Dissociating hippocampal and basal ganglia contributions to category learning using stimulus novelty and subjective judgments

Carol A. Seger a,b,⁎, Christina S. Dennison b, Dan Lopez-Paniagua a, Erik J. Peterson a,b, Aubrey A. Roark a

⁎ Corresponding author. Department of Psychology, Colorado State University, USA. Fax: +1 970 491 1032.
E-mail addresses: Carol.Seger@colostate.edu (C.A. Seger), csd83@lamar.colostate.edu (C.S. Dennison), lopezpan@rams.colostate.edu (D. Lopez-Paniagua), erik.exists@gmail.com (E.J. Peterson), aubreyant@gmail.com (A.A. Roark).

Article history:
Received 30 September 2010
Revised 13 December 2010
Accepted 11 January 2011
Available online 19 January 2011

Keywords:
Classification learning
Implicit learning
Striatum
Medial temporal lobe
Salience
Functional connectivity

ABSTRACT

We identified factors leading to hippocampal and basal ganglia recruitment during categorization learning. Subjects alternated between blocks of a standard trial and error category learning task and a subjective judgment task. In the subjective judgments task subjects categorized the stimulus and then instead of receiving feedback they indicated the basis of their response using 4 options: Remember: Conscious episodic memory of previous trials. Know-Automatic: Automatic, rapid response accompanied by conscious awareness of category membership. Know-Intuition: A "gut feeling" without fully conscious knowledge of category membership. Guess: Guessing. In addition, new stimuli were introduced throughout the experiment to examine effects of novelty. Categorization overall recruited both the basal ganglia and posterior hippocampus. However, basal ganglia activity was found during Know judgments (both Automatic and Intuition), whereas posterior hippocampus activity was found during Remember judgments. Granger causality mapping indicated interactions between the basal ganglia and hippocampus, with the putamen exerting directed influence on the posterior hippocampus, which in turn exerted directed influence on the posterior caudate nucleus. We also found a region of anterior hippocampus that showed decreased activity relative to baseline during categorization overall, and showed a strong novelty effect. Our results indicate that subjective measures may be effective in dissociating basal ganglia from hippocampal dependent learning, and that the basal ganglia are involved in both conscious and unconscious learning. They also indicate a dissociation within the hippocampus, in which the anterior regions are sensitive to novelty, and the posterior regions are involved in memory based categorization learning.

© 2011 Elsevier Inc. All rights reserved.

Introduction

Humans are endowed with a wide variety of learning and memory systems that allow us to adapt to our environment. One system underlies procedural learning and includes the basal ganglia; another underlies declarative memory and includes the hippocampus. Both systems are often recruited during categorization tasks, in which subjects learn to associate stimuli with responses indicating category membership via trial and error. The present study is aimed at further characterizing and dissociating the roles played by the basal ganglia and hippocampus in categorization learning. First, we used subjects’ subjective judgments to dissociate trials performed on the basis of memory (and found to recruit the hippocampus), from trials performed in a subjectively automatic or intuitive way (found to recruit the basal ganglia). Second, we manipulated stimulus novelty and found that it dissociated the roles of anterior and posterior hippocampus in category learning.

The basal ganglia in categorization learning

There is broad convergent evidence that the basal ganglia are important for categorization learning (see Seger and Miller, 2010; Shohamy et al., 2008; Ashby and O’Brien, 2005 for reviews). Although the basal ganglia are recruited during a wide variety of categorization tasks, they are particularly important for feedback-based categorization, in which subjects learn via trial and error. Basal ganglia activation during trial and error categorization learning is universally reported in functional imaging studies (Poldrack and Foerde, 2008; Seger, 2008; Nomura et al., 2007). Learning is impaired in patients with basal ganglia damage due to Parkinson and Huntington diseases (Knowlton et al., 1996a,b; Shohamy et al., 2004; Filoteo et al., 2007). Trial and error categorization tasks are similar to many commonly used tasks in the rodent and nonhuman primate literatures (such as arbitrary visuomotor learning and instrumental conditioning): in all subjects perceive a stimulus, make a contingent behavioral response,
and finally receive feedback or reward (Seger, 2009). Lesion and electrophysiological research with these species provide additional evidence that the basal ganglia are critical for trial and error categorization learning.

Recent research has focused on the question of what specific role or roles the basal ganglia play in categorization learning. We have argued that different corticostratial loops (interacting networks passing through particular cortical and basal ganglia regions) subserve different processes within categorization (Seger et al., 2010). Our views are based on the established function of the basal ganglia in modulating cortical processing to enable selection of responses, behaviors, and strategies appropriate to the current behavioral context (Seger, 2008; Frank, 2005; Gurney et al., 2004). Categorization requires first evaluating the stimulus and mapping it to the appropriate category; we argue that the visual loop passing through the posterior caudate and projecting to the pre-SMA is important for this process (Ashby et al., 1998). Subsequently, it is necessary to select the appropriate motor response indicating category membership; this function recruits the motor loop passing through the posterior putamen. Finally, categorization learning requires evaluating the feedback received and updating stimulus-category representations; these processes recruit the executive and motivational loops passing through the anterior caudate and ventral striatum, respectively.

The hippocampus in categorization learning

The evidence for the involvement of the hippocampus in categorization learning is mixed. Early research using the most studied categorization task, the probabilistic classification or “weather prediction” task, found that amnesics showed preserved early learning (Knowlton et al., 1994), indicating independence from the hippocampus. However, later research found amnesics were impaired in the task (Hopkins et al., 2004). The first functional imaging study of probabilistic classification found decreased anterior hippocampal activity across blocks of training (Poldrack et al., 1999). A follow-up study found a more complex pattern of results: an increase in anterior hippocampal activity at the beginning of training which rapidly dropped to below baseline levels (Poldrack et al., 2001; Little et al., 2006). Foerde and colleagues (Foerde et al., 2006) found that under single task learning conditions anterior hippocampal activity correlated with measures of learning, whereas under dual task conditions this relationship did not hold. Studies with categorization tasks other than the probabilistic classification task, including the information integration task (Cincotta and Seger, 2007) and trial and error prototype learning tasks (Zeithamova et al., 2008; Little et al., 2006) have reported hippocampal activity, but in relatively posterior regions. These patterns of results indicate that the anterior and posterior hippocampus may be playing different roles in categorization learning.

The transient anterior hippocampal activation at the beginning of probabilistic classification learning implies a role in initial processing of each novel stimulus. This is consistent with recent research finding anterior hippocampal sensitivity to stimulus novelty (Daselaar et al., 2006; Strange et al., 1999; Blackford et al., 2010; Wittmann et al., 2007; Köhler et al., 2005; Poppenk et al., 2010). Meeter and colleagues (Meeter et al., 2008) argued that the hippocampus may be required to encoding the stimulus and setting up representations that may then be accessed by other neural systems.

