thus, activation patterns and their relation to performance were qualitatively different in session 20 compared to earlier sessions. Taken together, these results all suggest that according to current criteria, participants were responding automatically during session 20, but not during sessions 2, 4, or 10.

4.1 Pre-Automatic Performance Depended on the Putamen but not the Caudate Nucleus

For the majority of participants, no voxels in the caudate nucleus showed significant categorization-related activation in any scanning session. In contrast, most participants did have significant putamen activation on every scanning day. Furthermore, at least in session 10, activation differences across participants within the putamen were significantly correlated with categorization accuracy. In other words, during session 10, participants who showed more putamen activation were more accurate in their categorization decisions than participants who showed less activation.

These results stand somewhat in contrast to several previous studies that have reported significant learning-related activation in the caudate nucleus during II categorization (Cincotta & Seger, 2007; Nomura et al., 2007; Seger & Cincotta, 2002). However, there are at least two important differences between the present study and these previous studies that might account for these seemingly discrepant results. First, even on our first scanning session (session 2), participants already had 600 trials of previous training. The COVIS theory of category learning (Ashby et al., 1998) predicts that during initial II category learning, participants will experiment with explicit categorization rules and that this process of hypothesis testing is mediated by a broad neural network that includes the caudate nucleus (as well as other structures, including the prefrontal cortex, anterior cingulate, and hippocampus). Many studies of rule-based category learning support this prediction (Hélie, Roeder, & Ashby, 2010; Konishi et al., 1998; Lie, Specht, Marshall, & Fink, 2006; Monchi, Petrides, Petre, Worsley, & Dagher, 2001). So one possibility is that at least some of the caudate activation noted in previous studies came from trials where the participant was experimenting with explicit rules. This hypothesis suggests that we may have observed caudate activation if our first scanning session was on day 1 rather than on day 2.

A second important difference is that the designs of Nomura et al. (2007) and Seger and Cincotta (2002) were not able to separate activation due to feedback processing from activation due to categorization. In the present study, feedback presentation was modeled as a separate event and therefore we were able to isolate activation due to categorization. As a result, one possibility is that at least some of the caudate activation reported in previous fMRI studies of II categorization may have been driven primarily by feedback processing. This hypothesis is supported by two studies in which humans or monkeys learned arbitrary

stimulus-response associations (Haruno & Kawato, 2006; Williams & Eskandar, 2006). Both studies reported that caudate activity was mostly related to feedback and reward processing, whereas putamen activity was related to learning. We found no feedback-related activation in the caudate, but again this may have been because of the extensive previous training all of our participants received before their first scanning session.

As mentioned above, there is strong theoretical reason to expect learning-related activation in the putamen during II category learning. The rationale is as follows. A number of studies have shown that switching the response keys interferes with II categorization performance (but not with rule-based categorization), even in the first session of training (Ashby et al., 2003; Maddox et al., 2004b; Maddox et al., 2010; Spiering & Ashby, 2008). The caudate and putamen both project to cortex via the globus pallidus/substantia nigra pars reticulata and the thalamus. The sensorimotor (i.e., posterior) regions of the putamen project primarily into motor areas of cortex and SMA (or SMA-proper) via the ventral lateral (VL) nucleus of the thalamus (Matelli & Luppino, 1996). The SMA is densely interconnected with primary motor cortex and with other premotor areas (Dum & Strick, 2005). In contrast, the caudate and anterior putamen project to cortex primarily via the medial dorsal (MD) and ventral anterior (VA) thalamic nuclei. The MD nucleus projects widely into all anterior areas of frontal cortex, including PFC, whereas VA projects most heavily into preSMA and SEF (Matelli & Luppino, 1996). The SEF projects strongly to the frontal eye fields. However, SEF and preSMA are both densely interconnected with the PFC (Wang, Isoda, Matsuzaka, Shima, & Tanji, 2005), and neither area sends direct projections to any premotor or motor areas that mediate finger movements (Dum & Strick, 2005). As a result, a strong case can be made that preSMA is most properly classified as a prefrontal region (Akkal, Dum, & Strick, 2007). Thus, with the exception of the SEF, the caudate projects primarily to prefrontal regions of cortex. If the caudate was the only critical striatal region for II category learning, one would not expect a switch of the response buttons to interfere with learning because of the caudate's poor connections to premotor and motor cortex. Since such interference does occur, the sensorimotor regions of the putamen may play a more prominent role in II category learning than previously assumed. Our results support this hypothesis. This conclusion is also consistent with single-unit recording results that show learning-related changes in the firing of cells in SMA during a traditional procedural-learning task (Lee & Quessy, 2003). In addition, several other fMRI studies of II categorization have reported significant task-related activation in the putamen (e.g., Cincotta & Seger, 2007; Seger & Cincotta, 2002).

