

# Multiple stages of learning in perceptual categorization: evidence and neurocomputational theory

George Cantwell<sup>1</sup> · Matthew J. Crossley<sup>2</sup> · F. Gregory Ashby<sup>3</sup>

Published online: 28 April 2015  
© Psychonomic Society, Inc. 2015

**Abstract** Virtually all current theories of category learning assume that humans learn new categories by gradually forming associations directly between stimuli and responses. In information-integration category-learning tasks, this purported process is thought to depend on procedural learning implemented via dopamine-dependent cortical-striatal synaptic plasticity. This article proposes a new, neurobiologically detailed model of procedural category learning that, unlike previous models, does not assume associations are made directly from stimulus to response. Rather, the traditional stimulus-response (S-R) models are replaced with a two-stage learning process. Multiple streams of evidence (behavioral, as well as anatomical and fMRI) are used as inspiration for the new model, which synthesizes evidence of multiple distinct cortical-striatal loops into a neurocomputational theory. An experiment is reported to test a priori predictions of the new model that: (1) recovery from a full reversal should be easier than learning new categories equated for difficulty, and (2) reversal learning in procedural tasks is mediated within the striatum via dopamine-dependent synaptic plasticity. The results confirm the predictions of the new two-stage model and are incompatible with existing S-R models.

**Keywords** Categorization · Procedural learning · Striatum

## Introduction

Categorization is the process of assigning unique responses to different groups of stimuli. A variety of different category-learning theories have been proposed, yet virtually all assume that category learning is a process via which a single association is formed between each stimulus and every possible response (henceforth referred to as S-R models). However, recent empirical investigations have challenged this seemingly simple assumption, suggesting that a fundamental revision of current theories may be in order (Maddox et al., 2010; Kruschke, 1996; Wills et al., 2006). This article proposes a new biologically detailed model of procedural learning that successfully addresses these challenges. New behavioral data are presented to further support the general necessity of model revision and to justify the specific details of the proposed new model.

A strong prediction of any theory that assumes category learning is mediated by S-R associations is that reversing the correct responses for all stimuli should cause catastrophic interference because recovery from a full reversal would require unlearning all prior S-R associations, followed by new learning of the reversed associations. In contrast, creating new categories from the same stimuli in any other way should be less disruptive, because only some of the associations would have to be relearned, but not all. Existing empirical data, however, indicate that reversal learning is easier than learning novel categories (Kruschke, 1996; Maddox et al., 2010; Sanders, 1971; Wills et al., 2006). These results challenge the validity of S-R learning assumptions that underlie the existing category-learning theories.

---

✉ F. Gregory Ashby  
ashby@psych.ucsb.edu

<sup>1</sup> University of California, Santa Barbara, CA, USA

<sup>2</sup> University of California, Berkeley, CA, USA

<sup>3</sup> Department of Psychological & Brain Sciences,  
University of California, Santa Barbara, CA 93106, USA

Although virtually all existing category-learning models are of the S-R type, a number assume that the S-R learning process is mediated via the interaction of multiple systems (Ashby et al., 1998; Erickson & Kruschke, 1998), or that S-R associations are formed via context-dependent Bayesian inference (Anderson, 1991; Gershman et al., 2010; Redish et al., 2007). Thus, the basic intuition about full reversals in S-R models may not be directly applicable to these models.

The multiple memory systems framework assumes that S-R associations are formed via procedural learning, whereas declarative memory is used to formulate and test explicit strategies (Ashby & O'Brien, 2005). Such a framework may, in principle, be able to account for the observation that recovery from full reversal is easier than new learning. However, since these models assume that procedural category learning is an S-R process, the only way this seems possible is if the reversal is dominated by declarative mechanisms. If declarative mechanisms can somehow be ruled out, then these models are clearly lacking in their current form. Likewise, a context-sensitive model might explain the results of reversal experiments as the effects of context shifts. The exact details of how this would be done are unclear, but even with the addition of context effects, if the underlying learning was S-R then the observed ease of reversals relative to new categories still appears mysterious.

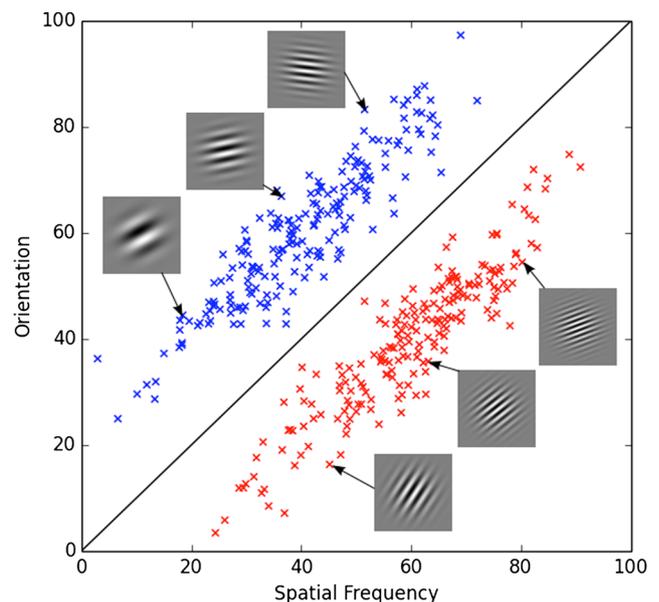
The literature on reversal learning is enormous, dating back at least to Spence (1940). Much of this work focuses on the learning of reversed reward associations. Many animal and neuroimaging studies of such reversals implicate a distributed neural network that includes orbitofrontal cortex and the ventral striatum (e.g., Clarke et al., 2008; Cools et al., 2002; McAlonan & Brown, 2003). Of course, if feedback is given, and the category assignments of all stimuli are reversed, then we would expect activity in a similar reward-learning network. Our interest, however, is not in the learning of the reversed reward associations, but in the learning of reversed motor responses. This literature is considerably smaller, and even within this reduced literature, our focus is on the learning of reversed motor associations in tasks mediated by procedural learning.

The empirical demonstrations that it is indeed easier for participants to recover from a full reversal than it is to learn novel categories (Maddox et al., 2010; Kruschke, 1996; Wills et al., 2006) provide a hint that category learning might include two separate stages or processes—one in which the category structure is learned, and one in which each structure is associated with a motor goal (e.g., press the button on the left). In such two-stage models, a full reversal only requires unlearning and relearning of the response associations (the second stage), since the category structures remain unchanged. In contrast, a switch to novel categories

constructed from the same stimuli could require unlearning and re-learning at both stages. Thus, it seems natural for a two-stage model to predict less interference from a full reversal than from a switch to new categories, in agreement with the empirical findings cited above.

Within the categorization domain, the best evidence for procedural learning comes from the information-integration (II) task. In an II category-learning task, stimuli are assigned to categories in such a way that accuracy is maximized only if information from two or more noncommensurable stimulus dimensions is integrated at some predecisional stage (Ashby & Gott, 1988). Typically, the optimal strategy in II tasks is difficult or impossible to describe verbally (which makes it difficult to discover via logical reasoning). An example of categories that might be used in an II task is shown in Fig. 1. In this case, the two categories are each composed of Gabor patterns, which are circular sine-wave gratings that vary in bar width (i.e., spatial frequency) and bar orientation. The diagonal line denotes the category boundary. Note that no simple verbal rule correctly separates the disks into the two categories. Nevertheless, many studies have shown that with enough practice people reliably learn such categories (Ashby & Maddox, 2005).

II categorization tasks are often contrasted with rule-based (RB) tasks, in which the categories can be learned via some explicit reasoning process that requires selective attention. Many studies have documented a wide variety of qualitative differences in how RB and II tasks are initially learned (see Ashby & Maddox, 2005, 2011 for reviews), and all of these are consistent with the hypothesis that



**Fig. 1** An example of stimuli that might be used in an information-integration (II) category-learning task. Gabor patches that differ on spatial frequency and orientation are assigned to either category A or B

procedural memory is required to learn II structures, whereas RB category learning is mediated by declarative memory (Ashby & O'Brien, 2005; although see Newell et al., 2011).

This article proposes a new biologically detailed two-stage model of procedural category learning. The model generalizes the best existing neurocomputational model of II learning and is motivated by multiple findings in the literature. We then derive several a priori predictions from the model about a variety of different reversal conditions. Next, we describe the results of a new experiment that tests and supports these predictions. Specifically, our results show that: (1) learning new categories is more difficult than recovering from a full reversal, and (2) recovering from a reversal depends on procedural, rather than declarative mechanisms.

## A new biologically detailed model of procedural category learning

### Motivation

There is considerable evidence that II category learning depends critically on the striatum (e.g., for a review, see Ashby & Ennis, 2006). The only current theory of category learning that posits a neurobiological locus of II category learning is COVIS (Ashby and Waldron, 1999; Ashby et al., 1998). COVIS assumes that direct associations from stimuli to motor response are learned during II categorization. Stimuli are associated with categorization responses via changes in synaptic strength at cortical-striatal synapses between pyramidal neurons in visual association areas and medium spiny neurons in the striatum. Although COVIS accounts for an impressive variety of data (e.g., Ashby et al., 2007; Crossley et al., 2013, 2014; Filoteo et al., 2014; Hélie et al., 2012; Valentin et al., 2014), because it postulates only one stage of learning, it predicts either worse transfer performance in II tasks after a full reversal than after a switch to new categories, or a switch to suboptimal declarative strategies during reversals.

The most relevant study to the question of whether procedural learning during II categorization is mediated by one or two stages was reported as experiment 3 by Maddox et al. (2010). This study found evidence that after training on one set of II categories, recovery was faster when the training categories were reversed than when participants were transferred to novel categories constructed from the same stimuli. Below, a similar experiment is reported that confirms this conclusion.