Positive posterior hippocampal activation in some tasks indicates that this region may be recruited to aid task dependent performance. Studies that found positive posterior hippocampal activity typically used larger sets of visually complex or variable stimuli, and had stimulus–category relationships that were deterministic rather than probabilistic. Both factors may make using a declarative memory strategy more useful and thus increase reliance on memory strategies.

We designed our task in accordance with these results to maximize the chance that we would modulate activity in both anterior and posterior hippocampal regions. We used a large study set of visually complex stimuli that were deterministically assigned to category, thus maximizing the utility of memory-based strategies. We also manipulated stimulus novelty separately from training block by introducing novel stimuli across the entire experiment. This avoids the confound between stimulus novelty and early overall learning present in previous studies. On the basis of previous studies, we predicted that the anterior hippocampus would be more sensitive to the novelty manipulation, whereas the posterior hippocampus would be recruited across the categorization learning process.

Interactions between basal ganglia, hippocampus, and dopaminergic systems

A number of patterns of interaction between the basal ganglia and hippocampal systems have been observed across species and tasks. In some tasks, including the probabilistic categorization in humans, and place versus response (or “habit”) learning in rodents the relationship is competitive: basal ganglia increases are paired with hippocampal decreases, and vice versa (Lee et al., 2008). In other situations, notably spatial learning, basal ganglia and hippocampus are recruited in parallel (Doeller et al., 2008). It is unknown whether interactions between basal ganglia and hippocampus are direct, or mediated by other neural systems. There is some evidence that the prefrontal cortex can mediate the balance during probabilistic classification learning (Poldrack and Rodriguez, 2004), and that emotional arousal can mediate, via the amygdala, habit versus place learning (Wingard and Packard, 2008).

The basal ganglia and hippocampus are both targets of dopamine projections from the midbrain (ventral tegmental area and substantia nigra). The dopamine signal reflects the valence of the feedback received, and the degree to which the feedback was or was not expected (often referred to as prediction error; Hollerman and Schultz, 1998). Dopamine is crucial for synaptic plasticity within the striatum; long term potentiation only occurs in the presence of dopamine (Reynolds and Wickens, 2002). In the hippocampus, long term potentiation can occur without dopaminergic input, but is increased when dopamine is present (Düzel et al., 2010; Bethus et al., 2010). Dopamine plays an important role in ensuring that memories are motivationally relevant and adaptive for the organism (Shohamy and Adcock, 2010).

We used Granger causality mapping (GCM) to explore interactions between basal ganglia, hippocampus, and the dopaminergic midbrain. GCM identifies directed influences across time to and from neural “seed” regions. We predicted that directed influences within the basal ganglia would proceed from putamen, to anterior caudate, to posterior caudate, as we have found in previous categorization studies (Lopez-Paniagua and Seger, (in press); Seger et al., 2010). We also predicted that the dopaminergic midbrain would exert directed influence on both the hippocampus and basal ganglia, in accordance with the known anatomy of dopaminergic projections to these regions. We did not have a priori hypotheses as to how the hippocampus and basal ganglia would interact due to the paucity of knowledge about anatomical and functional connections between these regions.

Subjective measures in the study of memory

Subjective measures have been used widely in the memory literature to dissociate forms of memory. One approach assesses whether subjects are subjectively aware (explicit) or not aware (implicit) of what they have learned (Seger, 1994). A more recently developed subjective judgment task is the Remember–Know task, which can be used to dissociate the contributions of full recollection from those of familiarity in recognition memory (Gardiner et al.,
Subjects make a Remember response if they have a full episodic memory of the stimulus including features such as the time and place at which they first encountered it. A Know response is used when subjects believe that they have encountered the stimulus before, but their memory is not accompanied by any episodic details. Within the medial temporal lobe, recollection and familiarity generally recruit the hippocampus and parahippocampal cortex (particularly perirhinal cortex), respectively (Eichenbaum et al., 2007; Diana et al., 2007). There is controversy over whether familiarity is fully dissociable from recollection, or if it might just reflect a weaker form of memory (Kirwan et al., 2008). However, it widely agreed that Remember responses identify strong episodic memories.

We adapted the Remember-Know task to the category learning domain in order to identify on a trial by trial basis when subjects were relying on episodic memory and when they were making their response based on some other information. During pilot research subjects indicated that they were able to use the Remember response to indicate trials on which they responded based on episodic memory. However, subjects indicated that the Know response was difficult to use because they actually experienced two subjectively different non-memory based states. One state was described as performing automatically, knowing the response without thinking about it. The other was described as having a sense of category response via intuition or a gut feeling. We examined these two states separately in our study, and referred to them as “Know-Automatic” and “Know-Intuition”; full descriptions of each are given in the appendix. Within our study, we predicted that Remember trials would recruit the hippocampus, and Know trials would recruit the basal ganglia. Within Know trials, if the basal ganglia are limited to relatively implicit forms of learning (in the sense of not being fully verbalizable; see the Discussion for a more detailed treatment of this issue), than basal ganglia recruitment should be greatest for Know-Intuition trials. Alternatively, if the basal ganglia are recruited across both implicit and explicit learning, then they may be recruited by both Know-Automatic and Know-Intuition.

Overview of the current study

We taught subjects to categorize complex abstract visual stimuli (fractal designs) into two separate categories. All stimuli were deterministically assigned to category. Assignment was arbitrary: that is, there was no pattern or rule that could be used to categorize stimuli and the category membership for each stimulus must be learned independently. Across the blocks new stimuli were continually added to the learning set so that within each block, stimuli varied in the numbers of times they had been previously encountered. This allowed us to separate processes associated with processing novel stimuli from those associated with overall categorization learning. Subjects alternated performing blocks of a standard trial and error categorization learning task and a subjective judgment task. In the category learning task, they viewed a single stimulus, pressed a button to indicate its category membership, then received feedback. In the subjective judgment task, instead of receiving feedback subjects were prompted to report the subjective basis of their response using one of four options: Remember, Know-Automatic, Know-Intuition, and Guess. Finally, we examined correlations between basal ganglia, hippocampus, and dopaminergic midbrain using Granger causality mapping across all trials.

Materials and methods

Subjects

Eleven subjects were recruited from the University of Colorado School of Medicine, Denver community (Aurora, CO). All subjects were healthy, right-handed adults (7 males, 4 females) with an average age of 25.5 years (range: 18–31). Subjects were fluent speakers of English, and were screened for a history of neurological and psychiatric disorders, use of psychoactive substances, and contraindications to MRI (i.e., metallic implants and claustrophobia). Each subject participated in one scanning session, completing the procedure described below. Functional data from one additional participant was excluded from analyses due to a high rate of missed trials (over 40%) combined with below chance accuracy on the completed trials.

Categorization tasks

Two tasks were utilized: a categorization task, and a subjective judgment task, which alternated in blocks of 28 trials. There were 6 blocks of each task, divided evenly across 3 separate scans. The categorization task was similar in structure to ones used in previous studies in our lab (Seger and Cincotta, 2005). A “weather prediction” cover story was used, in which subjects were informed that they should learn which stimuli predict rain and which predict sun. Each trial consisted of participants being presented with a single arbitrary visual stimulus and making a button press to indicate which category the stimulus belonged to. Following each response, participants were given feedback as to whether their response was correct or incorrect. During scanning, responses were made via right and left handed response boxes; “sun” was indicated via a left index finger button press, and “rain” via a right index finger button press.