4.3 Automatic Performance Depended only on Cortical Regions

Although we found strong evidence that the putamen played an important role in mediating the categorization response up to session 10, we found no evidence that the basal ganglia made any meaningful contribution to categorization performance in session 20. First, more than half of all participants had no significant activation in session 20 in the caudate nucleus, the globus pallidus, or any thalamic nuclei that are targets of basal ganglia output (i.e., MDN, VA, and VL). Second, although there was significant activation in the putamen on session 20, there were no voxels in the putamen in which activation was correlated with accuracy.

As mentioned earlier, one prominent proposal is that the development of automaticity is mediated by a gradual transfer of control from the associative striatum (e.g., the caudate nucleus) to the sensorimotor striatum (e.g., the posterior putamen) (Belin et al., 2009; Costa, 2007; Yin & Knowlton, 2006). This hypothesis predicts the presence of significant putamen activation in session 20, which we observed, but it also seems to predict that this activation should be correlated with behavioral performance. In contrast to this prediction, putamen activation on session 20 was uncorrelated with accuracy.

On the other hand, several studies have reported that a small group of neurons within the sensorimotor striatum increase their task-related activity with extended training, even while the total number of active neurons is decreasing (Barnes, Kubota, Hu, Jin, Graybiel, 2005; Tang, Pawlak, Prokopenko, & West, 2007). More subtle changes of this type could lead to reduced correlations with performance as training progresses because there would be fewer neurons mediating the behavior. This hypothesis seems to predict that putamen activation should decrease with practice, along with the correlations. In contrast, we observed reduced correlations but not reduced activation (see Tables 5 and 6).

A contrasting theory predicts that automaticity is associated with a transfer of control from the basal ganglia to cortico-cortical projections from the relevant sensory areas directly to the premotor areas (e.g., SMA) that initiate the behavior (Ashby, Ennis, & Spiering, 2007). The idea is that the basal ganglia use dopamine-mediated reinforcement learning (i.e., at cortico-striatal synapses) to gradually activate the correct post-synaptic targets in premotor cortex, which thereby enables Hebbian learning at cortico-cortical synapses to acquire the correct associations. In this way, the basal ganglia train up the cortical representations. The only ROIs in our study with voxels in which activation was correlated with accuracy on session 20 were SMA and preSMA. Thus, the present results are consistent with this view of automaticity.

A number of other results also support this theory. First, two separate studies examined neural changes as automaticity developed during rule-based categorization. One study used human participants and fMRI (Hélie, Roeder, & Ashby, 2010) and one used monkeys and single-unit recordings (Muhammad, Wallis, & Miller, 2006). Both studies reported vigorous cortical and striatal activity after automaticity had developed, and they both reported evidence that only the cortical activity played a role in mediating the categorization behavior. Second, Turner, McCairn, Simmons, and Bar-Gad (2005) reported that temporarily blocking striatal output to cortex (by injecting a GABA agonist into the internal segment of the globus pallidus) had little or no effect on the ability of monkeys to produce a highly practiced motor sequence. Third, several studies have shown that disconnecting the bird homologue of the basal ganglia completely blocks new song learning, but has little effect on the expression of well-learned songs (Doupe, Perkel, Reiner, & Stern, 2005). Fourth, some Parkinson's disease patients are able to emit an automatic motor response when presented with a familiar visual cue (e.g., kicking a ball), despite difficulties in initiating novel voluntary movements (Asmus, Huber, Gasser, & Schöls, 2008).

The theory that the development of automaticity is mediated by a transfer of control to cortex predicts the performance-related activation in SMA on session 20, but not in preSMA. As mentioned above, the connections of the preSMA classify it more as a prefrontal region than as a premotor region (Akkal et al., 2007), so it is unlikely that the preSMA has any direct role in selecting the categorization response on each trial. Instead, Hikosaka and Isoda (2010) have hypothesized that the preSMA is crucial for switching between controlled and automatic responding. The preSMA contains neurons that have a suppressive effect on such switches and others that have a facilitatory effect (Hikosaka & Isoda, 2010). So one intriguing, but speculative possibility is that the role of the preSMA in the present experiment is to prevent prefrontal or declarative memory strategies from interrupting the automatic response process that has emerged after 11,000 trials of practice.

