Published in final edited form as:
Ann N Y Acad Sci. 2011 April ; 1224: 147–161. doi:10.1111/j.1749-6632.2010.05874.x.

Human Category Learning 2.0

F. Gregory Ashby and
University of California, Santa Barbara

W. Todd Maddox
University of Texas, Austin

Abstract

During the 1990's and early 2000's, cognitive neuroscience investigations of human category learning focused on the primary goal of showing that humans have multiple category learning systems and on the secondary goals of identifying key qualitative properties of each system and of roughly mapping out the neural networks that mediate each system. Many researchers now accept the strength of the evidence supporting multiple systems, and as a result, during the past few years, work has begun on the second generation of research questions – that is, on questions that begin with the assumption that humans have multiple category learning systems. This article reviews much of this second generation of research. Topics covered include: 1) How do the various systems interact? 2) Are there different neural systems for categorization and category representation? 3) How does automaticity develop in each system?, and 4) Exactly how does each system learn?

Keywords

categorization; memory system interactions; automaticity; striatum

Cognitive neuroscience investigations of human category learning began in the second half of the 1990's.[1,2] During the ensuing decade, much of this work focused on the primary goal of showing that humans have multiple category learning systems and on the secondary goals of identifying key qualitative properties of each system and of roughly mapping out the neural networks that mediate each system. Although this work continues, many researchers now accept the strength of the evidence supporting multiple systems. As a result, during the past few years, work has begun on a second generation of research questions. By second generation, we mean questions that begin with the assumption that humans have multiple category learning systems. Included in this list are questions such as: 1) How do the various systems interact? 2) Do different neural networks mediate categorization and category representation? 3) How does automaticity develop in each system?, and 4) Exactly how does each system learn? Although a number of recent publications review work on the first generation of research questions,[3,4] no existing reviews focus on these second generation questions. This article reviews this more recent work.

Brief Review of First Generation Research

With the realization that humans might have different category-learning systems that are each adept at learning about different types of category structures came more careful

Correspondence concerning this article should be addressed to F. Gregory Ashby, Department of Psychology, University of California, Santa Barbara, CA 93106 (ashby@psych.ucsb.edu) or W. Todd Maddox, Department of Psychology, University of Texas, Austin, TX, 78712 (maddox@psy.utexas.edu).

attempts to characterize existing category-learning tasks. One popular nomenclature includes rule-based (RB), information-integration (II), prototype-distortion, and unstructured category-learning tasks.[5] Briefly, RB tasks are those where the optimal strategy is easy to verbalize (e.g., long objects are in one category and short objects are in another) and the categories can be learned via a logical reasoning process (e.g., by hypothesis testing). Optimal performance in II tasks requires integrating perceptual information from different stimulus components at a pre-decisional level. In II tasks the optimal strategy is not easily verbalized. In prototype distortion tasks, the category exemplars are created by randomly distorting a category prototype. Finally, in unstructured (or ad hoc) tasks the category exemplars are defined arbitrarily, rather than according to similarity (as in prototype distortion) or on the basis of some abstract logical (as in RB) or mathematical (as in II) relationship. There are proposals that these four tasks primarily load on different memory systems,[5] although this hypothesis is far from resolved.

Of these four tasks, the most widely studied is the RB task, followed in order by the II, the prototype distortion, and the unstructured tasks. As described below, much evidence suggests that learning in RB tasks depends on an explicit, hypothesis-testing system that is mediated by a broad neural network that includes the prefrontal cortex, anterior cingulate, head of the caudate nucleus, and the hippocampus and other medial temporal lobe structures. In contrast, the evidence suggests that II tasks recruit a procedural learning system that depends heavily on the striatum and on dopamine-mediated reinforcement learning at cortical-striatal synapses. Less is known about prototype distortion and unstructured tasks, but several studies suggest that the most popular prototype distortion task recruits the perceptual representation memory system.[6,7]

The evidence for multiple systems came primarily from two types of studies – behavioral experiments that reported empirical dissociations between different category-learning tasks and cognitive neuroscience studies that used either neuroimaging or else studied various neuropsychological patient groups. The dissociation studies manipulated variables that affected the structure of the categories, the placement, timing, and richness of the feedback, the response characteristics associated with the task, the number of categories and decision bounds, and the working memory load, to name a few. For example, delaying feedback by a few seconds,[8,9] switching the location of the response keys,[10,11] or informing participants of the category label before the stimulus rather than after the response[12] all interfere with performance in II tasks much more than in RB tasks. In contrast, adding a secondary (dual) task[13,14] or reducing the time available to process the feedback[15] interferes with performance in RB tasks much more than in II tasks. No single-system account of all these dissociations has been proposed.