The subjective judgment task was identical to the categorization task with one notable exception. Rather than receiving feedback as to whether their choice was correct or incorrect, subjects were asked to make a judgment as to the process used to categorize the previous stimulus as either rain or sun. To do so, participants used one of four types of responses: Remember, Know-Automatic, Know-Intuition, or Guess. Although we refer to these responses by these names here, subjects learned to refer to each response with a neutral letter: B (Remember), A (Know-Automatic), C (Know-Intuition), and D (Guess). These neutral letter names were used because previous research has shown that subjects vary in their preexisting ideas of what is meant by “Remember” and “Know” and we wanted to avoid biasing that might be entailed by using these words as a shorthand (McCabe and Geraci, 2009). A Guess response was included because previous research has shown that without a Guess response, the Know response is often used by subjects to indicate low memory confidence (Eldridge et al., 2002; Gardiner et al., 2002).

Our Remember instructions were based on earlier research using the Remember-Know task in recognition memory, following the instructions initially developed by Rajaram (1993), and utilized by McCabe and Geraci (2009). In these instructions, subjects are given a list of several dimensions that can be used as a basis for a remember response, including “something personal from the time you studied it”, “something that happened in the room (noise)”, and are also given examples of real world situations in which they might have experienced recollective or familiarity based memory. We used similar examples for our Remember instructions (see the Appendix). We developed our instructions for the two Know responses based on extensive piloting and interviews with subjects. As for Remember responses, we gave subjects several dimensions that might correspond to each type of Know response, and included real world examples of when they might have experienced a similar form of learning.

Subjects were given extensive description of each type of response (full instructions are included in the Appendix). They requested to read the task instructions at home before the experimental session. At the beginning of the session, they read the instructions again and discussed the meaning of each of the options with the experimenter. Subjects then performed 30–50 practice trials under the supervision.
of the experimenter; the stimuli used in the practice trials were different from those used during the scanning session. During scanning, judgments A through D were made using the following
fingers on the left and right hand response boxes: A by the left middle
finger, B by the left index finger, C by the right index finger, and D by
the right middle finger so that there was a left to right spatially
compatible mapping of letters onto fingers.

Stimuli were fractal images (Seattle Fractals Digital Art, Seattle,
WA) created using Tearazon v29 fractal drawing software (Stephen
Ferguson, Houston, TX). The 24 images selected for use in the study
were chosen on the basis of being distinctive patterns without easily
verbalized patterns or features in common across categories. The
relationship between each stimulus and the two possible responses
was deterministic: half of the stimuli were always associated with
"rain" and half were always associated with "sun". In the initial block
of the study, subjects were exposed to 6 stimuli. After 30 (± 8) trials,
one stimulus was removed and replaced with a novel one. The same
replacement procedure was carried out throughout the entire length
of the task in order to continually expose subjects to novel stimuli and
to ensure that stimuli with each block varied in number of times they
were previously seen. Analyses examining effects of stimulus
repetition number and novelty were limited to blocks 2–5; block 1
was excluded because all of the stimuli were being presented for the
first through fourth time, and including this block would result in a
confound between stimulus repetition number and overall amount of
training on the categorization task. Similarly, stimuli appearing for
more than 8 repetitions could not be evenly distributed across blocks
(none were present in block 2, and few in blocks 3 and 4); so these
stimuli were also excluded so as to avoid a confound stimulus
repetition and overall amount of training. One subject was excluded
from the novelty analysis due to technical problems resulting in his
receiving a different distribution of repetition numbers across blocks.

Within each block, trials were arranged pseudo-randomly with the
constraint that stimuli were never repeated in immediate succession.
The trial length varied depending on a number of factors. In the
categorization portion of the task, stimuli were presented for
1500 ms, during which subjects were required to make their response
indicating category membership. Following the response, there was a
brief delay lasting 250 ms, before subjects received trial-specific
feedback presented for 1250 ms. After feedback, subjects were
presented with a fixation cross which signaled the beginning of a
new trial. This inter-trial interval was jittered. The values for the jitter
were 1500 ms, 3000 ms, 4500 ms, and 6000 s, even multiple of the TR.

These values were randomly sampled from a geometric distribution
using Matlab (The MathWorks Inc, Natick, MA), i.e., 50% of the trials
1500 ms inter-trial interval, 25% of the trials had 3000 ms inter-trial
interval, 12.5% of the trials had 4500 ms inter-trial interval, and 12.5%
of the trials had 6000 ms inter-trial interval. In the response judgment
portion of the task, trial length and timing of individual events was
similar to that of the categorization portion. However, instead of the
feedback screen, subject were presented with a screen that allowed
them to choose either a type A, B, C or D response, which lasted
2750 ms. Following this response screen, subjects were once again
presented with a jittered inter-trial fixation screen.

fMRI image acquisition

Images were obtained on a research-dedicated 3.0 T whole-body
MRI scanner (GE Healthcare, Milwaukee, WI) at the Brain Imaging
Center at the University of Colorado Denver (Aurora, CO). The scanner
was equipped with an 8-channel, high-resolution phased array head
coil using GE’s Array Spatial Sensitivity Encoding Technique (ASSET)
software. Anatomical images were collected using a T1-weighted
SPGR sequence (minimal TR; TE, 3.95 ms; TI, 950 ms; FA, 10°; FOV,
220-mm; 256 x 256 coronal matrix; 166 1.2-mm slices). The structural
images were used to verify proper slice selection and to determine the
sites of functional activation (i.e., voxels that were found to be
significantly activated during the functional scan were overlaid on the
high-resolution structural images). Functional images were re-
structured from 26 axial oblique slices obtained using a T2*-weighted
EPI-Gradient-Recalled Echo sequence (TR, 1500 ms; TE, 30 ms; FA,
64°; FOV, 220-mm; 64° x 64 matrix; 4.0-mm slices; no inter-slice gap),
in order to measure BOLD signal change. Additionally, the first five
volumes, recorded before longitudinal magnetization reached a
steady state, were discarded.

Visual stimuli were presented to subjects using a magnet-compatible
projector that projects visual images onto a mirror attached to the RF
head coil. A computer running E-Prime 2.0 experiment software
(Psychology Software Tools Inc., Pittsburg, PA) was used to control
stimulus presentation and interface with two two-button magnet
compatible response boxes placed one in each hand. Earplugs and
headphones were provided to protect the subjects’ hearing. Head
movement was minimized using small foam pads placed on each side of
the head inside the RF head coil.