4.4 Feedback Effects

The main effect of feedback was deactivation, which was seen most consistently in the vIPFC and in motor and premotor regions of cortex. There are several reasons why widespread feedback-related activations should not be expected in the present experiment. First, all participants had 600 trials of practice before the first scanning session, so feedback

was never as important in this study as in experiments that examine early learning. Second, feedback processing requires attention and effort in rule-based category-learning tasks, but in the II task used here feedback processing is largely automatic (Maddox et al., 2004a). Our results reinforce the hypothesis that there is no general feedback processing network, and instead that feedback may be processed differently by different memory systems.

4.5 General Practice Effects

Although we know of no previous fMRI studies that examined the development of automaticity during II categorization, a number of studies have examined the development of automaticity in other tasks. As mentioned above, results have been different depending on whether the task recruits declarative or procedural memory systems (e.g., Kelly & Garavan, 2005). Declarative memory tasks (e.g., working memory) generally show reductions in neural activation with practice, especially in regions associated with executive attention (Chien & Schneider, 2005). Included in this list are various regions within the PFC, orbitofrontal cortex, mACC, and the insula. However, procedural-learning tasks (e.g., sequence learning) have often reported increases in neural activation with extended practice. Generally these increases are found primarily in motor regions (Hazeltine, Grafton, & Ivry, 1997; Honda et al., 1998; Karni et al., 1995, 1998).

We found evidence for changes in categorization-related activation with training in only four regions. One of these was M1, where Table 6 shows an increasing trend across sessions. This increase is consistent with previous studies of procedural-learning tasks. An increasing trend was also found in extrastriate visual area² V2. Such increases in visual areas are frequently seen in studies where subjects receive extended training with a limited set of similar visual stimuli (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; Op de Beeck, Baker, DiCarlo, & Kanwisher, 2006; Weisberg, van Turennout, & Martin, 2007). Thus, overall, our results are consistent with prior studies that have examined the effects of extended training on procedural-learning tasks.

5. Conclusions

Many real world categorization decisions require information integration (e.g., deciding whether an animal is a wolf or German shepherd; deciding whether an x-ray displays a tumor). This article described the results of an extensive experiment in which participants received more than 11,000 trials of feedback training on the same II categories. Sessions 2, 4, 10, and 20 were conducted inside an MRI scanner, thus allowing an extended look at how neural activation changes as automaticity develops in this important task. The results suggested that between sessions 2 and 10 learning depended on the putamen but not the caudate nucleus, and that automatic judgments were mediated exclusively within cortex.

Acknowledgments

This research was supported by NIH Grant R01 MH3760-2. The author's would like to thank Jessica Hein, Jessica L. Roeder, Erik Rush, and Maria Schellenberger for help with data collection and/or analysis. We would also like to thank each of our anonymous reviewers and Sebastien Hélie for their helpful comments.

References

Akkal D, Dum RP, Strick PL. Supplementary motor area and presupplementary motor area: Targets of basal ganglia and cerebellar output. *J Neurosci.* 2007; 27:10659–10673. [PubMed: 17913900]

²V2 difference maps were created from the mid-level t-statistics that compared sessions 2 and 20. As before, a spatial-extent cluster-based threshold was used to correct for multiple comparisons (with the experiment-wise error rate set to .01). Results showed voxels with greater activation during session 20, but none with greater activation during session 2.