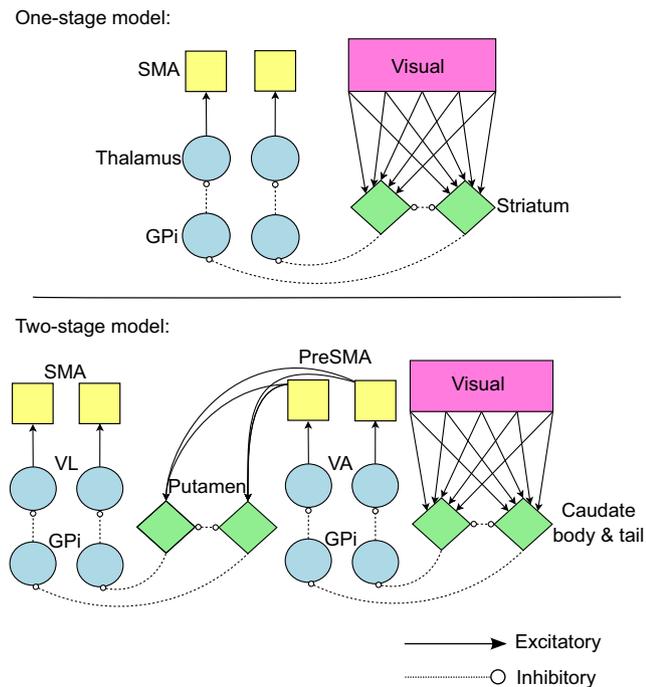
So how should the procedural-learning model of COVIS be generalized to include a second learning stage? One clue comes from fMRI studies of II category learning, which

have not reported a consistent site of task-related activation within the striatum. Some studies have reported activation in both the caudate nucleus and the putamen (Cincotta & Seger, 2007; Seger & Cincotta, 2002), while other studies have reported task-related activity either only in the putamen (Waldschmidt & Ashby, 2011) or only in the caudate (Nomura et al., 2007). Thus, one interpretation of the literature is that both the caudate nucleus and the putamen are, or can be, relevant for II category learning.

Other clues come from anatomical and physiological studies that support the existence of (at least partially) segregated loops through the basal ganglia that each serves a different function. For example, DeLong et al. (1983) distinguished between an associative loop through the caudate nucleus and a motor loop through the putamen. More recent authors suggest four cortical-striatal loops: motor, visual, executive/spatial and motivational/affective (Seger, 2008; Lawrence et al., 1998). The classic view is that these loops are parallel and closed—that is, the striatum projects back to the same region of cortex from which it receives its input (Alexander et al., 1986; Middleton & Strick, 2000). Even so, there is now considerable evidence that at least some loops through the basal ganglia are open and interconnected (Lopez-Paniagua & Seger, 2011; Joel & Weiner, 1994; Hikosaka et al., 2006; Nakano et al., 1992; McFarland & Haber, 2000).

COVIS can be generalized to include two learning stages in a way that is consistent with these results if we hypothesize that II category learning is mediated by double cortical-striatal-cortical loops. The current literature is insufficient to make any strong claims about the exact anatomical locus of these loops, but an initial plausible hypothesis<sup>1</sup> is illustrated in Fig. 2. The model begins with projections from cortical visual association areas into the body and tail of the caudate nucleus, which projects through the globus pallidus and ventral anterior nucleus of the thalamus to the pre-supplementary motor area (preSMA). Next, preSMA projects to the posterior putamen then through the globus pallidus and ventral lateral nucleus of the thalamus to the SMA. The two sites of learning in this model are at the two different types of cortical-striatal synapses. The learning of category structure is mediated by plasticity at visual cortical-caudate synapses, and learning to associate each structure with a specific response is mediated at preSMA-putamen synapses. Although we are reticent to make strong,

<sup>1</sup>This model only includes the so-called direct pathway through the basal ganglia, and not the indirect or hyperdirect pathways. As mentioned earlier, the COVIS model of procedural learning, which also only includes a direct pathway, is nevertheless consistent with a wide variety of data (Ashby et al., 2007; Crossley et al., 2013, 2014; Filoteo et al., 2014; Hélie et al., 2012; Valentin et al., 2014). Following the computational cognitive neuroscience simplicity heuristic (Ashby and Hélie, 2011), we do not include structure unless it is critical for function or is required by existing data.



**Fig. 2** Architectures for both models simulated in this article. *Diamonds* represent Izhikevich medium spiny neuron units, *circles* represent quadratic integrate-and-fire units, *squares* represent Izhikevich regular spiking neuron units. (*SMA* = supplementary motor area, *GPI* = internal segment of the globus pallidus, *VL* = ventral-lateral nucleus of the thalamus, *VA* = ventral-anterior nucleus of the thalamus)

precise anatomical claims due to a current lack of evidence, this model nevertheless posits that both learning stages are at cortical-striatal synapses and mediated by dopamine-gated reinforcement learning. As such, the model makes the strong prediction that recovery from an II reversal should be impaired in the presence of delayed feedback. This prediction is discussed, tested, and supported in the experiment described below.

Despite its intuitive appeal, however, it is not immediately clear that this model is capable of learning using a global reinforcement learning rule of the kind thought to characterize synaptic plasticity at cortical-striatal synapses (Ashby & Helie, 2011). This is because of a credit-assignment problem. For example, if this model produces an incorrect response and negative feedback is delivered, then dopamine levels will drop below baseline and all recently active synapses will be weakened. However, the error could have occurred because one stage was incorrect while the other was correct. In this case, the synapses at the stage that was correct would be weakened, reducing the probability of a correct action on future trials. Therefore, it is critical to test whether this model can learn when combined with a biologically realistic model of reinforcement learning, and whether it can actually provide better accounts of the various data than the single-stage version of COVIS.

If the two-stage model can solve the credit-assignment problem, then it seems likely that the model will recover from a full reversal more quickly than from a partial reversal (i.e., a switch to new categories constructed from the same stimuli). Again, this is in direct opposition to what we expect from a one-stage model. Before comparing these models to human behavior however, it is worth confirming these expectations. In the following section, we test whether the two-stage model can solve the credit assignment problem, and we describe simulations of both the one-stage and two-stage models. It should be noted that these simulations effectively present a priori predictions of the models since no consideration was given to empirical data in order to produce them. This is possible, in part, because unlike many mathematical models, neurobiologically constrained models are severely limited in their range of behavior.

## Simulations

We constructed biologically detailed computational cognitive neuroscience (CCN) models of both networks shown in Fig. 2. CCN models are constructed to mimic neural networks that are faithful to known neuroanatomy and that include spiking neuron units in each of the modeled brain regions (Ashby & Helie, 2011). The models use learning rules that mimic biologically plausible synaptic plasticity. Typically, models of each neuron type are fixed by fitting the individual unit models to appropriate single-unit recording data from the literature. In real brains, connections between brain regions do not appear or disappear from task to task, nor does the qualitative nature via which a neuron responds to input. For these reasons, the architecture of the network and the models of each individual unit (including numerical values of all parameters) remain fixed throughout all applications (called the Set-in-Stone Ideal; Ashby & Helie, 2011).

Compared to traditional cognitive models, CCN models have a number of attractive advantages (for more details see Ashby & Helie, 2011). First, they are more constrained than cognitive models because of their fixed architecture and the rigid way that each unit responds to input. For example, the models shown in Fig. 2 include a parameter specifying the strength of synapses between thalamus and SMA. Neuroanatomy requires that these are excitatory projections, so no matter the value of these parameters, increasing activity in thalamus can only increase activity in SMA<sup>2</sup>. Second, attending to the neuroscience data can expose relationships between seemingly unrelated behaviors. For example, the architectures shown in Fig. 2 as

<sup>2</sup>Of course, if the synaptic weights are set too low, then no amount of thalamic activity will cause a response in SMA. But in this case, the model never makes a motor response, so these values are avoided.

possible models of II category learning are similar to the neural networks that have been proposed to underlie implicit sequence learning (e.g., Grafton, Hazeltine, & Ivry, 1995), suggesting these two seemingly disparate behaviors might share some previously unknown deep functional similarity. Third, in many cases, studying the underlying neuroscience leads to surprising and dramatic behavioral predictions that would be difficult or impossible to derive from a purely cognitive approach. For example, the prediction that II category learning should be impaired with a feedback delay came from a CCN approach. Fourth, CCN models are especially amenable to a converging operations approach to model testing because they make predictions about both behavioral and neuroscience data. Thus, rather than simply testing them against behavioral data, it should also be possible to test CCN models against a variety of neuroscience data, including single-unit recording data, lesion data, psychopharmacological data, fMRI data, and possibly even EEG data.

The two CCN models we constructed included all brain areas shown in Fig. 2. The one-stage model is identical to the procedural-learning system of COVIS (Ashby et al., 1998, 2007; Ashby & Waldron, 1999; Crossley et al., 2013). The two-stage model was constructed in an identical manner, but with the addition of extra brain regions. Each model was constructed to mimic the architectures shown in the figure and included spiking neuron units in each of the depicted brain regions. The models used well-established methods (Ashby & Helie, 2011) and similar parameter values to previously published models (Ashby & Crossley, 2011; Crossley et al., 2013). In both models, learning occurred only at cortical-striatal synapses and was mediated via a biologically plausible reinforcement-learning algorithm (Ashby & Helie, 2011). Therefore, in the one-stage model, learning occurred only at synapses between sensory cortical units and the medium spiny units in the striatum. In the two-stage model, learning occurred at these same synapses but also at synapses between preSMA units and medium spiny units in the putamen. Two units are used for each cortical-striatal stage—one for each possible category or response. Mathematical details are given in the [Appendix](#).