Image preprocessing

Image analysis was performed using BrainVoyager QX 2.1 (Brain
Innovation, Maastricht, The Netherlands). Functional data was first
subjected to preprocessing, consisting of 1) three dimensional motion
correction using trilinear interpolation, 2) slice scan time correction
using cubic spline interpolation, 3) temporal data filtering with a
high-pass filter of 3 cycles in the time course and 4) linear trend
removal. Each subject’s high-resolution anatomical image was
normalized to the Talairach and Tournoux (1988) brain template.
The normalization process consisted of two steps: an initial rigid body
translation into the AC–PC plane, followed by an elastic deformation
into the standard space performed on 12 individual sub-volumes. The
resulting set of transformations was applied to the subject’s functional
image volumes to form volume time course representations to be
used in subsequent statistical analyses. Finally, the volume time
course representations were spatially smoothed using a Gaussian
kernel, full-width at half maximum (FWHM) of 6.0 mm.

Whole brain analyses

Whole brain analyses were performed using the general linear
model (GLM) implemented in Brain Voyager. The epochs for each
condition were convolved with a prototypical hemodynamic function.
Conditions were then compared by running the GLM using separate
subject predictors, which treated subjects as a random effect. In order
to increase the probability of identifying functionally significant
clusters of activation while controlling for the rate of false-positives,
one of two correction methods was used. Our primary method was
the False Discovery Rate (FDR) method with a threshold of q < 0.05 was
used (Genovese et al., 2002). This approach is most suitable for
controlling the overall level of false positives across the entire data set
without making any assumptions about particular features of the data
set, such as the smoothness of the data (Bennett et al., 2009). The
other is the cluster level method (Forman et al., 1995) extended to 3D
maps (Goebel et al., 2006) and implemented in the Brain Voyager
Cluster Threshold plug-in. The minimum voxel cluster size was
determined for each statistical map using Monte Carlo simulations
with 1000 iterations, with an alpha threshold of .05. To effectively
control for false positives, this method requires that data smoothness
be calculated separately for each analysis; the final cluster size based
on these estimates can vary widely from analysis to analysis, and can
lead to type II errors when the cluster size is large. Given the
limitations on the cluster size threshold approach, we deemed it more
suitable for analyses for which we had strong a priori hypotheses
about the neural systems that should be recruited; specifically, the
hippocampus for Remember judgments and the basal ganglia for
Know judgments. The particular correction method used is indicated for each analysis in the figures and table.

Two whole brain analyses were performed. The first analysis broke trials into conditions on the basis of the task was performed (Standard Categorization: “Cat” and categorization followed by subjective judgments “Judge”) and whether the subject correctly categorized the stimulus or not (“Corr” and “Inc”, respectively). This resulted in four explicitly defined conditions: Cat-Corr, Cat-Inc, Judge-Corr, and Judge-Inc. In addition, an implicit baseline condition was defined as the mean of all remaining time points. Epochs were defined as the first two image volumes acquired following the presentation of the stimulus (TR = 1500 ms; epoch length = 2TR = 3 s); this epoch length included the entire time for stimulus presentation and category response in both Categorization and Judgment trials. However, the feedback and judgment parts of the trial followed so soon afterwards that it is not possible to state with certainty that only initial stimulus–response processes were reflected in the measures of neural activity.

The second whole brain analysis was limited to judgment trials and examined the four judgment types. This analyses included four conditions: Remember, Know-Automatic, Know-Intuition, and Guess, in addition to the implicit baseline. Epochs for this analysis started at the time of stimulus presentation and extended across a total of 4 TRs (6 s), which encompassed the entire amount of time allowed for all parts of the trial. We performed a follow-up analysis incorporating only correct trials to control for the possibility that any differences between judgment types were due to differences in proportion of correct trials. There were very few correct Guess trials, so they were excluded from this analysis (mean of 4.2, 1.4, and 1.2 correct Guess trials per subject in runs 1–3, respectively). The overall pattern of results did not differ qualitatively from the analysis including both correct and incorrect trials; however, some of the differences did not reach the set statistical threshold, most likely due to lower power that resulted from the smaller number of trials.

### Anatomical ROI analyses of the hippocampus

In order to avoid problems with circular analysis that could result from using functionally based ROIs from the whole brain analyses as the bases of our novelty comparisons (Kriegeskorte et al., 2009), we instead identified regions of the hippocampus anatomically. Four ROIs were identified, from the anterior to posterior axis of the hippocampus bilaterally (8 ROIs total); they are illustrated in the right panel of Fig. 7 and referred to as the Anterior, Mid-Anterior, Mid-Posterior, and Posterior ROIs. Statistical tests within these ROIs were performed using the Brain Voyager ROI GLM tool. All contrasts were calculated as within subject random effects analyses, controlling for between subject variability, and subjected to an alpha value of $p < .05$.

### Granger causality mapping

Granger causality mapping (GCM) was used to explore effective connectivity between the striatum and other brain regions. This study applied Roebroeck and colleagues’ (2005) procedure, as implemented within BrainVoyager, for creating causality maps that provide a measure of directed influence. Reference (or “seed”) regions were defined on the basis of the contrast comparing correctly categorized trials with baseline; they included right and left regions within the basal ganglia (putamen, head of the caudate, and body of the caudate) and the hippocampus (both anterior and posterior). We also included a seed region from the dopaminergic brainstem (general region encompassing the VTA/SNr) given the theoretical interest in how dopaminergic systems modulate both basal ganglia and hippocampus (Shohamy and Wagner, 2008; Adcock et al., 2006).

Target regions were defined as any voxel not included in the reference region (y). Influence measures were then calculated from the reference to target region (FX → Y), target to reference region (FY → X) and total linear dependence between the reference and target regions (FX, Y) by repeatedly pairing the time-course maps of each voxel in these regions. Time course data was sampled from all trials in every condition across all scans. The subsequent GCM analysis was performed using preprocessed data, which included spatial smoothing. Directed influences to and from the reference region were calculated by subtracting the influence of the target reference region from the influence from the reference to the target region (FX → Y – FY → X) for every voxel to calculate a difference (dGCM). Thus, effective connectivity was described as: dGCM = FX–Y − FY–X (see Roebroeck et al., 2005 for details). A positive difference value indicates FX → Y (reference → volume) influence, whereas negative difference values depict FY → X (volume → reference) influence. That is, activity in region Y is said to cause activity in region Y if the past activity in X can be used to statistically predict activity in Y more accurately than merely using the past patterns of activation in region Y. Effective connectivity maps were computed by first creating individual maps for each VOI for each subject, then comparing activation across maps using a voxelwise t-test examining whether activity was significantly different from zero; $p < .05$, cluster threshold of 20.

We confirmed each pattern of directed influence by using seeds in both regions; for example the directed influence from putamen to anterior caudate was detected both as an influence from the putamen seed region to the anterior caudate, and an influence onto the anterior caudate seed region from the putamen. In addition, we required that the pattern be present in both the right and the left hemispheres. It should be noted that GCM has some limitations. One is that it assumes that the hemodynamic response has the same shape and latency in all regions (David et al., 2008; Deshpande et al., 2009); it is unknown whether this assumption holds across all the regions we have examined.