- Anderson JLR, Jenkinson M, Smith SM. Non-linear registration, aka spatial normalisation. FMRIB technical report TR07JA2. 2007
- Ashby FG, Alfonso-Reese LA, Turken AU, Waldron EM. A neuropsychological theory of multiple systems in category learning. *Psychol Rev.* 1998; 105:442–481. [PubMed: 9697427]
- Ashby FG, Boynton G, Lee WW. Categorization response time with multidimensional stimuli. *Percept Psychophys.* 1994; 55:11–27. [PubMed: 8036090]
- Ashby FG, Crossley MJ. Interactions between declarative and procedural-learning categorization systems. *Neurobiol Learn Mem.* 2010; 94:1–12. [PubMed: 20304078]
- Ashby FG, Ell SW, Waldron EM. Procedural learning in perceptual categorization. *Mem Cognit.* 2003; 31:1114–1125.
- Ashby FG, Ennis JM. The role of the basal ganglia in category learning. *Psychology of Learning and Motivation.* 2006; 46:1–36.
- Ashby FG, Ennis JM, Spiering BJ. A neurobiological theory of automaticity in perceptual categorization. *Psychol Rev.* 2007; 114:632–656. [PubMed: 17638499]
- Ashby FG, Gott RE. Decision rules in the perception and categorization of multidimensional stimuli. *J Exp Psychol Learn Mem Cogn.* 1988; 14:33–53. [PubMed: 2963894]
- Ashby FG, Maddox WT. Human category learning. *Annu Rev Psychol.* 2005; 56:149–178. [PubMed: 15709932]
- Ashby FG, Maddox WT, Bohil CJ. Observational versus feedback training in rule-based and information-integration category learning. *Mem Cognit.* 2002; 30:666–677.
- Ashby FG, Turner BO, Horvitz JC. Cortical and basal ganglia contributions to habit learning and automaticity. *Trends Cogn Sci.* 2010; 14:191–232. [PubMed: 20363177]
- Ashby, FG.; Valentin, VV. Multiple systems of perceptual category learning: Theory and cognitive tests. In: Cohen, H.; Lefebvre, C., editors. *Categorization in cognitive science*. New York: Elsevier; 2005.
- Asmus F, Huber H, Gasser T, Schöls L. Kick and rush: Paradoxical kinesis in Parkinson disease. *Neurology.* 2008; 71:695. [PubMed: 18725599]
- Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature.* 2005; 437:1158–1161. [PubMed: 16237445]
- Belin D, Jonkman S, Dickinson A, Robbins TW, Everitt BJ. Parallel and interactive learning processes within the basal ganglia: Relevance for the understanding of addiction. *Behav Brain Res.* 2009; 199:89–102. [PubMed: 18950658]
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol.* 1995; 57:289–300.
- Brainard DH. The psychophysics toolbox. *Spat Vis.* 1997; 10:433–436. [PubMed: 9176952]
- Chein JM, Schneider W. Neuroimaging studies of practice-related change: fMRI and meta-analytic evidence of a domain-general control network for learning. *Brain Res Cogn Brain Res.* 2005; 25:607–623. [PubMed: 16242923]
- Cincotta CM, Seger CA. Dissociation between striatal regions while learning to categorize via feedback and via observation. *J Cogn Neurosci.* 2007; 19:249–265. [PubMed: 17280514]
- Costa RM. Plastic corticostriatal circuits for action learning: What's dopamine got to do with it? *Ann N Y Acad Sci.* 2007; 1104:172–191. [PubMed: 17435119]
- Crossman EREW. A theory of the acquisition of speed-skill. *Ergonomics.* 1959; 2:153–166.
- Doupe AJ, Perkel DJ, Reiner A, Stern EA. Birdbrains could teach basal ganglia research a new song. *Trends Neurosci.* 2005; 28:353–363. [PubMed: 15935486]
- Dum RP, Strick PL. Frontal lobe inputs to the digit representations of the motor areas on the lateral surface of the hemisphere. *J Neurosci.* 2005; 25:1375–1386. [PubMed: 15703391]
- Erickson MA, Kruschke JK. Rules and exemplars in category learning. *J Exp Psychol Gen.* 1998; 127:107–140. [PubMed: 9622910]
- Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC. Assessing the significance of focal activations using their spatial extent. *Hum Brain Mapp.* 1994; 1:210–220.

- Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC. Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nat Neurosci.* 1999; 2:568–573. [PubMed: 10448223]
- Haruno M, Kawato M. Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *J Neurophysiol.* 2006; 95:948–959. [PubMed: 16192338]
- Hazeltine E, Grafton ST, Ivry R. Attention and stimulus characteristics determine the locus of motor-sequence encoding. A PET study. *Brain.* 1997; 120:123–140. [PubMed: 9055803]
- Heaton, RK. A manual for the Wisconsin Card Sorting Test. Odessa, FL: Psychological Assessment Resources; 1981.
- Hélie S, Roeder JL, Ashby FG. Evidence for cortical automaticity in rule-based categorization. *J Neurosci.* 2010; 30:14225–14234. [PubMed: 20962243]
- Hélie S, Waldschmidt JG, Ashby FG. Automaticity in rule-based and information-integration categorization. *Atten Percept Psychophys.* 2010; 72:1013–1031. [PubMed: 20436197]
- Hikosaka O, Isoda M. Switching from automatic to controlled behavior: Cortico-basal ganglia mechanisms. *Trends Cogn Sci.* 2010; 14:154–161. [PubMed: 20181509]
- Honda M, Deiber MP, Ibanez V, Pascual-Leone A, Zhuang P, Hallett M. Dynamic cortical involvement in implicit and explicit motor sequence learning. A PET study. *Brain.* 1998; 121:2159–2173. [PubMed: 9827775]
- Jenkinson M, Bannister PR, Brady JM, Smith SM. Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *Neuroimage.* 2002; 17:825–841. [PubMed: 12377157]
- Jenkinson M, Smith SM. A global optimisation method for robust affine registration of brain images. *Med Image Anal.* 2001; 5:143–156. [PubMed: 11516708]
- Karni A, Meyer G, Jezard P, Adams MM, Turner R, Ungerleider LG. Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature.* 1995; 377:155–158. [PubMed: 7675082]
- Karni A, Meyer G, Rey-Hipolito C, Jezard P, Adams MM, Turner R, Ungerleider LG. The acquisition of skilled motor performance: Fast and slow experience-driven changes in primary motor cortex. *Proc Natl Acad Sci U S A.* 1998; 95:861–868. [PubMed: 9448252]
- Kelly AMC, Garavan H. Human functional neuroimaging of brain changes associated with practice. *Cereb Cortex.* 2005; 15:1089–1102. [PubMed: 15616134]
- Konishi S, Nakajima K, Uchida I, Kameyama M, Nakahara K, Sekihara K, Miyashita Y. Transient activation of inferior prefrontal cortex during cognitive set shifting. *Nat Neurosci.* 1998; 1:80–84. [PubMed: 10195114]
- Lee D, Quessy S. Activity in the supplementary motor area related to learning and performance during a sequential visuomotor task. *J Neurophysiol.* 2003; 89:1039–1056. [PubMed: 12574479]
- Lie CH, Specht K, Marshall JC, Fink GR. Using fMRI to decompose the neural processes underlying the Wisconsin Card Sorting Test. *Neuroimage.* 2006; 30:1038–1049. [PubMed: 16414280]
- Logan GD. Shapes of reaction-time distributions and shapes of learning curves: A test of the instance theory of automaticity. *J Exp Psychol Learn Mem Cogn.* 1992; 18:883–914. [PubMed: 1402715]
- Love BC, Medin DL, Gureckis TM. SUSTAIN: A network model of category learning. *Psychol Rev.* 2004; 111:309–332. [PubMed: 15065912]
- Maddox WT, Ashby FG, Bohil CJ. Delayed feedback effects on rule-based and information-integration category learning. *J Exp Psychol Learn Mem Cogn.* 2003; 29:650–662. [PubMed: 12924865]
- Maddox WT, Ashby FG, Ing AD, Pickering AD. Disrupting feedback processing interferes with rule-based but not information-integration category learning. *Mem & Cogn.* 2004a; 32:582–591.
- Maddox WT, Bohil CJ, Ing AD. Evidence for a procedural learning-based system in perceptual category learning. *Psychon Bull Rev.* 2004b; 11:945–952. [PubMed: 15732708]
- Maddox WT, Glass BD, O'Brien JB, Filoteo JV, Ashby FG. Category label and response location shifts in category learning. *Psychol Res.* 2010; 74:219–236. [PubMed: 19471959]