All parameters in the models were fixed, except the learning rates. Specifically, within each striatal compartment, one learning rate governed the rate of synaptic strengthening (long-term potentiation or LTP) and one governed the rate of synaptic weakening (long-term depression or LTD). Thus, in the one-stage model, two learning rates were estimated (the rates of LTP and LTD at synapses between visual cortex and the caudate), whereas four learning rates were estimated in the two-stage model (one pair for each stage of learning).

To extract the a priori predictions of the models, we ran them through full Reversal and New Category conditions

(see Fig. 4 for category distributions, the exact details of which can be found in the Methods of the experiment reported below). Learning rates were optimized to extract the best possible performance of the models across both conditions (i.e., to maximize accuracy). Of course, it is implausible that the rates of LTP and LTD in the human striatum evolved specifically to optimize accuracy in the II tasks described in this article. If not, then the parameter values that maximize model accuracy are not good estimates of human striatal learning rates. Even so, examining performance of the models under conditions in which they learn as quickly as possible demonstrates the natural behavior of the model architectures, showcasing the best performance one can hope for with each model.

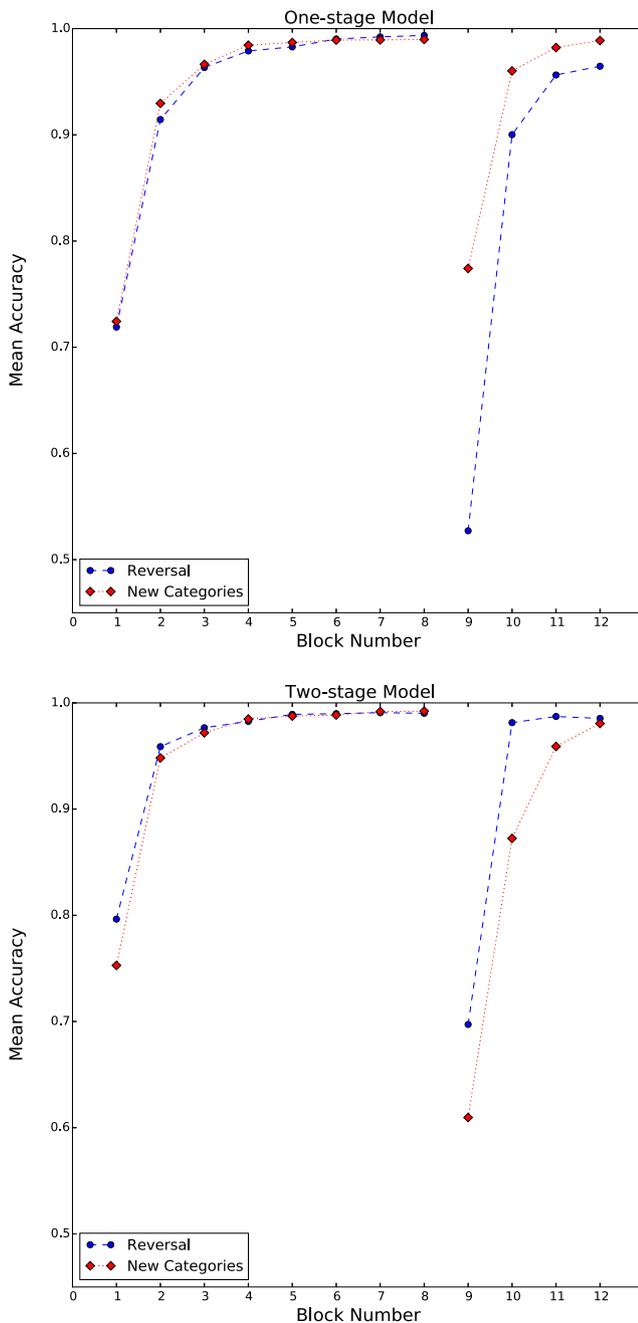
To find the best-fitting values of the learning rates, we used particle swarm optimization (e.g., Clerc, 2012), which creates a population of potential solutions (the “particles”) and then iteratively moves these particles in parameter space according to both their historically best position, and the best-known position of their neighborhood. Due to the stochastic nature of the models, the “function” to be optimized (proportion of correct responses) is not strictly a function at all. Hence, particle swarm optimization, which makes very few assumptions about the form of the problem, is an appropriate tool where traditional optimization routines will fail. After parameter estimation was complete, 100 simulations were run with the best-fitting parameter values and the model predictions were computed by taking the mean across all 100 simulations.

The results are presented in Fig. 3. First note that, as expected, the one-stage model performs worse in the Reversal condition than in the New Categories condition. This seems an inescapable consequence of the fact that all stimuli change their category membership in the Reversal condition, whereas only half change their membership in the New Categories condition. Any one-stage model must therefore reverse all associations in the Reversal condition, but only half of the associations in the New Categories condition.

Second, note that the two-stage model learns consistently, both during training and during transfer. Thus, this model successfully solves the credit assignment problem.

Third, note that the two-stage model predicts better transfer performance in the Reversal condition than in the New Categories condition. It does this when it maximizes accuracy across both tasks with no regard for human data. In this sense, it is an a priori prediction of the model.

Fourth, note that initial learning is actually faster for the two-stage model than for the one-stage model when both models are optimized for maximum performance. This is because the one-stage model maximizes accuracy when it recovers quickly from the reversal. A quick recovery requires a high LTD rate (an order of magnitude larger than the rate of LTP). This high rate of LTD hastens recovery



**Fig. 3** The most accurate performance allowed by each model

from the reversal, but at the cost of slower initial learning. Thus, despite having more synapses to train (two stages of learning versus one), the two-stage model still learns faster. Further simulations showed that learning rates can be found that allow the one-stage model to learn initially as fast (or faster) than the two-stage model, but performance after the reversal takes a serious hit with these parameter values. In particular, transfer accuracy of the model is below chance for a considerable time. Thus, the one-stage model can either do well initially or during the reversal, but not both.

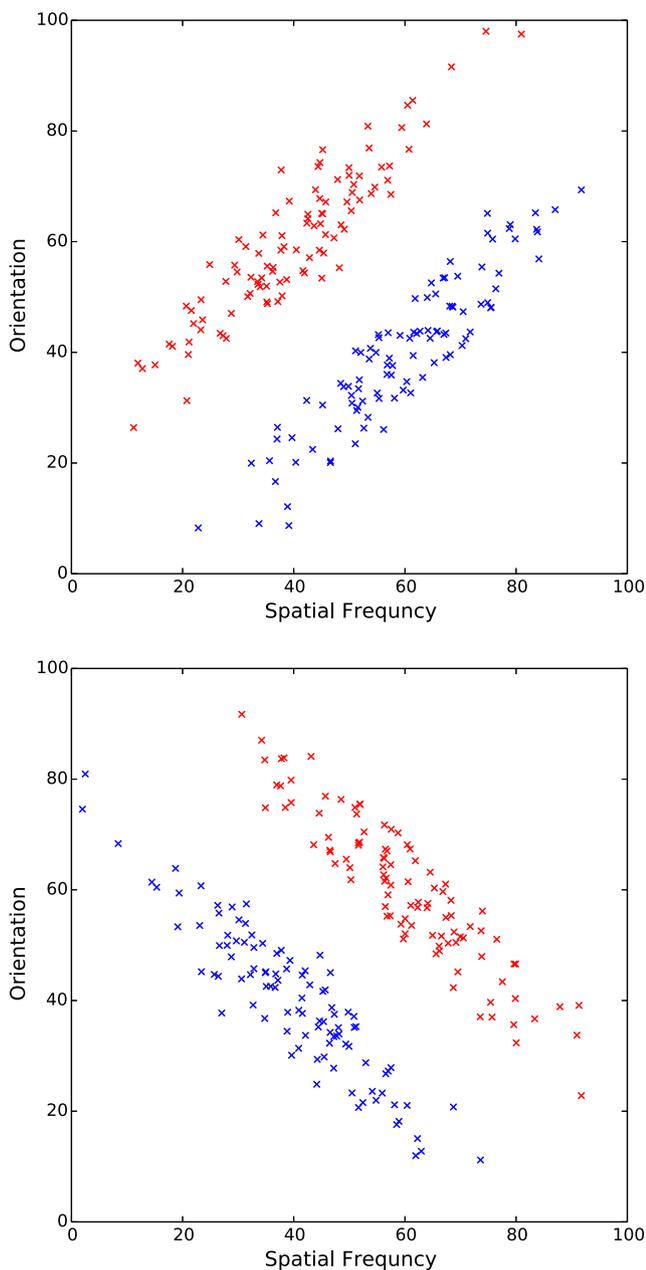
The natural flexibility of the two-stage model provides an inherent advantage over the one-stage model, even in training, with quite minimal constraints on transfer performance (i.e., at least as good as chance).

The next section describes an experiment to test this new model.

## An experimental test

Participants were trained on the II categories shown in Fig. 1. After they had learned the categories, without any warning, they were transferred to one of three conditions: (1) Reversal: a reversal shift in which the category assignment of every stimulus was reversed; (2) New Categories: new categories created from the same stimuli; or (3) Reversal/Delay: a reversal shift with a feedback delay. The new categories were created by rotating the category bound shown in Fig. 1 by 90 degrees counterclockwise (see Fig. 4). Note that this rotation reverses the category membership of half of the stimuli whereas the other half retain their same category membership as during initial training. Thus, simple S-R models predict that performance should be worse in the Reversal conditions than in the New Categories condition.