Neuroimaging and neuropsychological patient studies helped map out the neural circuits active in these different tasks. Neuroimaging studies of all four tasks have been reported, and a variety of patient groups have been tested, including patients with Parkinson's disease, Huntington's disease, frontal-lobe damage, medial temporal lobe amnesia, and schizophrenia.[16] Perhaps the most consistent result from all of this work is the finding that almost all forms of category learning depend, to some extent, on the striatum (although different tasks may depend on different striatal subregions).

One theoretical reason why the striatum is seen as a plausible site for initial category learning is that the conditions under which cortical-striatal synapses are strengthened and weakened closely match the conditions for reinforcement learning[17–19] with dopamine serving as the reinforcement training signal. In accordance with this observation, Parkinson's patients, who have reduced brain dopamine levels, are impaired in a wide variety of category learning tasks.[2,20,21] In addition, II category learning, which is thought

to rely primarily on reinforcement learning at cortical-striatal synapses,[1] is disrupted by feedback delays just as predicted by reinforcement learning.[8,9] For a thorough review of the evidence implicating the striatum in category learning see Ashby and Ennis[22] or Seger[23].

Although many cognitive theories of categorization have been proposed,[3] only a few theories have been developed in enough neurobiological detail to make predictions in cognitive neuroscience experiments. The COVIS theory of category learning assumes separate explicit and procedural-learning categorization systems that compete for access to response production.[1,24] The explicit system selects and tests simple verbalizable hypotheses about category membership. The procedural system gradually associates categorization responses with regions of perceptual space via reinforcement learning. COVIS assumes that RB categorization is mediated by a broad neural network that includes the prefrontal cortex, anterior cingulate, head of the caudate nucleus, and the hippocampus. The network that includes the prefrontal cortex, anterior cingulate, and head of the caudate nucleus is assumed to select, test, and switch among alternative hypotheses using working memory and executive attention,[25] whereas the primary role of the hippocampus in COVIS is to mediate the long-term retention of RB learning.[26] To perform well in RB tasks, participants must remember which rules they have already tested and rejected, in order to avoid revisiting these failed rules again. Thus, COVIS predicts normal learning by medial temporal lobe amnesiacs in simple RB tasks where the correct rule can be discovered before the list of rejected hypotheses is lost from working memory. In more difficult RB tasks, the search for the correct rule will exceed working memory capacity, so COVIS predicts that in these cases medial temporal lobe amnesiacs will be impaired. Much evidence supports the former prediction.[27,28] The latter prediction has not been as rigorously tested, although several studies have reported normal performance by amnesiacs on the first 50 trials of a difficult task, but impaired performance later on.[2,29]

The key structures in the COVIS procedural-learning system are the putamen and the premotor cortex (i.e., the supplementary motor area). Early versions of COVIS assumed that the striatal regions most critical to the procedural system were the body and tail of the caudate nucleus.[1] More recent evidence however, suggests that the procedural system has a strong motor association,[10,11] which has caused the focus to switch to the putamen. The key site of learning in this model is at cortical-striatal synapses, and this synaptic plasticity is presumed to be facilitated by a dopamine-mediated reinforcement training signal from the substantia nigra pars compacta.

Interactions among the Systems

Once the multiple systems hypothesis is accepted, a next obvious question is how the various systems interact. The available evidence suggests competition. Less clear is whether this competition is at the learning or output stages.

Behavioral Studies

An ideal task for studying interactions between two systems is one in which each system is active on different trials and the experimenter is reasonably confident of which responses are controlled by each system. Two studies have attempted to achieve these goals.[30,31] Both studies used a hybrid category-learning task in which perfect accuracy required applying an RB strategy on some trials and an II strategy on the others. In the Ashby and Crossley[30] study sine-wave gratings with an orientation of greater than 50 degrees required an RB strategy, whereas gratings with a shallower orientation required an II strategy. In the Erickson experiments,[31] three cues signaled participants which strategy to apply. First, the stimuli requiring an II strategy were perceptually distinct from the stimuli requiring an RB strategy. Second, stimuli requiring an II strategy were presented in one color, whereas

stimuli requiring an RB strategy were presented in a different color. Third, the II categories required different responses from the RB categories (i.e., A and B versus C and D).