### Results

#### Behavioral results

**Learning across scans and stimulus repetitions**

As shown in Fig. 1 (top), mean accuracy increased across scans of the study; each scan included 2 categorization and 2 judgment blocks. A $2 \times 3$ within subjects ANOVA with factors of condition (Categorization vs. Judgment) and Scan (1–3) revealed a main effect of scan, $F(2,18) = 29.49, p < .001$, and a main effect of condition, $F(2,18) = 17.40, p = .002$, but no interaction. The main effect of condition is likely due to two factors: judgment half-blocks always followed categorization half-blocks, and new stimuli were always introduced and repeated for the first time in a categorization block.

As shown in Fig. 1 (bottom), subjects were at chance at the first presentation of each stimulus, and accuracy rose across subsequent presentations, with most of the increase in accuracy completed by trial 7. A one-way within subjects ANOVA indicated a main effect of stimulus repetition number, $F(23,192) = 2.7, p < .001$.

**Judgment type distribution**

The distribution of the four judgment types across scans is shown in Fig. 2 (top). A $4 \times 3$ ANOVA with factors of Judgment Type (Remember, Know-Automatic, Know-Intuition, and Guess) and Scan (1–3) showed a main effect of judgment type $F(1.68,15.08) = 10.77, p = .002$, and a significant interaction between scan and judgment type, $F(2,14,19.24) = 6.57, p = .006$. Maluchy's test indicated that the assumption of sphericity had been violated for response type (chi-square = 15.54) and the scan by response type interaction (chi-square = 44.00), so degrees of freedom were adjusted using Greenhouse–Geisser estimates of sphericity (epsilon = .558 and .356, respectively). Post hoc tests using a Bonferroni correction revealed that subjects made significantly made
more Know-Automatic, Know-Intuition, and Remember responses compared to Guess ($p = 0.021, 0.027, \text{ and } 0.003,$ respectively).

We examined each Judgment type separately to see if its use changed across scans using one-way repeated measures ANOVA. For Know-Automatic judgments there was a significant main effect of Scan, $F(2,18) = 10.34, p = .001.$ Post hoc tests confirm that subjects made more Know-Automatic judgments across time: there was a significant increase from Scan 1 to Scan 2 ($p = .055),$ and from Scan 1 to Scan 3 ($p = .006),$ but not between Scans 2 and 3. For Know-Intuition judgments there was a significant main effect of Scan, $F(2,18) = 5.83, p = .011,$ but pairwise comparisons revealed no significant differences between scans once corrections for multiple comparisons were applied. There was no significant main effect of scan for Remember judgments. For Guess judgments there was a significant main effect of scan, $F(1.14,10.29) = 6.97, p = .022.$ Post hoc tests with a Bonferroni correction revealed a significant decrease in Guess responses between Scan 1 and Scan 3 ($p = .05$). Overall there were more Know-Automatic responses over time, and fewer Guess responses, but Remember and Know-Intuition responses did not change substantially across blocks.

Relation between judgment types and accuracy

Fig. 2 (bottom) shows the accuracy rate associated with each judgment type. A one-way ANOVA revealed significant effect of response type $F(3,32) = 94.29, p < .001.$ Tukey’s post hoc test showed that accuracy was higher for Know-Automatic than Remember ($p = 0.02),$ Know-Intuition ($p = 0.001$) or Guess ($p < 0.001$) responses. Remember responses were significantly more accurate than Guess responses ($p < 0.001),$ but did not significantly differ in accuracy from Know-Intuition ($p = 0.09$). Finally, there was a trend towards Know-Intuition being more accurate than Guess responses ($p = 0.052$).

Whole brain analyses

Activity during categorization and judgments

We first examined activity during correct categorization (Cat-Corr $-$ Baseline). This revealed patterns of activity similar to previous categorization learning studies, confirming that the differences in design in the present study did not significantly change the recruited neural systems. Within the basal ganglia (see Fig. 3, middle left), activity extended across all regions of the striatum, including ventral striatum, anterior and posterior putamen, and anterior and posterior caudate. As in previous studies, we identified a region of the medial temporal lobe that was overall less active during categorization than during baseline: this region is relatively anterior and closely matches the region identified in previous categorization studies (Poldrack et al., 1999, 2001; Foerde et al., 2006; Seger and Cincotta, 2006). In addition, a posterior region of the hippocampus was more active during categorization (see Fig. 3, lower left). Extensive bilateral frontoparietal networks were also recruited; of particular note these include the inferior frontal gyrus, lateral premotor cortex, dorsolateral prefrontal cortex, medial frontal cortex/supplementary motor area, and both superior and inferior posterior parietal regions (see Fig. 3, top left).

We then examined activity during subjective judgment trials (Judge-Corr $-$ Baseline), and compared standard categorization to judgment task (Judge-Corr $-$ Cat-Corr; recall that for both Judge-Corr and Cat-Corr the epoch was limited to the first 2TRs, 3 s total, following stimulus onset and did not included the time during which the subjective judgment was made). The regions recruited during judgment trials were very similar to the regions recruited in categorization. As in categorization, there was activity across all regions of the striatum, posterior hippocampus, and frontoparietal networks, combined with decreased activity in the anterior hippocampus. When directly compared, overall levels of activity were higher in Judge-Corr than Cat-Corr, but no regions of activity were exclusive to either the standard categorization or the judgment task. These analyses confirm that the initial phases of the Judgment task (viewing the stimulus and making the categorical response) were similar to the corresponding phases of the Categorization task.
Remember and know activations in basal ganglia and hippocampus

We first examined Remember and Know judgments separately, focusing on the basal ganglia and hippocampus. These conditions are compared with activation patterns during Categorization in Fig. 3. Overall, the hippocampus was associated with Remember trials, and the basal ganglia with Know trials. Specifically, in Remember trials (Remember\(^{-\text{N Baseline}}\)), we found bilateral activation of the posterior hippocampus in addition to deactivation of the anterior hippocampus. During Know trials (Know-Automatic + Know-Intuition\(^{-\text{N Baseline}}\)) there was bilateral activity in the head of the caudate. No suprathreshold basal ganglia activity was found for Remember trials, or suprathreshold hippocampal activity in Know trials. The recruitment patterns for the caudate and both hippocampal regions across all three judgment types are illustrated in Figs. 4 and 5, respectively.

A direct comparison of Remember and Know trials (Remember\(^{-}\)Know-Automatic + Know-Intuition; Table 1a) revealed greater activity in the right and left head of the caudate for Know than Remember trials, consistent with the pattern shown in Fig. 4. Although activation and deactivation in the posterior and anterior hippocampi, respectively, was found in Remember\(^{-}\)Baseline and not in Know\(^{-}\)Baseline, these differences were not significant in the direct comparison. As shown in Fig. 5, this may be due to differences in hippocampal recruitment in the two Know subconditions, with the larger differences being between Remember and Know-Automatic. To explore this possibility, we performed pairwise comparisons between Remember and each of the two Know conditions individually. We found greater activity in the right anterior hippocampus during Know-Automatic than Remember trials, consistent with the pattern illustrated in Fig. 5. There were no other activation differences in the basal ganglia or hippocampus.