A number of previous studies have included similar experimental conditions. For example, numerous studies have reported that reversing the category response keys interferes with the expression of II learning, but not the expression of simple RB learning (Ashby et al., 2003; Maddox et al., 2004; Spiering & Ashby, 2008). None of this work however, focused on recovery and no attempt was made to compare the effects of a reversal to learning new categories. Our Reversal and New Categories conditions are most similar to Experiment 3 of Maddox et al. (2010). However, in our experiment, participants were not informed of any changes to the categories prior to transfer, whereas in Maddox et al. (2010) participants were explicitly informed that the category structures had changed at the time of the reversal or when new categories were introduced. Especially in reversal conditions, informing participants when a change has occurred raises the possibility of explicit intervention. For example, one possibility is that the faster recovery from a full reversal reported by Maddox et al. (2010) was not a natural property of procedural learning, but instead was due to explicit intervention. This possibility is somewhat reduced in our design because we did not explicitly inform participants of a change. Another critical difference is that we included the Reversal/Delay condition to explore whether reversal learning is mediated via procedural or declarative mechanisms. In this condition, feedback was delayed by 2.5 seconds after the response. The delay was introduced in the final block of training to avoid its introduction coinciding with transfer.



**Fig. 4** Scatter plots showing stimuli that were used during transfer. *Top* transfer stimuli for Reversal and Reversal/Delay conditions. *Bottom* transfer stimuli for New Categories condition

This way, the introduction of the delay itself would not alert participants to any changes in category membership.

Motivation for including the Reversal/Delay condition comes from the neuroscience literature on synaptic plasticity within the basal ganglia. Much evidence suggests that procedural learning depends critically on the striatum (Badgaiyan et al., 2007; Grafton et al., 1995; Jackson & Houghton, 1995; Knopman & Nissen, 1991) and in particular, on synaptic plasticity at cortical-striatal synapses

that is facilitated by a dopamine (DA) mediated reinforcement learning signal (Valentin et al., 2014). The idea is that phasic DA levels in the striatum increase following positive feedback, and decrease following negative feedback, causing the strengthening (i.e., via LTP) or weakening (i.e., via LTD) of recently active synapses, respectively. In this way, the DA response to feedback serves as a teaching signal for which successful behaviors increase in probability and unsuccessful behaviors decrease in probability.

According to this account, synaptic plasticity can only occur when the visual trace of the stimulus and the post-synaptic effects of DA overlap in time. The cortical excitation induced by the visual stimulus results in glutamate release into the striatum, which initiates several post-synaptic intracellular cascades that alter the cortical-striatal synapse (e.g., Rudy, 2014). One such cascade, which seems especially important for cortical-striatal synaptic plasticity, is mediated by NMDA receptor activation and results in the phosphorylation of calcium/calmodulin-dependent protein kinase II (CaMKII). During a brief period of time (thought to be several seconds) when CaMKII is partially phosphorylated, a chemical cascade that is initiated when DA binds to D1 receptors can potentiate the LTP-inducing effects of CaMKII (e.g., Lisman et al., 2002).

Thus, the effects of feedback should be greatest when the peak effects of the DA-induced cascade overlap in time with the period when the CaMKII is partially phosphorylated. The further apart in time these two cascades peak, the less effect DA will have on synaptic plasticity. In fact, a number of studies have reported II category learning results consistent with these predictions. First, Worthy et al. (2013) reported that II learning is best with feedback delays of 500 ms and slightly worse with delays of 0 or 1000 ms. Second, several studies have reported that feedback delays of 2.5 s or longer impair II learning, whereas delays as long as 10 s have no effect on RB learning (Dunn et al., 2012; Maddox et al., 2003; Maddox & Ing, 2005).

In summary, if reversal learning is mediated via procedural memory, and in particular by cortical-striatal dynamics, then it should be impaired when feedback is delayed. In contrast, if reversal learning is mediated cortically (e.g., by declarative mechanisms) then it should be unaffected by delayed feedback. Evidence reported below that reversal learning is indeed impaired when feedback is delayed support the proposed new model, which assumes that the reversal is mediated via procedural memory.

## Methods

**Participants** Seventy-nine participants completed the study and received course credit for their participation. All participants had normal or corrected to normal vision. Each participant was assigned randomly to one of three

conditions. Participants were excluded from subsequent analyses if they failed to reach a mean accuracy of 70% during the final two blocks of training. Sixty-three participants met this criterion (Reversal  $N = 19$ ; New Categories  $N = 25$ ; Reversal/Delay  $N = 19$ ). A priori power calculations using G\*Power 3.1.9 (Faul et al., 2007) suggest that with this sample size, power is approximately 0.8 for a moderate effect size ( $f = 0.25$ ) with  $\alpha = 0.05$  and a between-measure correlation of 0.3.

**Stimuli and category structure** Stimuli were sine-wave gratings with Gaussian masks (Gabor patches). They differed across trials only in orientation and spatial frequency. Values for orientation and frequency were assigned by sampling from two bivariate normal distributions—one for each category. These distributions had means  $\mu_1 = (40, 60)$  and  $\mu_2 = (60, 40)$ , and identical covariance matrices,  $\Sigma = \begin{bmatrix} 185 & 170 \\ 170 & 185 \end{bmatrix}$ . The optimal decision bound for these distributions is the line  $y = x$ , which would achieve 100% accuracy.

Two hundred random samples from each distribution were selected for the training categories and 100 random samples were selected for the transfer categories. Outliers (more than three Mahalanobis distance units from the mean) were discarded and replaced with new samples. Each set of sample stimuli was linearly transformed so that the sample means, variances, and covariances exactly matched the population values. In the Reversal and Reversal/Delay conditions correct category membership during transfer was the reverse of during training. For the New Category condition, the stimuli were rotated 90 degrees in stimulus space to form the transfer categories. Figure 4 shows scatter plots for the transfer stimulus distributions that were used.

To convert the sampled values into orientations and spatial frequencies the following transformations were chosen:

- Spatial frequency =  $(x + 7.5)/60$  (cycles per degree)
- Orientation =  $0.9y + 20$  (degrees)

These values were selected so as to give roughly equal salience to both dimensions.

**Procedure** Each participant was randomly assigned to one of three conditions: Reversal, New Categories, Reversal/Delay. The experiment was run on computers using PsychoPy (Peirce, 2007). Before the experiment, participants were told that they would learn to categorize novel stimuli. An initial set of on-screen instructions told participants that two buttons would be used (the ‘f’ and ‘j’ keys), and that to start with they would have to guess, having not seen such stimuli before. For all conditions, the total session consisted of 12 blocks of 50 trials each. Between each block there was a participant-terminated rest period. The

stimuli were shuffled and presented in a random order for each participant.

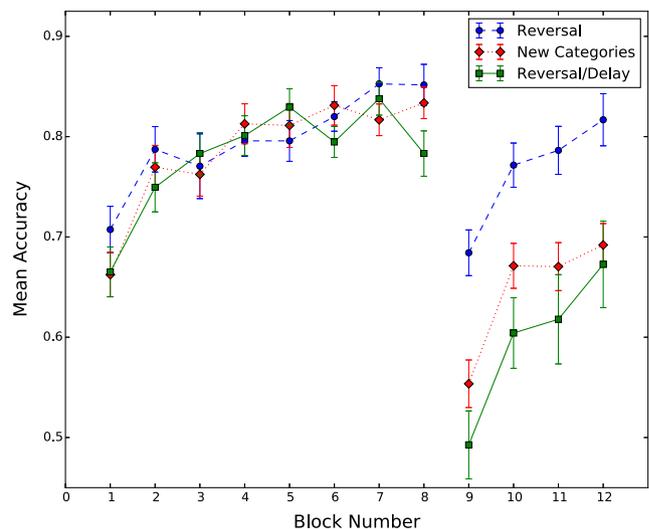
For the Reversal and New Category conditions, there were eight blocks of training followed by four blocks of transfer. No warning or prompt was given about transfer. For the Reversal/Delay condition there were likewise eight blocks of training followed by four blocks of transfer and no warning or prompt was given about transfer. The feedback delay was introduced in the final block of training.

At the start of each trial, a fixation crosshair was presented for 600 ms. Following this, a response-terminated stimulus was presented for a maximum of 3000 ms. Auditory feedback was then given. For a correct response, a happy sound was played: the notes E, G#, and B were played in quick succession with an organ sound. Incorrect responses were met with a negative sound: G then E with a klaxon sound. Auditory feedback was delivered immediately after each response throughout the Reversal and New Categories conditions and during the first seven blocks of the Reversal/Delay condition. For the final five blocks of Reversal/Delay, feedback was given 2.5 s after the response. During the 2.5 s between stimulus and feedback, a random stimulus from the stimulus space was presented.

## Results

**Accuracy analysis** The accuracy results, averaged across participants, are shown in Fig. 5. The data were split between training and transfer for the purpose of statistical analysis.

To examine whether there were any differences in how the three groups learned the identical training categories, a 3 condition  $\times$  8 block repeated-measures ANOVA was computed. The main effect of block was significant



**Fig. 5** Mean accuracy per block, for each condition

[ $F(7, 420) = 26.9, \eta^2 = 0.196, \eta_p^2 = 0.300, p < 0.001$ ], which suggests significant category learning. More importantly, however, there were no differences among conditions [ $F(2, 60) = 0.34, p > 0.5$ ] and no interaction between block and condition [ $F(14, 420) = 1.33, p = 0.19$ ]. Thus, all groups learned the training categories in a similar manner.