Participants in the Erickson[31] study successfully switched between RB and II strategies on a trial-by-trial basis. In contrast, participants in the Ashby and Crossley[30] experiments did not switch. Instead, almost all of these participants used either a single RB or II strategy on all trials. By far, the most common strategy was RB.

If the two systems can operate independently then people should have performed well in both tasks. Taken together, these data suggest that switching between systems does not occur automatically. The prevalence of RB strategies in the Ashby and Crossley[30] experiments suggests that use of an RB strategy might limit access to the procedural system, either because the procedural system is prevented from learning or because it is denied access to motor output systems. The Erickson[31] results however, suggest that this inhibition can be overcome if enough cues are provided to signal the participant which system should be used on each trial.

Cognitive Neuroscience Studies

A number of neuroimaging studies have reported negative correlations between medial temporal lobe and striatal activation that are consistent with a competitive relationship between declarative and procedural memory systems during category learning.[32–37] Similar results have been reported within the more general memory-systems literature.[38–40] In addition, a number of animal lesion studies have reported that medial temporal lobe lesions can improve performance in striatal-dependent habit-learning tasks, and conversely that striatal lesions can improve performance in medial temporal lobe-dependent spatial learning tasks.[41–43]

All data reviewed so far suggest that the use of declarative memory either inhibits procedural learning, or denies access of the procedural-learning system to control of the response. Although this issue is far from resolved, some results support the latter of these two hypotheses – that is, that the inhibition is at the output stage, and that the use of one system does not necessarily inhibit learning in the other. First, Packard and McGaugh reported that animals displaying hippocampal-dependent (place learning) behavior immediately exhibited behavior that showed prior striatal-dependent (response) learning following inactivation of the hippocampus.[44] Similarly, animals displaying striatal-dependent behavior immediately exhibited hippocampal-dependent behavior following inactivation of the striatum. Second, using fMRI, Foerde, Knowlton, and Poldrack reported that the introduction of a secondary task shifted the brain region that correlated with learning from the hippocampus to the striatum, but that the overall level of striatal activation was equal in the two conditions.[45] Furthermore, Foerde, Poldrack, and Knowlton reported evidence that this striatal activation had a behavioral effect.[46] In particular, they showed that a dual task that impaired learning in declarative memory systems in probabilistic classification did not prevent implicit learning of the correct cue–response associations. These results are consistent with the hypothesis that the use of strategies that depend on declarative memory systems does not prevent simultaneous striatal-mediated procedural learning, but it does restrict access of that procedural learning to motor output systems.

Although these cognitive neuroscience results are consistent with the hypothesis that there is inhibition between declarative and procedural memory systems during category learning, it is important to note that they do not necessarily imply that this inhibition is between the COVIS explicit and procedural learning systems. For example, another possibility is that the declarative memory system tapped in these (human) studies may have been episodic

memory-based explicit memorization, rather than working memory-based hypothesis-testing.

A Tentative Theory

What brain mechanisms could be mediating this type of output inhibition? Ashby and Crossley tentatively proposed that frontal cortex and the subthalamic nucleus might control system interactions via the hyperdirect pathway through the basal ganglia.[30] The hyperdirect pathway begins with direct excitatory projections from frontal cortex to the subthalamic nucleus.[47,48] Some evidence suggests that the key cortical input may come from the pre-supplementary motor area.[49] The subthalamic nucleus then sends excitatory projections directly to the internal segment of the globus pallidus (GPi).[50,51] This extra excitatory input to the GPi tends to offset inhibitory input from the striatum, making it more difficult for striatal activity to affect cortex. Hence, the hyperdirect pathway could permit (by reducing subthalamic activity), or prevent (by increasing subthalamic activity) signals coming from the striatum from influencing cortex.

Note that this hypothesis accounts for asymmetries in the Ashby and Crossley[30] data. Specifically, Ashby and Crossley found that the use of an explicit strategy prevented access to procedural knowledge, but they failed to find evidence of the opposite influence. The hyperdirect pathway provides a mechanism via which the prefrontal cortex can inhibit a response selected by the striatum, but it does not allow the striatum to inhibit a response selected by the prefrontal cortex. Furthermore, because frontal cortex controls the excitatory input to the hyperdirect pathway, this hypothesis could also account for the success of Erickson's[31] participants. The extra cues present in the Erickson study could be sufficient to inform participants when to turn this signal on and off. In addition, note that this model is consistent with the neuroscience data suggesting that the inhibition between systems is at the output stage. The hyperdirect pathway has no direct effect on processing within the striatum, which has frequently been identified as a key site of procedural learning. Thus, this hyperdirect pathway hypothesis predicts that when declarative memory systems control behavior, procedural learning and memory operates normally but is blocked from (cortical) motor output systems.