We then examined whether the two types of Know judgments differed in their neural correlates (Know-Automatic\(^{-}\)Know-Intuition; see Table 1b). There were no differences between the two judgment types in their recruitment of basal ganglia or hippocampal systems. We also examined activation during Guess trials (Guess\(^{-}\)Baseline). There were no suprathreshold regions activated, which is likely due to the small number of Guess trials leading to low statistical power. The Guess condition was included mainly to reduce the contamination of the Know-Intuition condition by guess responses, and is not theoretically important in and of itself.

Remember and know activations in cortex

Cortically, widespread prefrontal and parietal lobe regions were recruited in both Know and Remember trials. As shown in Fig. 3, top row, there was a trend towards greater left prefrontal activity in Remember, and more right prefrontal activity in Know, although the only region that reached significance in the direct comparison was
greater activity in the right inferior prefrontal gyrus/anterior insula for Know.

There was a laterality difference in parietal lobe activity in the postcentral gyrus (primary somatosensory cortex), with greater activity in the Remember>Know and Know-Automatic>Know-Intuition contrasts in the right hemisphere, and the opposite pattern in the left hemisphere (greater activity for Know-Intuition>Know-Automatic and Know>Remember). This pattern is likely to be due to the sensorimotor demands of the task: both Remember and Know-Automatic responses were made by pressing buttons on the left hand response box, whereas Know-Intuition was indicated through a button on the right hand response box.

Novelty effects in the hippocampus

We examined the effects of stimulus novelty by comparing activity at the first presentation of each stimulus against later presentations (Presentation 1>Presentations 5–8) within each of the 8 anatomically defined hippocampal ROIs (see Anatomical ROI analyses of the hippocampus for details). For purposes of comparison, the Anterior ROI corresponds closely to the anterior region deactivated in the Correct categorization contrast; the Posterior ROI is close the posterior region recruited in Correct Categorization and in Remember. The Mid-Anterior and Mid-Posterior ROIs have intermediate anatomical positions. The three most anterior ROIs all showed significant novelty effects. Activity was significantly greater for the first stimulus presentation in both right and left Anterior ROIs ($t(8) = 2.4, p < .05$ and $t(8) = 3.8, p < .01$, respectively), in the left Mid-Anterior ROI ($t(8) = 3.8, p < .01$), and in both the right and left Mid-Posterior ROIs ($t(8) = 4.8, p < .005$ and $t(8) = 3.8, p < .01$, respectively). The effect in the Mid-Anterior ROI did not reach significance, $t (8) = 1.6, p = .14$. There was no significant effect of novelty in both Posterior ROIs (left Posterior $t(8) = 1.2, p > .1$). To further explore the nature of the novelty effect, we calculated beta values for each stimulus repetition within each of the ROIs. As shown in Fig. 6, the Anterior and Mid-Anterior ROIs showed strong above baseline activation for the first one or two presentations of each stimulus, followed by a sharp drop to baseline or below activity across all subsequent presentations. In contrast, the Posterior ROI showed no clear novelty effect, but instead a complex pattern of activity across stimulus repetitions. Notably, there was higher activity for odd numbered stimulus repetitions in contrast with even repetitions (Presentations 1, 3, 5, and 7>Presentations 2, 4, 6, and 8; left Posterior $t (8) = 2.7, p < .05$; right Posterior $t (8) = 2.87, p < .05$). The odd repetitions correspond to the first time each stimulus appeared in each half block (each stimulus appeared twice in each standard categorization and judgment half-block; repetitions 1 and 2, and 5 and 6 were always in standard categorization, and 3, 4, 7, and 8 in judgment). The Mid-Posterior ROI appears to reflect a combination of the two patterns of activity: a novelty effect plus stronger activity on odd repetitions. The odd vs. even repetition effect was significant for the right Mid-Posterior ROI, $t (8) = 2.6, p < .05$, but not the left $p > .1$.

Interactions between regions

We used Granger causality mapping to examine interactions between the primary basal ganglia, hippocampal, cortical, and brain stem regions involved in the Categorization and subjective judgment tasks, as described above in Materials and methods. We first examined directed influences between basal ganglia regions. As shown in Fig. 7, we found a pattern from putamen to anterior caudate.
to posterior caudate. The putamen seed region received directed influence only from other regions of the putamen, and exerted directed influence on the entire caudate. The anterior caudate nucleus received directed influence from the putamen and adjacent anterior regions of the caudate, and exerted directed influence on the posterior caudate. The posterior caudate (seed region in the body of the caudate) received directed influence from the putamen and anterior regions of the caudate, and exerted directed influence on more posterior regions of the caudate (the tail of the caudate). This pattern replicates previous research from our laboratory (Lopez-Paniagua and Seger, (in press); Seger et al., 2010).

We then examined how the basal ganglia interacted with the hippocampus and dopaminergic brain stem. As shown in Fig. 8, the posterior hippocampus received directed influence from the putamen and exerted directed influence on the caudate nucleus. There were no reliable patterns of directed interaction to or from the anterior

Fig. 5. Recruitment of anterior (blue) and posterior (red) hippocampal regions across judgment types. Top: Activated regions overlaid on a coronal and sagittal slices. Bottom: Beta parameter plots indicating activity in the left and right activated regions in the Remember, Know-Automatic, and Know-Intuition conditions.
hippocampal ROIs. Finally, the brain stem exerted directed influence on both the basal ganglia and the hippocampus, consistent with the known projections of dopaminergic neurons from this region. The overall pattern of interactions between regions is summarized in Fig. 9.

Discussion

The results reveal complex patterns of interaction between hippocampus and basal ganglia subserving learning, with the basal ganglia associated with two different kinds of “knowing”, and the hippocampus with memory based categorization decisions. Furthermore, they indicate that two different regions of the hippocampus are recruited by different task demands; the anterior hippocampus during processing of novel stimuli, and a more posterior region associated with categorization learning overall, memory based trials in particular.

Categorization and the basal ganglia

Overall we found that the basal ganglia were activated during categorization tasks, consistent with a large number of previous studies. We have extended that work by finding that the basal ganglia are particularly associated with subjects reporting that they responded on the basis of “knowing” the answer. The basal ganglia were further associated with two different subjective Knowing states; a subjectively automatic, rapid, high-confidence state (Know-Automatic) and a subjectively uncertain intuitive state (Know-Intuition).

Table 1b

Direct comparison of Know-Automatic and Know-Intuition.