To examine differences during transfer, we ran a 3 condition  $\times$  4 block repeated measures ANOVA. The main effects of condition [ $F(2, 60) = 12.46, \eta^2 = 0.181, \eta_p^2 = 0.294, p < .001$ ] and block [ $F(3, 180) = 29.02, \eta^2 = 0.124, \eta_p^2 = 0.323, p < .001$ ] were both significant, and the condition  $\times$  block interaction was not significant [ $F(6, 180) = 0.38, p > 0.5$ ].

We also completed post hoc repeated measures ANOVAs for each pair of conditions during transfer. These analyses revealed that recovery from the reversal with immediate feedback was significantly different than recovery from the new categories [ $F(1, 42) = 21.7, \eta^2 = 0.149, \eta_p^2 = 0.341, p < .001$ ] and from the reversal with delayed feedback [ $F(1, 36) = 18.61, \eta^2 = 0.231, \eta_p^2 = 0.341, p < .001$ ], but recovery from the new categories was not significantly different than recovery from the reversal with delayed feedback [ $F(1, 42) = 2.005, p = 0.16$ ].

**Decision-bound modeling** The accuracy-based analyses demonstrate that recovery from a reversal is faster than transfer to new categories or a reversal with feedback delay. These results are consistent with two plausible hypotheses. One is that the recovery differences are due entirely to properties of procedural learning. Another however, is that the reversal and/or new categories caused participants to switch to non-optimal declarative memory-based strategies. To test between these possibilities, a number of different decision bound models (e.g., Ashby & Gott, 1988; Maddox & Ashby, 1993) were fit to the response data of each individual participant.

Decision-bound models assume that participants partition the perceptual space into response regions. On every trial, the participant determines which region the percept is in, and then emits the associated response. Three different types of models were fit to each participant's responses:

models that assumed an explicit RB strategy, models that assumed an II strategy, and models that assumed random guessing. The RB models assume that participants set a decision criterion on a single stimulus dimension. For example, a participant might base his or her categorization decision on the following rule: "Respond A if the bar width is small, otherwise respond B". Two versions of the model were fit to the data. One version assumed a decision based on spatial frequency (i.e., bar width), and the other assumed a decision based on orientation. These models have two parameters: a decision criterion along the relevant perceptual dimension, and a perceptual noise variance. The II model assumes that participants divide the stimulus space using a linear decision bound. One side of the bound is associated with an A response and the other side is associated with a B response. These decision bounds require linear integration of both stimulus dimensions, thereby producing an II decision strategy. This model has three parameters: the slope and intercept of the linear decision bound, and a perceptual noise variance. Two models assumed the participant guessed randomly on every trial. One version assumed that each response was equally likely to be selected. This model has no free parameters. A second model assumed that the participant guessed response A with probability  $p$  and guessed B with probability  $1 - p$ , where  $p$  was a free parameter. This model is useful for identifying participants who are biased towards pressing one response key.

Model parameters were estimated using the method of maximum likelihood, and the statistic used for model selection was the Bayesian information criterion (BIC; Schwarz, 1978), which is defined as:  $BIC = r \ln N - 2 \ln L$ ; where  $r$  is the number of free parameters,  $N$  is the sample size, and  $L$  is the likelihood of the model given the data. The BIC statistic penalizes models for extra free parameters. To determine the best-fitting model within a group of competing models, the BIC statistic is computed for each model, and the model with the smallest BIC value is the winning model. All models were separately fit to the last 100 training trials of each participant and the last 100 transfer trials.

The results of the model fitting are presented in Table 1. Note that more than 90% of the participants successfully acquired an II strategy during training and that this

**Table 1** Number of participants whose responses were best fit by decision-bound models of three different types

	II strategy	1-D rule	guessing
Reversal training	17	2	0
Reversal transfer	14	5	0
New Categories training	23	2	0
New Categories transfer	11	9	5
Reversal/Delay training	17	2	0
Reversal/Delay transfer	6	12	1

percentage did not differ much across conditions. Thus, in all three conditions, the vast majority of participants appeared to learn the categories using a procedural memory strategy. After transfer in the Reversal condition, explicit one-dimensional rule use increased slightly but II strategies remained in use by the majority of participants. In stark contrast, almost two-thirds of participants in the Reversal/Delay condition used rules after transfer. Fisher's exact test supports the conclusion that the distribution of strategies is significantly different between conditions ( $p = 0.02$ ).

The fact that participants continued using II strategies after transfer in the Reversal condition indicates that performance observed here is due directly to properties of procedural learning. The greater move away from II strategies in the New Categories and Reversal/Delay conditions makes the inferences about procedural learning slightly less direct. Since the transfer was introduced without any prompt or warning, there were no direct cues to participants to change strategy. Rather, a change in strategy would likely be prompted by the failing of the currently employed strategy. That this "failing" occurred after transfer in the New Categories and Reversal/Delay conditions, but not in the Reversal condition, is indicative of poorer performance of procedural learning in these transfer scenarios. It should also be noted that the delay in the Reversal/Delay condition was introduced before transfer, and so the delay itself does not appear to have caused a move away from II strategies.

As a further analysis, and to verify our reasoning above, we repeated the accuracy-based analyses reported in the previous section for the subset of participants who used procedural strategies throughout the course of the experiment. Thus, the data of any participant that were best fit by an RB or guessing strategy, either during training or transfer, were excluded from this analysis. Qualitatively, the accuracy plots remained in agreement with Fig. 5. In particular, the same ordinal relations were observed among conditions. To assess for significance, ANOVAs were run as before. Despite the greatly reduced sample sizes (Reversal  $N = 14$ ; New Categories  $N = 10$ ; Reversal/Delay  $N = 6$ ), the effect due to condition was still significant during transfer [ $F(2, 27) = 4.03$ ,  $\eta^2 = 0.113$ ,  $\eta_p^2 = 0.230$ ,  $p = 0.029$ ], with performance better in the Reversal condition than in either the New Categories or Reversal/Delay conditions.

## Discussion

Our data analyses revealed a number of notable results. First, in line with previous findings, participants recovered more quickly from a full category reversal than from a rotation of the category structures that reversed the category membership of only half the stimuli. Decision-bound

modeling showed that this difference was likely not due to a switch from a procedural strategy to suboptimal one-dimensional rules.

Second, note that the initial learning curve for the Reversal group appears similar to the learning curve for this group during transfer. There are several differences though between the training and transfer phases which make it difficult to draw strong conclusions from this similarity. One is that at the beginning of initial training, accuracy is at chance (i.e., 50%), whereas at the beginning of transfer, accuracy is considerably below chance. Another critical difference is that participants must overcome the natural human tendency to persevere with response strategies (e.g., Ramage et al., 1999) during transfer but not during initial training.

Third, feedback-delay, which is known to impair striatal-based learning, drastically undermined recovery from a full reversal. This result is consistent with the hypothesis that reversal learning is dependent on the striatum.

This qualitative pattern of results is predicted by the two-stage model, but not by the one-stage model. Later in this section we will examine the quantitative fits of both models to the human data. First though, we will explore the implications of these results for other existing models of category learning.

The most successful models of category learning assume that S-R associations are formed through a sophisticated process. For example, consider ATRIUM (Erickson & Kruschke, 1998) and COVIS (Ashby et al., 1998). Both models assume that category learning is mediated by the interaction of two systems—one that learns S-R associations directly and one in which S-R associations are formed by applying an explicit rule. ATRIUM assumes that the direct system is a standard exemplar model, whereas COVIS assumes it is the striatal pattern classifier. Although both models are capable of accounting for a wide range of categorization behavior, they both seem incompatible with our results. During training, both ATRIUM and COVIS predict that the direct system will dominate performance (because optimal performance cannot be obtained with a simple rule). There are two possibilities for transfer: either the direct system will continue to dominate, in which case the models essentially reduce to simple single-system S-R models, or the models predict a switch to rule-based strategies. Neither of these options is consistent with our results. Single-system S-R models make the wrong predictions about the relative difficulty of recovering from a reversal compared to learning new categories while rule use predicts that decision bound models that assume a simple one-dimensional rule should dominate during transfer.

Models that assume S-R associations are formed via sophisticated context-dependent Bayesian inference

(Anderson, 1991; Gershman et al., 2010; Redish et al., 2007) might avoid catastrophic interference in the Reversal condition by recognizing the transfer phase as a new context. In this case, the previously learned S-R associations would simply be set aside, and a new set of associations would be learned from scratch. In our opinion, such context-sensitive models provide the best opportunity for single-stage models of learning to account for our results. Even so, such models face some major challenges. Transfer performance in context-sensitive models depends on their ability to recognize the transfer environment as a new context. In general, this is easy for environments that share few attributes, and difficult for those that are highly similar. In our experiments, there are no cues of the traditional type that signal transfer is occurring. Error rate increases in all conditions, and in the New Categories condition there are some novel stimuli (around 30%), but it is not clear how these changes could be used to define a new context. Furthermore, for an S-R model to account for our results it would have to recognize a new context in the Reversal condition, but not in the New Categories or Reversal/Delay conditions. We know of no current models capable of this discrimination. But more work is needed to determine whether such a model is possible.