- Matelli M, Luppino G. Thalamic input to mesial and superior area 6 in the macaque monkey. *J Comp Neurol*. 1996; 327:59–87. [PubMed: 8841922]
- Monchi O, Petrides M, Petre V, Worsley K, Dagher A. Wisconsin Card Sorting revisited: Distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *J Neurosci*. 2001; 21:7733–7741. [PubMed: 11567063]
- Muhammad R, Wallis JD, Miller EK. A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum. *J Cogn Neurosci*. 2006; 18:974–989. [PubMed: 16839304]
- Newell, A.; Rosenbloom, PS. Mechanisms of skill acquisition and the law of practice. In: Anderson, JR., editor. *Cognitive Skills and their Acquisition*. Hillsdale, NJ: Erlbaum; 1981. p. 1-55.
- Nolte, J. *The human brain: An introduction to its functional anatomy*. 6. Mosby; Philadelphia: 2008.
- Nomura EM, Maddox WT, Filoteo JV, Ing AD, Gitelman DR, Parrish TB, Mesulam MM, Reber PJ. Neural correlates of rule-based and information-integration visual category learning. *Cereb Cortex*. 2007; 17:37–43. [PubMed: 16436685]
- Nosofsky RM, Palmeri TJ. An exemplar-based random walk model of speeded classification. *Psychol Rev*. 1997; 104:266–300. [PubMed: 9127583]
- Op de Beeck HP, Baker CI, DiCarlo JJ, Kanwisher NG. Discrimination training alters object representations in human extrastriate cortex. *J Neurosci*. 2006; 26:13025–13036. [PubMed: 17167092]
- Parker A, Ngu H, Cassaday HJ. Odour and Proustian memory: Reduction of context-dependent forgetting and multiple forms of memory. *Appl Cogn Psychol*. 2001; 15:159–171.
- Paul EJ, Boomer J, Smith DJ, Ashby FG. Information-integration category learning and the human uncertainty response. *Mem Cognit*. (in press).
- Picard N, Strick PL. Imaging the premotor areas. *Curr Opin Neurobiol*. 2001; 11:663–672. [PubMed: 11741015]
- Reber PJ, Gitelman DR, Parrish TB, Mesulam MM. Dissociating conscious and nonconscious knowledge with fMRI. *J Cogn Neurosci*. 2003; 15:574–685. [PubMed: 12803968]
- Schneider W, Shiffrin RM. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol Rev*. 1977; 84:1–66.
- Seger CA. How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. *Neurosci Biobehav Rev*. 2008; 32:265–278. [PubMed: 17919725]
- Seger CA, Cincotta CM. Striatal activity in concept learning. *Cogn Affect Behav Neurosci*. 2002; 2:149–161. [PubMed: 12455682]
- Serences J. A comparison of methods for characterizing the event-related BOLD timeseries in rapid fMRI. *Neuroimage*. 2004; 21:1690–1700. [PubMed: 15050591]
- Shiffrin RM, Schneider W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychol Rev*. 1977; 84:127–190.
- Smith S. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002; 17:143–155. [PubMed: 12391568]
- Spiering BJ, Ashby FG. Response processes in information-integration category learning. *Neurobiol Learn Mem*. 2008; 90:330–338. [PubMed: 18550397]
- Tang C, Pawlak AP, Prokopenko V, West MO. Changes in activity of the striatum during formation of a motor habit. *Eur J Neurosci*. 2007; 25:1212–1227. [PubMed: 17331217]
- Tintera J, Gawehn J, Bauermann T, Vucurevic G, Stocter P. New partially parallel acquisition technique in cerebral imaging: preliminary findings. *Eur Radiol*. 2004; 14:2273–2281. [PubMed: 15300396]
- Turner, RS.; McCairn, K.; Simmons, D.; Bar-Gad, I. Sequential motor behavior and the basal ganglia: Evidence from a serial reaction time task in monkeys. In: Bolam, JP.; Ingham, CA.; Magill, PJ., editors. *The basal ganglia VIII. Advances in Behavioral Biology*. Vol. 56. New York: Springer; 2005. p. 563-574.
- Vogt BA. Pain and emotion interactions in subregions of the cingulate gyrus. *Nat Rev Neurosci*. 2005; 6:533–544. [PubMed: 15995724]

- Waldron EM, Ashby FG. The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychon Bull Rev.* 2001; 8:168–176. [PubMed: 11340863]
- Wang Y, Isoda M, Matsuzaka Y, Shima K, Tanji J. Prefrontal cortical cells projecting to the supplementary eye field and presupplementary motor area in the monkey. *Neurosci Res.* 2005; 53:1–7. [PubMed: 15992955]
- Weisberg J, van Turennout M, Martin A. A neural system for learning about object function. *Cereb Cortex.* 2007; 17:513–521. [PubMed: 16581980]
- Williams ZM, Eskandar EN. Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nat Neurosci.* 2006; 9:562–568. [PubMed: 16501567]
- Woolrich MW. Robust group analysis using outlier inference. *Neuroimage.* 2008; 41:286–301. [PubMed: 18407525]
- Yin HH, Knowlton BJ. The role of the basal ganglia in habit formation. *Nat Rev Neurosci.* 2006; 7:464–476. [PubMed: 16715055]
- Zeithamova D, Maddox WT. Dual-task interference in perceptual category learning. *Mem Cognit.* 2006; 34:387–398.
- Zeithamova D, Maddox WT. The role of visuospatial and verbal working memory in perceptual category learning. *Mem Cognit.* 2007; 35:1380–1398.

Figure 1.

The category structures for this experiment. Each stimulus was a circular sine-wave grating that varied across trials in spatial frequency (i.e., bar width) and bar orientation. Each plus denotes an exemplar of category A and each star denotes an exemplar of category B. The disk in the upper left is the prototype of category A (i.e., the category mean) and the disk in the lower right is the prototype of category B. The dashed line is the optimal category decision boundary.

Figure 2.

Timing of a trial scaled in TR (one $TR = 2,000$ ms). The number of blank TR s between stimulus and feedback events was jittered with a truncated geometric distribution with $p = 0.5$ (*max.* 5 TR s). When at least one blank TR was inserted between the feedback and the following stimulus ($\sim 50\%$ of the trials), a crosshair was displayed in the second half of the TR immediately preceding stimulus presentation.