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>37</td>
<td>−30</td>
<td>59</td>
<td>6323</td>
</tr>
<tr>
<td>60</td>
<td>−30</td>
<td>28</td>
<td>773</td>
</tr>
<tr>
<td>55</td>
<td>−13</td>
<td>−10</td>
<td>1025</td>
</tr>
<tr>
<td>−55</td>
<td>−60</td>
<td>17</td>
<td>1300</td>
</tr>
<tr>
<td>−31</td>
<td>−33</td>
<td>−10</td>
<td>768</td>
</tr>
<tr>
<td>−11</td>
<td>−84</td>
<td>16</td>
<td>458</td>
</tr>
</tbody>
</table>

The role of the basal ganglia in implicit learning

The basal ganglia are often considered to be components within an implicit learning system. This originated from cognitive neuropsychology research that mapped the explicit vs. implicit learning dichotomy from cognitive psychology onto the declarative vs. nondeclarative dichotomy in neuroscience (e.g., Squire and Zola-Morgan, 1988). It is important to note that the original distinction between explicit and implicit learning in cognitive psychology was based on differences in conscious access what was learned: explicit learning was defined as resulting in fully conscious and usually verbalizable knowledge, whereas implicit learning was defined as resulting in unconscious knowledge that was not fully verbalizable (Seger, 1994). In contrast, the distinction between declarative and nondeclarative learning was based on neural systems, with declarative knowledge reliant on the medial temporal lobe memory system, and nondeclarative knowledge reliant on a collection of other neural systems (Squire and Zola-Morgan, 1988). Despite these qualitatively different definitions, the mapping of explicit–implicit onto declarative–nondeclarative held up well for a number of years. Across a broad range of tasks, including perceptual priming (Cermak et al., 1985), eye blink conditioning (Clark and Squire, 1998), motor skill learning (Nissen and Bullemer, 1987) and categorization (Knowlton and Squire, 1993), it was found that variants that required explicit, or conscious, memory in healthy adults were also impaired in persons with amnesia due to damage to the hippocampus. In contrast, tasks that healthy adults could perform implicitly (without full conscious knowledge) were preserved in amnesia. Some implicit learning tasks were shown to be dependent on the basal ganglia, including motor sequence learning (Siegert et al., 2006) and probabilistic classification learning (Shohamy et al., 2008). However, there are limitations to equating the two definitions of implicit learning. There have been a number of recent studies that have found that learning within nondeclarative systems can sometimes be accessible to consciousness, whereas learning within the declarative memory system is sometimes not verbalizable (Schendan et al., 2003).

Our results provide additional evidence for the existence of nondeclarative learning supported by the basal ganglia that is explicit, in the sense of being accessible to consciousness. We found that the basal ganglia were associated with Know responses, across both relatively explicit (Know-Automatic) and implicit (Know-Intuition) trials. This contributes to the growing body of evidence that the basal ganglia are not exclusively associated with implicit or unconscious forms of learning. The basal ganglia are active when performing tasks in which subjects are aware of what they have learned, such as rule learning via hypothesis testing (Seger and Cincotta, 2006; Monchi et al., 2001). Basal ganglia activity is also found in both implicit and explicit sequence learning (Keele et al., 2003; Destrebecqz et al., 2005), and persons with basal ganglia disorders are impaired on explicit as well as implicit sequence learning (Wilkinson et al., 2009).

The basal ganglia and familiarity

We found basal ganglia recruitment for both kinds of Know judgments. In recognition memory studies, Know judgments are thought to reflect familiarity based memory processing, in contrast with full recollection expressed in Remember judgments. This raises the question of whether the basal ganglia contribute to familiarity-based memory. It is important to note that recognition tests differ in several ways from the categorization learning tasks we used here. In recognition, a subject views the stimulus, and makes a yes/no, or old/new, memory decision. In categorization, recognition of the stimulus may well occur but needs to be combined with retrieval and execution of the categorical response. Therefore, the Know judgments in the current study may reflect some combination of the familiarity of the stimulus and/or other factors, possibly including the fluency of the decision making process and motor response.
Despite these differences in task demands, there is some research indicating a role of the basal ganglia in familiarity based recognition memory. Several studies of patients with basal ganglia impairment due to Parkinson’s disease have found that these patients are impaired on familiarity across a variety of tasks, including the Remember-Know task (Davidson et al., 2006; Weiermann et al., 2010) but see also (Edelstyn et al., 2009). Functional imaging studies of recognition memory have typically focused on medial temporal lobe recruitment and have deemphasized subcortical activations. However, Montaldi et al. (2006) reported an association between increased caudate activity and reported levels of familiarity.

Categorization, memory, and novelty processing in the hippocampus

We found a dissociation between the anterior and posterior regions of the hippocampus during category learning. The anterior region was overall reduced in activity during categorization, particularly during Remember trials, and was primarily affected by stimulus novelty. This region corresponds well to the region reported as being deactivated in previous category learning studies (Poldrack et al., 1999, 2001; Seger and Cincotta, 2006).

We also found a posterior hippocampal region that was active overall during categorization, and was activated during Remember trials. The location of this region matches that found for Remember trials in many recognition memory studies using pictorial information (Uncapher and Rugg, 2005; Fenker et al., 2005; Slotnick, 2009; Montaldi et al., 2006). Our results broaden the conditions in which Remember judgments are associated with posterior hippocampal recruitment from relatively simple recognition tasks, to more complex uses of memory required in the categorization tasks. Our finding that subjects can use memory based strategies to succeed at learning categorization tasks is consistent with the finding that Parkinson’s patients recruit medial temporal lobe regions when learning to categorize (Moody et al., 2004), and have implications

**Fig. 6.** Novelty related responses along the anterior–posterior axis of the hippocampus, displayed from top to bottom. Left: Graphs of beta parameters indicating recruitment of left and right ROIs across stimulus presentations 1 through 8. Right: Anatomically defined hippocampal ROIs, overlaid on coronal slices. Each ROI was an $8 \times 8 \times 8$ voxel cube. The Talairach coordinates of the center voxel of each ROI are indicated beneath each image.
for understanding the development of compensatory strategies in memory or learning disorders.

In the memory literature, two theories have been recently proposed characterizing the different functions of anterior and posterior hippocampus. One theory associates anterior hippocampal activity with novelty processing and more posterior regions with recollection (Daselaar et al., 2006; Strange et al., 1999). This theory is clearly consistent with our reported results finding that the anterior hippocampus was sensitive to stimulus novelty and the posterior to memory based categorization. An alternative view is that the anterior hippocampus is associated with relational encoding (Chua et al., 2007) and flexible use of relational information, whereas posterior regions are more concerned with reinstatement of the context within which the stimulus was originally encountered (Giovanello et al., 2009). This theory is also compatible with our results. When encountering a novel stimulus in a categorization task subjects are required to not only encode the stimulus itself but in addition to set up a relational memory between the stimulus and the category it belongs to. During memory based categorization of repeated items, subjects need to recall the situation in which the stimulus was previously encountered in order to retrieve its category membership from memory.