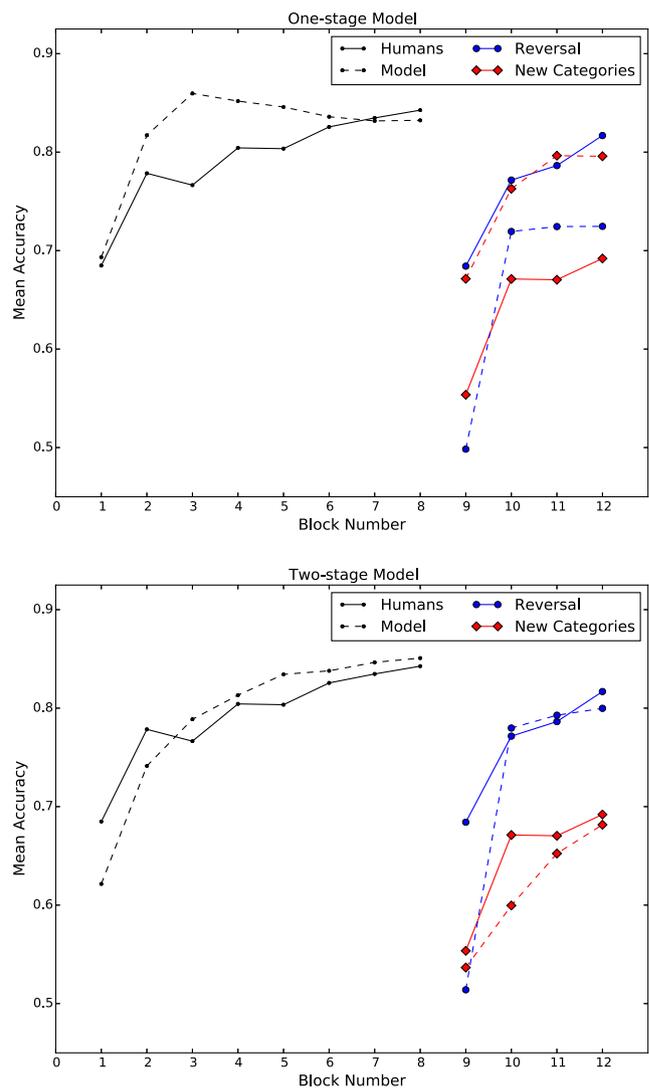
Even so, it is important to note that the issue of context-sensitive learning is logically unrelated to whether procedural learning is mediated by one or two stages. In fact, we have proposed elsewhere that striatal learning is context dependent (Crossley et al., 2013, 2014). So our hypothesis is that the best model of procedural learning will probably combine context sensitivity with two stages of learning.

In summary, previous models of procedural category learning seem unable to account for why reversal learning is easier than learning new categories. To evaluate the ability of the two-stage model to provide a good quantitative fit to our results, we fit it and the one-stage model described above to the data using particle swarm optimization. Using this approach, we found parameter estimates that minimized the sum of squared errors between the model and the human data.<sup>3</sup>

We chose to fit the models to the learning and transfer data of all participants, rather than just to the data from the subset of participants whose responses were best fit by a model that assumed a procedural strategy. First, some of the

sample sizes of participants in this latter group were small, so standard errors were large. Second, as mentioned previously, the data of the two groups were highly similar. Third, the decision bound modeling, like any statistical method, is susceptible to errors.

As in our earlier simulations, after parameter estimation was complete, 100 simulations were run with the best-fitting parameter values and the model predictions were computed by taking the mean across all 100 simulations. The results are presented in Fig. 6. As expected, the one-stage model again performs worse in the Reversal condition than in the New Categories condition. Thus, this prediction is made regardless of whether the model attempts to mimic human performance or to maximize accuracy.



**Fig. 6** Human data and the best-fitting predictions of both models to training and transfer for Reversal and New Categories conditions. Human data is plotted with *solid lines*, simulated with *dashed lines*

<sup>3</sup>We made no attempt to fit the data from the Reversal/Delay condition. The one-stage model has successfully accounted for the effects of various feedback delays (i.e., from 0 to 2.5 s) (Valentin et al., 2014), so it would be straightforward to model these data. Even so, we chose not to include such modeling. First, both models predict poorer performance in the Reversal/Delay condition than in the Reversal condition, because all learning occurs in the striatum in both models. Second, modeling the Reversal/Delay data would require extra free parameters.

To test the generality of this result, we repeated these simulations using a single-layer perceptron (Bishop et al., 2006; Rosenblatt, 1958). This is a popular machine-learning model that learns linear classification boundaries. The results were qualitatively similar to the single-stage model. The best performing version of the perceptron learned the new categories faster than the reversed categories. A brute force exploration of the parameter space failed to find parameters for which the reversal was learned faster than the new categories. Thus, even taking a more general machine-learning approach and abandoning biological constraints, single-stage models appear unable to account for the results of this experiment.

In contrast, the two-stage model provides good quantitative and qualitative fits to the human data. Qualitatively, it captures the major features of the data, including (i) consistent incremental learning during initial training, (ii) faster recovery from a reversal than from transfer to new categories, and (iii) that recovery in both conditions never reaches final-block training accuracy. Quantitatively, the fit is also quite good, at least for 15 of the 16 data points. The single exception is the first transfer block (i.e., block 9) in the Reversal condition, where the two-stage model performs notably worse than the human participants. Figure 3 shows that this is not a necessary prediction of the model, however. Rather, it appears to be a feature of the particular set of parameter values that minimize sum of squared errors.

## General discussion

We proposed a revision of a classic procedural-learning model that assumes a single cortical-basal ganglia loop. The new model includes two stages implemented via two sequentially connected basal ganglia loops with DA-dependent synaptic plasticity at both sets of cortical-striatal synapses. Stage 1 learns about category structure and stage 2 learns about response mappings. We reported new experimental results that were strongly incompatible with classic S-R models of procedural learning but consistent with this new account. Our experimental data showed that recovery after a full reversal, in which all category labels are reversed, is faster than learning new categories in which only half the stimuli reverse category membership, but only when feedback is provided immediately after the response. This difference is completely abolished if feedback is delayed by just a few seconds. This, we argue, supports the hypothesis that the categories were learned in two stages (since full-reversals are less impaired) and that reversals are mediated within the basal ganglia (since delayed feedback removes this effect).

In addition to the considerable evidence that there are interactions between otherwise “distinct” cortical-striatal loops, our proposal demonstrates a potential advantage to such an open interconnectedness. Such interactions may allow for significantly more flexible behaviors, and could even provide benefits in fixed scenarios, such as the increased learning rate described above.

The model we propose assumes that stage 1 learning is mediated via projections between visual cortical areas and the caudate body and tail, whereas stage 2 is assumed to be mediated through projections between preSMA and the posterior putamen. There is already good evidence that the body and tail of the caudate receive similar DA inputs as the posterior putamen (e.g., Haber et al., 2006). For this reason, we assumed that learning in both stages is mediated by a single global DA signal that is identical at every cortical-striatal synapse. Global learning signals of this type greatly reduce model flexibility, and in the present case, they make the model potentially susceptible to the credit-assignment problem. Thus, it is noteworthy that our model was able to learn at all. More importantly, however, it also provides a significantly improved account of the behavioral data. This was achieved despite attempts to significantly constrain model construction to biologically grounded mechanisms and anatomy.

The two-stage model in Fig. 2 is superficially similar to a traditional three-layer connectionist network since both allow for learning at two successive stages. In three-layer connectionist networks, the weights are modifiable between the input layer and the intermediate or hidden layer, and between the hidden and output layers. Adding a hidden layer to connectionist networks has long been known to add to their flexibility in accounting for a range of behaviors. Three-layer connectionist networks, however, are fundamentally different from the model proposed here in at least two important ways. First, the weights in connectionist networks are typically adjusted through some gradient-descent learning algorithm (e.g., back-propagation) and therefore each synaptic weight is trained via its own independent teaching signal. In contrast, as already noted, our network assumes a more biologically plausible global reinforcement learning signal that is identical at all synapses. Second, connectionist models are typically not constrained by neuroanatomy, whereas the architecture of the models described in Fig. 2 was chosen to mimic known neuroanatomy.

In summary, empirical results demonstrate the need for a move away from classic S-R models of category learning but are compatible with the assumption of a multiple-stage process. We proposed a two-stage generalization of a classic one-stage procedural category-learning model, in

which both stages are learned striatally. Our own empirical investigations support the hypothesis that a second stage of category learning, which associates a response with a category structure, is mediated within the basal ganglia. Via simulations, the model was shown to provide a good account of the data. Future research should investigate the possible effects of a second category-learning stage on tasks other than the simple reversal and new category tasks studied here.

**Author Notes** This research was supported in part by NINDS grant P01NS044393, NIMH grant 2R01MH063760, and by AFOSR grant FA9550-12-1-0355.

### Appendix: Computational cognitive neuroscience modeling

**Neural network** Visual input was simulated using an ordered array of 10,000 sensory units, each tuned to a different stimulus. We assumed that each unit responds maximally when its preferred stimulus is presented and that its response decreases as a Gaussian function of the distance in stimulus space between the stimulus preferred by that unit and the presented stimulus. Activation in each unit was either 0 or equal to some positive constant value during the duration of stimulus presentation. Specifically, we assumed that when a stimulus is presented, the activation of sensory cortical unit  $k$  is given by:

$$I_k = A \exp\left(\frac{-d^2}{a}\right), \quad (\text{A1})$$

where  $d$  is the distance in stimulus space between the preferred stimulus of unit  $k$  and the presented stimulus.  $A$  and  $a$  are both constants, determining the maximum levels of activation and the width of the tuning curves.