Figure 3.
 (a) Individual accuracies of participants across the 20 sessions. (b) The mean accuracy across sessions. Error bars are standard errors. Scanning occurred on sessions 2, 4, 10, and 20.

Figure 4. Response times for individual subjects. Note that sessions 2, 4, 10 and 20 were completed inside the scanner. The thick line is the power function that best fits the mean RTs from sessions outside the scanner.

Figure 5.

Group maps for the *stimulus – baseline* contrast (from the whole-brain analysis) for each scanning session. From left to right, 8 different axial slices are shown for each scanning session at the following MNI152 z coordinates: -8 , -4 , 0 , 4 , 8 , 12 , 16 , and 20 respectively.

Table 1

Coordinates of voxels showing peak *stimulus* – *baseline* activation within each significant cluster for every scanning session.

Brain region	Cluster Size	Max Z	x	y	z
Session 2					
Precuneus cortex	19154	4.48	-8	-74	40
R. Frontal pole	9958	4.74	38	64	-2
L. Insula	5170	4.40	-36	20	-2
Paracingulate gyrus	2058	4.71	4	28	38
Posterior cingulate gyrus	560	3.78	2	-32	20
Session 4					
R. Supramarginal gyrus	27008	4.60	46	-32	46
R. Frontal operculum	13571	4.59	42	18	-2
Session 10					
R. Cuneal cortex	22839	4.29	14	-74	34
R. Paracingulate gyrus	20327	4.55	12	30	28
L. M1	863	3.78	-24	-10	62
R. Middle frontal gyrus	783	4.02	26	4	52
Session 20					
R. Occipital fusiform gyrus	57455	4.50	28	-74	-18

Note: coordinates are in MNI152 space. The minimum cluster size thresholds used were 38, 42, 45, and 43 for sessions 2, 4, 10 and 20 respectively. A matlab file named mni2tal.m, which is freely available at several web sites, converts from MNI coordinates to Talairach coordinates.

Table 2

Coordinates of voxels showing peak activation in each cluster where *stimulus* – *baseline* activation changed significantly across sessions.

Cluster	Location	Max Z	x	y	z
1	Anterior cingulate cortex	3.64	0	16	32
	"	3.60	0	-2	30
	"	3.52	-4	22	28
	"	3.50	4	18	32
	"	3.10	-8	4	34
2	"	3.09	0	26	22
	Left Extrastriate cortex	4.31	-28	-90	6
	"	4.18	-18	-102	10
	"	3.87	-24	-94	8
	"	3.84	-24	-98	8
3	"	3.77	-20	-92	6
	"	3.70	-14	-102	8
	Left M1	4.12	-36	-22	64
	"	3.59	-28	-12	64
	Postcentral gyrus	3.44	-40	-26	48
	"	3.40	-46	-20	50
	"	3.20	-64	-22	38
	"	3.19	-56	-24	42

Note: The *F*-values from the ANOVA were converted to *z*-values via the FSL ftoz routine. Coordinates are in MNI152 space.

Table 3

Proportion of participants with active voxels in each ROI across sessions for the *baseline - feedback* contrast.

ROI	Session 2	Session 4	Session 10	Session 20
Cortical motor regions				
M1	.70	.90	.70	.90
Supplementary motor area	.60	.90	.50	.70
Dorsal premotor	.60	.90	.60	.90
Ventral premotor	.70	.90	.70	.70
Posterior anterior cingulate cortex	.60	.90	.60	.70
Non-motor cortical regions				
Ventrolateral prefrontal cortex	.70	.90	.70	.60
Presupplementary motor area	.60	.90	.40	.60
Middle anterior cingulate cortex	.60	.90	.40	.70
Hippocampus	.40	.60	.70	.30
Striatum				
Head of the caudate	.30	.60	.20	.50
Body and tail of the caudate	.20	.50	.30	.20
Putamen	.50	.50	.30	.50
Striatal output regions				
Globus pallidus	.10	.40	.20	.30
Medial dorsal nucleus	.30	.50	.40	.40
Ventral anterior/Ventral lateral nucleus of the thalamus	.40	.40	.50	.40

Note: All ROIs are bilateral.

Table 4

Median percent signal change for the *baseline - feedback* contrast (deactivation during feedback) in predefined ROIs for regions in which at least half of the participants had significant deactivation.