Interactions between basal ganglia, hippocampus, and dopaminergic midbrain

We found a clear interaction between dorsal striatal regions, in which the putamen exerts influence on the anterior caudate, which in turn exerts influence on the posterior caudate. This is the third study from our lab to identify this pattern, indicating that it holds across different types of stimuli (abstract fractals, faces and houses) and different stimulus–response contingencies (all deterministic in the present work, deterministic and random in Seger et al., 2010, and probabilistic and random in Lopez-Paniagua and Seger). This pattern largely follows the order in which we expect the different cortico-striatal loops to be of primary importance within categorization trials: stimulus categorization and motor response selection occur first, which should require the visual and motor loops (including the putamen), then are followed by feedback processing, which requires the executive loop through the anterior caudate nucleus. One puzzling pattern in our data is that the anterior caudate exerts influence on the posterior caudate, whereas the reverse should be true because visual processing recruiting the posterior caudate precedes feedback processing recruiting the anterior caudate. This may be due to the location of our posterior caudate seed region in a relatively anterior location within in the body of the caudate that may be strongly influenced by the executive loop; in previous studies the body of the caudate has exhibited characteristics of both the head and tail regions, consistent with its intermediate anatomical position (Lopez-Paniagua and Seger, [in press]; Seger et al., 2010). The patterns of influence we found are not in accordance with the primary anatomical interactions within the striatum (Haber, 2003), which proceed along a gradient from the most ventral, anterior, and medial regions (primarily ventral striatum) to the most dorsal, posterior, and lateral regions (primarily posterior putamen and tail of the caudate nucleus). However, there are a number of additional anatomical connections between
corticostriatal loops that do not follow this gradient that may underlie these patterns of influence (see Lopez-Paniagua and Seger, (in press), for a more detailed discussion of these issues).

The present study elicited robust hippocampal activity in two regions (anterior and posterior), which allowed us to examine how the medial temporal lobe interacts with the basal ganglia. The posterior hippocampus received directed influence from the putamen, and exerted directed influence on the anterior and posterior caudate. This indicates it may be playing an intermediate role in categorization: processing information about the stimulus and information about the response from putamen, and sending the resulting information to the anterior caudate to be integrated there with feedback. The directed influences from putamen to hippocampus may serve as input to the relational processing functions of the hippocampus. Relational processing may be of use in categorization in two ways: first, during encoding to bind together features e.g., the stimulus and the response options during categorization), and second at retrieval to reconstruct full memories on the basis of features (e.g., using the stimulus to recall the associated response). Finally, directed influence from the hippocampus to the caudate is consistent with integration of retrieved memories with the feedback received on each trial. It is not possible from the current data, however, to determine whether the participation of the hippocampus in this network contributes to categorization learning; it is possible that the hippocampus subserves conscious recollection of the stimulus that is independent of effective categorization.

As described in the Results section, there was no consistent pattern of directed influence from the anterior hippocampal seed regions. This is consistent with our finding that this region was not clearly associated with categorization, and appeared to be primarily sensitive to stimulus novelty. It is important to note that directed influences identified by GCM do not necessarily reflect anatomically direct connections; they may pass through or be mediated by additional regions. Previous research (Poldrack and Rodriguez, 2004; Graham et al., 2009) has found that the interactions between hippocampus and basal ganglia are mediated by the prefrontal cortex.

In addition, we found directed influence from the midbrain to both the hippocampus and the striatum, consistent with the known projections of dopaminergic neurons from the ventral tegmental area and substantia nigra. Dopamine neurons code for reward (firing rates are greatest when an unexpected reward is encountered), and provide a teaching signal to enable learning about reward contingencies. In both basal ganglia and hippocampus, dopamine affects learning via NMDA...
receptors. In the basal ganglia, synaptic plasticity requires the presence of dopamine (Reynolds and Wickens, 2002), whereas in the hippocampus encoding occurs in the absence of dopamine, but dopamine increases the persistence of memories (Duzel et al., 2010; Bethus et al., 2010). Dopamine projections from the midbrain may also play a role in the anterior hippocampal sensitivity to novelty. In addition to reward coding, dopamine neurons increase their firing rates for novel stimuli (Kakade and Dayan, 2002). Electrophysiological studies (Axmacher et al., 2010) of the hippocampus have found both early (187 ms) and later (482 ms) novelty effects; the former may contribute to initial identification of a stimulus as novel and serve as one of the inputs affecting dopamine neurons, and the latter may be a response to the dopamine signal and underlie stronger memory encoding for novel stimuli (Lisman and Grace, 2005).

Conclusions

We found that categorization learning recruits multiple learning systems, including the basal ganglia and hippocampus, and that these systems are associated with different subjective states. The basal ganglia are associated both with high confidence, automatic performance and with lower confidence intuitive performance. The hippocampus can play multiple roles in categorization. When a memory strategy is useful, posterior regions can be recruited, and are associated with a subjective sense of remembering. In addition, the anterior hippocampus can play a role in processing novel stimuli, which may be particularly important when there are many stimuli and/or the stimuli are complex. Subjective measures may be of use in future research to identify when different memory systems are being employed within complex tasks.

Acknowledgments

We would like to thank David McCabe for his advice about the design of the subjective judgment task, and Brian Spiering for useful comments on previous drafts of the manuscript. This research was supported by the Executive Functions program of the National Institute for Mental Health, R01 MH79182.

Appendix

Full descriptions of each judgment type given to subjects in the Instructions. Know-Automatic, referred to in the instructions as a Type A response:

“Is made when you have a feeling or belief that you know the answer but you don’t actually remember it. For example, you may know or have a feeling or “gut instinct” that the picture predicts rain. You may have an impulse to press one of the response keys. You may feel that you answer was right or wrong though you don’t know why you feel that way. You should choose this response if you have some sense about what the correct answer is, but it isn’t an automatic, quick response (in this case you would choose ‘A’), or based on your conscious memory of a previous trial (in this case you would answer ‘B’). However, if you have no sense of what that response should be or whether you are right or wrong, then you should choose type D.”

Guess, referred to in the instructions as a Type D response:

“Is made when you have no idea about what the correct answer is; you are guessing.”

Recollection, referred to in the instructions as a Type B response:

“Is made when you can consciously recollect when and how you learned the information. For example, when the picture appears you may remember it predicted rain (or sun) because you remember seeing it, the response you made, and/or the feedback you received. You may also recall other aspects of the last time you saw the picture—perhaps a feature in the picture reminded you of rain (or sun). You should choose this response when you remember things associated with seeing the picture on a previous trial and using that memory to make your response. Sometimes memories are accompanied by emotion, such as remembering how you feel if you answered incorrectly on a previous trial. Sometimes memories are accompanied by details of the experience. You might remember that part of the picture made you think about something outside of the experiment, or remember hearing a noise from the hallway, or you remember feeling a cold draft.”

Know-Intuition, referred to in the instructions as a Type C response:

“Is made when you have a feeling or belief that you know the answer but you don’t actually remember it. For example, you may know or have a feeling or “gut instinct” that the picture predicts rain. You may have an impulse to press one of the response keys. You may feel that you answer was right or wrong though you don’t know why you feel that way. You should choose this response if you have some sense about what the correct answer is, but it isn’t an automatic, quick response (in this case you would choose ‘A’), or based on your conscious memory of a previous trial (in this case you would answer ‘B’). However, if you have no sense of what that response should be or whether you are right or wrong, then you should choose type D.”

References