In the first striatal stage of the one-stage and two-stage models, there are two medium spiny units. Denote these as  $j$  and  $j'$ . The activation in striatal unit  $j$  at time  $t$ , denoted  $V_j(t)$ , was determined by the following version of the Izhikevich (2007) model of the medium spiny neuron:

$$50 \frac{dV_j(t)}{dt} = I_j(t) + [V_j(t) + 80][V_j(t) + 25] - U_j(t) \quad (\text{A2})$$

$$100 \frac{dU_j(t)}{dt} = -20[V_j(t) + 80] - U_j(t) \quad (\text{A3})$$

with

$$I_j(t) = \sum_k w_{jk}(n) I_k(t) - 3.5 f[V_{j'}(t)] + 100\epsilon(t) \quad (\text{A4})$$

where  $w_{jk}(n)$  is the strength of the synapse between sensory cortical unit  $k$  and striatal unit  $j$  on trial  $n$ , and  $\epsilon(t)$  Gaussian white noise. To produce spikes, when  $V_j(t)$  reaches 40 mV it is reset to -55 mV, and  $U(t)$  is reset to  $U(t) + 150$ . Input to the Izhikevich model is calculated according to Eq. A4, the second term of which models lateral inhibition from striatal unit  $j'$  using the alpha function  $f[x]$  (e.g., Rall, 1967), a standard method for modeling the temporal smearing that occurs postsynaptically when a presynaptic neuron fires a spike. Specifically, if striatal unit  $j'$  spikes at time  $t_0$ , the following input is delivered to striatal unit  $j$ :

$$\alpha(t') = \frac{t'}{\lambda} \exp\left(\frac{\lambda - t'}{\lambda}\right) \quad (\text{A5})$$

with  $t' = t - t_0$ . The alpha function is used consistently to model the post-synaptic effects of a spike in every unit of both models.

The two-stage model also has a second pair of medium spiny neurons, this time in the putamen (see Fig. 2). These are likewise governed by Eqs. A2 and A3 only this time the equation for  $I$  is replaced by:

$$I_j(t) = \sum_k w_{jk}(n) f[V_k(t)] - 3.5 f[V_{j'}(t)] + 100\epsilon(t) \quad (\text{A6})$$

where the sum over units  $k$  indicates the units from preSMA, and the unit  $j'$  is the lateral in putamen.

Activation in cortical regions (SMA and preSMA) is modeled using the Izhikevich equations as well. However, the parameters chosen here were those of “regular spiking neurons” (Izhikevich, 2007).

Activation in all other units is modeled with the quadratic integrate-and-fire (QIF) model (see e.g., Ermentrout & Kopell, 1986, or Izhikevich, 2007). The QIF model assumes that the intracellular voltage at time  $t$  is given by:

$$\frac{dV(t)}{dt} = I(t) + \beta + 0.117[V(t) + 60][V(t) + 40], \quad (\text{A7})$$

where  $\beta$  is a constant that determines the baseline firing rate,  $I$  is the total driving current,  $-60$  is the resting membrane potential, and  $-40$  is the instantaneous threshold potential. Equation A7 produces the upstroke of action potentials by itself, but not the downstroke. To create spikes, when  $V(t)$  reaches 35 it is reset to  $-50$ .

The architecture of both models is shown in Fig. 2. In the one-stage model, strong activation in the striatum inhibits the tonically active GPi, which in turn allows thalamus to

excite cortex. The two-stage model has analogous behavior for both loops through the striatum (i.e., caudate and putamen).

All differential equations were solved numerically using Euler’s method with time steps of 1 ms. To generate motor responses, thresholds were placed on the number of spikes produced by each SMA unit. As soon as one unit spiked enough to reach threshold, the model was considered to have made the corresponding response. If no response was made within 3000 ms, the trial ended and no response was recorded.

**Learning** Initially, the weights  $w_{jk}$ , which codify the strength of projections from cortex to striatum, were set at random uniformly in the [0.1, 0.2) interval. For the two-stage model, the  $w_{jk}$  from preSMA to putamen were set at random uniformly in the [1.75, 1.85) interval. Following this, all cortical-striatal synaptic weights were then updated after each trial via a biologically plausible reinforcement learning algorithm (Ashby & Helie, 2011). In the one-stage model, learning occurred only at synapses between the sensory cortical units and the medium spiny units in the striatum. In the two-stage model, learning occurred at these same synapses but also at synapses between preSMA units and putamen units. After feedback on trial  $n$ , the strength of the synapse between units  $a$  and  $b$  was adjusted as follows:

$$\begin{aligned}
 w_{ab}(n+1) = & w_{ab}(n) \\
 & + \alpha_w \int I_a(t) dt \left[ \int [V_b(t)]^+ dt - \theta_{\text{NMDA}} \right]^+ \\
 & \times [D(n) - D_{\text{base}}]^+ [w_{\text{max}} - w_{ab}(n)] \\
 & - \beta_w \int I_a(t) dt \left[ \int [V_b(t)]^+ dt - \theta_{\text{NMDA}} \right]^+ \\
 & \times [D_{\text{base}} - D(n)]^+ w_{ab}(n) \\
 & - \gamma_w \int I_a(t) dt \left\{ \left[ \theta_{\text{NMDA}} - \int [V_b(t)]^+ dt \right]^+ \right. \\
 & \left. - \theta_{\text{AMPA}} \right\}^+ w_{ab}(n) \tag{A8}
 \end{aligned}$$

where  $I_a$  is the driving current at  $b$  due to action potentials from  $a$ , and  $[g(t)]^+$  is equal to  $g(t)$  when  $g(t) > 0$  and 0 otherwise.  $\theta_{\text{NMDA}}$  and  $\theta_{\text{AMPA}}$  represent the activation thresholds for post-synaptic NMDA and AMPA glutamate receptors, respectively. In Eq. A8 these parameters act as thresholds so that synaptic strengthening (i.e., LTP) occurs only if post-synaptic activation exceeds  $\theta_{\text{NMDA}}$ . If activation is between  $\theta_{\text{AMPA}}$  and  $\theta_{\text{NMDA}}$  then the synapse is weakened, and no change in the synaptic weight occurs if activation is below  $\theta_{\text{AMPA}}$ .

$D(n)$  represents the amount of dopamine released in response to the trial  $n$  feedback and  $D_{\text{base}}$  represents the baseline dopamine level. Following previous applications

(e.g., Ashby & Crossley, 2011), we assumed that dopamine release was a piecewise linear function of the reward prediction error (RPE), which is defined as the difference between obtained reward on trial  $n$ ,  $R_n$ , and the predicted reward,  $P_n$ ; that is,

$$\text{RPE} = R_n - P_n \tag{A9}$$

$R_n$  is set to 1 when reward feedback is given, 0 in the absence of feedback, and  $-1$  when error feedback is given. The predicted reward,  $P_n$  is calculated using the single-operator model (Bush & Mosteller, 1951):

$$P_{n+1} = P_n + 0.025(R_n - P_n), \tag{A10}$$

Finally, we assumed that the amount of dopamine release is related to the RPE in the manner consistent with the dopamine firing data reported by (Bayer and Glimcher, 2005). Specifically, we assumed that

$$D(n) = \begin{cases} 1 & \text{if RPE} > 1 \\ .8 \text{ RPE} + .2 & \text{if } -.25 \leq \text{RPE} \leq 1 \\ 0 & \text{if RPE} < -.25 \end{cases} \tag{A11}$$

Note that the baseline dopamine level is 0.2 (i.e., when the RPE = 0) and that dopamine levels increase linearly with the RPE. However, note also the asymmetry between dopamine increases and decreases. As is evident in the Bayer and Glimcher (2005) data, a negative RPE quickly causes dopamine levels to fall to zero, whereas there is a considerable range for dopamine levels to increase in response to positive RPEs.

**Parameter estimation** Most parameters in the models were fixed during model construction. For example, the parameters in the medium spiny neuron units were taken from Ashby and Crossley (2007). The parameters for the units in the other regions were chosen so that a reasonable change in the firing rate of the pre-synaptic unit would cause a reasonable change in the firing rate of the post-synaptic unit. The baseline firing rate of all cortical units was set to a low value, whereas the baseline firing rate of thalamic and GPi units was set to a high value. These values were chosen because GPi neurons have a high tonic firing rate and so as to mimic other excitatory inputs to thalamus. This allows the thalamus to excite cortex when the tonic inhibition from the GPi is reduced by striatal firing. The height of the radial basis function was chosen to be large enough so that visual activation increased striatal firing, and the radial basis function width was set to 1/8th of the stimulus space (analogous to a standard deviation). The parameters in the dopamine model were taken from Ashby and Crossley (2011).

Thus, only the parameters  $\alpha_w$  and  $\beta_w$  of the learning equation (A8) were adjusted independently for the two models. In model 1 there is one  $\alpha$  and one  $\beta$ , corresponding to the rates of LTP and LTD at synapses from visual cortex into

**Table 2** Estimated learning rates for each model

	1-stage	2-stage
Best fit to human data		
$\alpha_1$	$1.09 \times 10^{-7}$	$2.13 \times 10^{-10}$
$\beta_1$	$1.45 \times 10^{-7}$	$2.14 \times 10^{-8}$
$\alpha_2$		$8.39 \times 10^{-7}$
$\beta_2$		$2.75 \times 10^{-8}$
Maximum accuracy		
$\alpha_1$	$5.73 \times 10^{-8}$	$3.45 \times 10^{-8}$
$\beta_1$	$1.31 \times 10^{-7}$	$4.04 \times 10^{-8}$
$\alpha_2$		$3.17 \times 10^{-7}$
$\beta_2$		$1.11 \times 10^{-7}$

striatum. Model 2 has two  $\alpha$  and two  $\beta$  – one pair for each stage of learning. Particle swarm optimization was used to find optimal values for these parameters. See Table 2 for values.