ROI	Session 2	Session 4	Session 10	Session 20
Cortical motor regions				
M1	0.06647**	0.1166***	0.08465***	0.0804***
Supplementary motor area	0.09549***	0.1477***	0.104**	0.1075***
Dorsal premotor	0.05279*	0.09831**	0.06652**	0.06348**
Ventral premotor	0.07825***	0.1341***	0.09782***	0.08934***
Posterior anterior cingulate cortex	0.06777***	0.09587***	0.07333***	0.08253***
Non-motor cortical regions				
Ventrolateral prefrontal cortex	0.03697*	0.07031***	0.0581***	0.07376***
Presupplementary motor area	0.1185***	0.1483***	-	0.1247***
Middle anterior cingulate cortex	0.09129***	0.1231***	-	0.09882***
Hippocampus	-	0.05278**	0.03491*	-
Striatum				
Head of the caudate	-	0.05936***	-	0.04113***
Body and tail of the caudate	-	0.05347***	-	-
Putamen	0.03165***	0.07424***	-	0.05954***
Striatal output regions				
Medial dorsal nucleus	-	0.08186***	-	-
Ventral anterior/Ventral lateral nucleus of the thalamus	-	-	0.05029**	-

Note: All ROIs are bilateral.

*
 $p < .05$,

**
 $p < .01$,

 $p < .001$

Table 5

Proportion of participants with active voxels in each ROI across sessions for the *stimulus - baseline* contrast.

ROI	Session 2	Session 4	Session 10	Session 20
Cortical motor regions				
M1	.80	.70	.70	.80
Supplementary motor area	.50	.60	.60	.50
Dorsal premotor	.60	.70	.50	.70
Ventral premotor	.70	.70	.60	.60
Posterior anterior cingulate cortex	.60	.60	.60	.60
Non-motor cortical regions				
Ventrolateral prefrontal cortex	.80	.70	.70	.70
Presupplementary motor area	.50	.60	.60	.50
Middle anterior cingulate cortex	.60	.60	.60	.60
Hippocampus	.30	.40	.40	.80
Striatum				
Head of the caudate	.20	.40	.20	.40
Body and tail of the caudate	.20	.20	.20	.30
Putamen	.80	.70	.60	.70
Striatal output regions				
Globus pallidus	.00	.10	.10	.20
Medial dorsal nucleus	.30	.30	.20	.30
Ventral anterior/Ventral lateral nucleus of the thalamus	.30	.20	.30	.40

Note: All ROIs are bilateral.

Table 6

Median percent signal change for the *stimulus - baseline* contrast in predefined ROIs for regions in which at least half of the participants had significant activation.

ROI	Session 2	Session 4	Session 10	Session 20
Cortical motor regions				
M1	0.01764	0.01024	.05764	0.07027*
Supplementary motor area	0.00912	0.01446	.05640	0.07028*
Dorsal premotor	0.00252	−0.00157	0.02993	0.04492
Ventral premotor	0.02742	0.01609	0.06036*	0.06970*
Posterior anterior cingulate cortex	0.00611	0.01858	.05069*	0.06629*
Non-motor cortical regions				
Ventrolateral prefrontal cortex	0.09476**	0.10480**	0.10030**	0.11420**
Presupplementary motor area	0.02142	0.02525	0.07103**	0.07802**
Middle anterior cingulate cortex	0.03897**	0.06071***	0.07946***	0.10260***
Hippocampus	-	-	-	0.04303
Striatum				
Putamen	.05906***	0.06308***	0.07309***	0.08341***

Note: all ROIs are bilateral.

*
 $p < .05$,

**
 $p < .01$,

 $p < .001$

Table 7

Correlations (across-subjects) between median percent signal change and behavioral accuracy for predefined ROIs (bilateral) for the *stimulus - baseline* contrast for regions in which at least half of the participants had significant activation.

ROI	Session 2	Session 4	Session 10	Session 20
Cortical motor regions				
M1	−0.003	0.548	0.715	0.006
Supplementary motor area	0.033	0.430	0.795	0.077
Dorsal premotor	−0.191	0.501	0.700	−0.040
Ventral premotor	−0.006	0.604	0.756	0.162
Posterior anterior cingulate cortex	−0.047	−0.023	0.752	−0.083
Non-motor cortical regions				
Ventrolateral prefrontal cortex	0.734	0.795	0.629	0.068
Presupplementary motor area	0.046	0.219	0.584	0.398
Middle anterior cingulate cortex	−.051	0.212	0.671	0.412
Hippocampus	-	-	-	−0.065
Striatum				
Putamen	0.296	0.333	0.573	−0.110

Note: All session 10 correlations are significant at a false discovery rate of .05. No session 20 correlations are significant at this same criterion.