## References

- Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, 9(1), 357–381.
- Ashby, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, 98(3), 409.
- Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, 105(3), 442.
- Ashby, F. G., & Crossley, M. J. (2011). A computational model of how cholinergic interneurons protect striatal-dependent learning. *Journal of Cognitive Neuroscience*, 23(6), 1549–1566.
- Ashby, F. G., Ell, S. W., & Waldron, E. M. (2003). Procedural learning in perceptual categorization. *Memory & Cognition*, 31(7), 1114–1125.
- Ashby, F. G., & Ennis, J. M. (2006). The role of the basal ganglia in category learning. *Psychology of Learning and Motivation*, 46(1).
- Ashby, F. G., Ennis, J. M., & Spiering, B. J. (2007). A neurobiological theory of automaticity in perceptual categorization. *Psychological Review*, 114(3), 632.
- Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 33–53.
- Ashby, F. G., & Helie, S. (2011). A tutorial on computational cognitive neuroscience: Modeling the neurodynamics of cognition. *Journal of Mathematical Psychology*, 55(4), 273–289.
- Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, 56, 149–178.
- Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy of Sciences*, 1224(1), 147–161.
- Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Science*, 2, 83–89.
- Ashby, F. G., & Waldron, E. M. (1999). On the nature of implicit categorization. *Psychonomic Bulletin & Review*, 6(3), 363–378.
- Badgaiyan, R. D., Fischman, A. J., & Alpert, N. M. (2007). Striatal dopamine release in sequential learning. *Neuroimage*, 38, 549–556.
- Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, 47(1), 129–141.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York: Springer.
- Bush, R. R., & Mosteller, F. (1951). A model for stimulus generalization and discrimination. *Psychological Review*, 58(6), 413.
- Cincotta, C. M., & Seger, C. A. (2007). Dissociation between striatal regions while learning to categorize via feedback and via observation. *Journal of Cognitive Neuroscience*, 19(2), 249–265.
- Clarke, H. F., Robbins, T. W., & Roberts, A. C. (2008). Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. *The Journal of Neuroscience*, 28(43), 10972–10982.
- Clerc, M. (2012). Standard particle swarm optimisation. *Open access archive HAL*.
- Cools, R., Clark, L., Owen, A. M., & Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *The Journal of Neuroscience*, 22(11), 4563–4567.
- Crossley, M. J., Ashby, F. G., & Maddox, W. T. (2013). Erasing the engram: The unlearning of procedural skills. *Journal of Experimental Psychology: General*, 142(3), 710–741.
- Crossley, M. J., Ashby, F. G., & Maddox, W. T. (2014). Context-dependent savings in procedural category learning. *Brain and Cognition*, 92, 1–10.
- DeLong, M., Georgopoulos, A., & Crutcher, M. (1983). Cortico-basal ganglia relations and coding of motor performance. *Experimental Brain Research*, 7, 30–40. suppl.
- Dunn, J. C., Newell, B. R., & Kalish, M. L. (2012). The effect of feedback delay and feedback type on perceptual category learning: the limits of multiple systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(4), 840.
- Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, 127(2), 107.
- Ermentrout, G. B., & Kopell, N. (1986). Parabolic bursting in an excitable system coupled with a slow oscillation. *SIAM Journal on Applied Mathematics*, 46(2), 233–253.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.
- Filoteo, J. V., Paul, E. J., Ashby, F. G., Frank, G. K., Helie, S., Rockwell, R., et al. (2014). Simulating category learning and set shifting deficits in patients weight-restored from anorexia nervosa. *Neuropsychology*, 28(5), 741–751.
- Gershman, S. J., Blei, D. M., & Niv, Y. (2010). Context, learning, and extinction. *Psychological Review*, 117(1), 197.
- Grafton, S. T., Hazeltine, E., & Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *Journal of Cognitive Neuroscience*, 4, 497–510.
- Haber, S. N., Kim, K. S., Maily, P., & Calzavara, R. (2006). Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. *The Journal of Neuroscience*, 26(32), 8368–8376.
- Hélie, S., Paul, E. J., & Ashby, F. G. (2012). Simulating the effects of dopamine imbalance on cognition: From positive affect to Parkinson's disease. *Neural Networks*, 32, 74–85.
- Hikosaka, O., Nakamura, K., & Nakahara, H. (2006). Basal ganglia orient eyes to reward. *Journal of Neurophysiology*, 95(2), 567–584.
- Izhikevich, E. M. (2007). *Dynamical systems in neuroscience*. Cambridge, MA: MIT Press.

- Jackson, S., & Houghton, G. (1995). Sensorimotor selection and the Basal Ganglia: A neural network model. In Houk, J. C., Davis, J. L., & Beiser, D. G. (Eds.) *Models of information processing in the Basal Ganglia* (pp. 337–368): Cambridge, MA: MIT Press.
- Joel, D., & Weiner, I. (1994). The organization of the basal ganglia-thalamocortical circuits: open interconnected rather than closed segregated. *Neuroscience*, *63*(2), 363–379.
- Knopman, D., & Nissen, M. J. (1991). Procedural learning is impaired in Huntington's disease: Evidence from the serial reaction time task. *Neuropsychologia*, *29*(3), 245–254.
- Kruschke, J. K. (1996). Dimensional relevance shifts in category learning. *Connection Science*, *8*(2), 225–247.
- Lawrence, A. D., Sahakian, B. J., & Robbins, T. W. (1998). Cognitive functions and corticostriatal circuits: insights from Huntington's disease. *Trends in Cognitive Sciences*, *2*(10), 379–388.
- Lisman, J., Schulman, H., & Cline, H. (2002). The molecular basis of CaMKII function in synaptic and behavioural memory. *Nature Reviews Neuroscience*, *3*(3), 175–190.
- Lopez-Paniagua, D., & Seger, C. A. (2011). Interactions within and between corticostriatal loops during component processes of category learning. *Journal of Cognitive Neuroscience*, *23*(10), 3068–3083.
- Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, *53*(1), 49–70.
- Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2003). Delayed feedback effects on rule-based and information-integration category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 650/662.
- Maddox, W. T., Bohil, C. J., & Ing, A. D. (2004). Evidence for a procedural-learning-based system in perceptual category learning. *Psychonomic Bulletin & Review*, *11*(5), 945–952.
- Maddox, W. T., Glass, B. D., O'Brien, J. B., Filoteo, J. V., & Ashby, F. G. (2010). Category label and response location shifts in category learning. *Psychological Research*, *74* (2), 219–236.
- Maddox, W. T., & Ing, A. D. (2005). Delayed feedback disrupts the procedural-learning system but not the hypothesis testing system in perceptual category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31* (1), 100–107.
- McAlonan, K., & Brown, V. J. (2003). Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. *Behavioural Brain Research*, *146*(1), 97–103.
- McFarland, N. R., & Haber, S. N. (2000). Convergent inputs from thalamic motor nuclei and frontal cortical areas to the dorsal striatum in the primate. *The Journal of Neuroscience*, *20* (10), 3798–3813.
- Middleton, F. A., & Strick, P. L. (2000). Basal ganglia output and cognition: evidence from anatomical, behavioral, and clinical studies. *Brain and Cognition*, *42*(2), 183–200.
- Nakano, K., Tokushige, A., Kohno, M., Hasegawa, Y., Kayahara, T., & Sasaki, K. (1992). An autoradiographic study of cortical projections from motor thalamic nuclei in the macaque monkey. *Neuroscience Research*, *13*(2), 119–137.
- Newell, B. R., Dunn, J. C., & Kalish, M. (2011). Systems of Category Learning: Fact or Fantasy? *Psychology of Learning and Motivation-Advances in Research and Theory*, *54*, 167.
- Nomura, E. M., Maddox, W. T., Filoteo, J. V., Ing, A. D., Gitelman, D. R., Parrish, T. B., et al. (2007). Neural correlates of rule-based and information-integration visual category learning. *Cerebral Cortex*, *17*(1), 37–43.
- Peirce, J. W. (2007). PsychoPy – Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13.
- Rall, W. (1967). Distinguishing theoretical synaptic potentials computed for different soma-dendritic distributions of synaptic input. *Journal of Neurophysiology*, *30*(5), 1138–1168.
- Ramage, A., Bayles, K., Helm-Estabrooks, N., & Cruz, R. (1999). Frequency of perseveration in normal subjects. *Brain and Language*, *66*(3), 329–340.
- Redish, A. D., Jensen, S., Johnson, A., & Kurth-Nelson, Z. (2007). Reconciling reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling. *Psychological Review*, *114*(3), 784–805.
- Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386.
- Rudy, J. W. (2014). *The neurobiology of learning and memory* (2nd Edition). Sunderland, MA: Sinauer.
- Sanders, B. (1971). Factors affecting reversal and nonreversal shifts in rats and children. *Journal of Comparative and Physiological Psychology*, *74*, 192–202.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.
- Seger, C. A. (2008). How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. *Neuroscience & Biobehavioral Reviews*, *32*(2), 265–278.
- Seger, C. A., & Cincotta, C. M. (2002). Striatal activity in concept learning. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(2), 149–161.
- Spence, K. W. (1940). Continuous versus non-continuous interpretations of discrimination learning. *Psychological Review*, *47*(4), 271–288.
- Spiering, B. J., & Ashby, F. G. (2008). Response processes in information-integration category learning. *Neurobiology of Learning and Memory*, *90*(2), 330–338.
- Valentin, V. V., Maddox, W. T., & Ashby, F. G. (2014). A computational model of the temporal dynamics of plasticity in procedural learning: sensitivity to feedback timing. *Frontiers in Psychology*, *5*, 643.
- Waldschmidt, J. G., & Ashby, F. G. (2011). Cortical and striatal contributions to automaticity in information-integration categorization. *Neuroimage*, *56*(3), 1791–1802.
- Wills, A., Noury, M., Moberly, N. J., & Newport, M. (2006). Formation of category representations. *Memory & Cognition*, *34*(1), 17–27.
- Worthy, D. A., Markman, A. B., & Maddox, W. T. (2013). Feedback and stimulus-offset timing effects in perceptual category learning. *Brain and Cognition*, *81*(2), 283–293.