Evidence supporting this model comes from studies using the stop-signal task. On a typical stop-signal trial, participants initiate a motor response as quickly as possible when a cue is presented. On some trials, however, a second cue is presented soon after the first and in these cases participants are required to inhibit their response. A variety of evidence implicates the subthalamic nucleus in this task.[52–54] A popular model is that the second cue generates a "stop signal" in cortex that is rapidly transmitted to the GPi via the hyperdirect pathway, where it cancels out the "go signal" being sent through the striatum. When declarative memory is controlling behavior, Ashby and Crossley proposed that a similar stop signal may be used to inhibit a potentially competing response signal generated by the procedural memory system.[30]

This hypothesis might also account for results from a recent study that examined category learning in older adults, and another recent study of category learning in frontal lesion patients.[55,56] Maddox et al. found an age-related deficit in both RB and II category learning.[55] When the analyses focused only on participants who used the task appropriate strategy, however, the age-related RB deficit disappeared whereas the II deficit remained. For this group of individuals, the II deficit was due to less consistent application of the task appropriate strategy by older adults, and over the course of learning these older adults shifted from an explicit hypothesis-testing strategy to the task-appropriate strategy later in learning. These data support a two-component model of II category learning that includes a striatal component that mediates procedural learning, and a frontal-cortical component

(possibly via the hyperdirect pathway) that mediates the transition from hypothesis-testing to procedural-learning strategies.

Schnyer et al. also found RB and II deficits for frontal lesion patients.[56] A more careful lesion analysis pointed to damage in a fairly circumscribed region of ventral medial prefrontal cortex as common to the impaired group of patients, whereas those patients without ventral medial PFC damage mostly performed normally. These results provide further evidence that the ventral medial PFC is critically important for the ability to monitor and integrate feedback in order to select, maintain, and switch strategies in the interest of obtaining optimal performance. Taken together, these data implicate frontal brain regions in both RB and II category learning.

One intriguing possibility is that the primary role of the ventral medial PFC in II tasks may be to control the hyperdirect pathway. According to this account, damage to this PFC region would disrupt the normal transition from RB to II strategies. Schnyer et al. reported that ventromedial PFC patients were impaired in II tasks because they were more likely to use explicit rules. This result suggests that the default state of the hyperdirect projection from frontal cortex to the subthalamic nucleus may be "on" and that one role of the ventral medial PFC may be to switch this excitatory projection off. If so, then as Schnyer et al. reported, damage to the ventral medial PFC would make it more difficult for participants to switch away from explicit RB strategies. Obviously, much more work is needed to test this hypothesis.

Categorization versus Category Representation

One question where significant progress has been made in recent years is whether there are different neural networks for categorization and category representation. Interest in this question arose with reports of a variety of category-specific agnosias that can result following lesions to certain areas of visual cortex.[57,58] The most widely known of such deficits, which occur with human faces (i.e., prosopagnosia), are associated with lesions to the fusiform gyrus in inferotemporal cortex. Subsequently, it was discovered that long periods of practice with a category cause well-documented changes in how visual cortex responds to exemplars of that category.[59,60] For example, Gauthier, Tarr, Anderson, Skudlarski, and Gore reported that 7 hours of training with categories of novel 3-dimensional objects called greebles recruited a greeble-sensitive area in fusiform gyrus.[61] Similar results were later reported for other novel categories and other regions of visual cortex.[62,63]

These results all seem to suggest that much categorization might be mediated within visual cortex. Despite these results, however, there is more recent evidence that the primary role of visual cortex might not be to categorize, but rather to create the visual representation that other brain areas categorize. Categorization is a behavior that requires the subject to take some action, and therefore categorization requires linking a percept to an action. There is good reason to believe that this linkage is not encoded in visual cortex.[64] First, several studies have reported that following categorization training, cells in inferotemporal cortex showed enhanced sensitivity to diagnostic features compared with features that were irrelevant to the categorization judgment.[65,66] Such changes are consistent with the widely held view that category learning is often associated with changes in the allocation of perceptual attention.[67] Second and most critical are the studies showing that categorization training did not make inferotemporal cortex neurons more likely to respond to other stimuli in the same category, or less likely to respond to stimuli belonging to the contrasting category.[65,68–72] Third, Rolls, Judge, and Sanghera showed that the firing properties of cells in inferotemporal cortex did not change when the motor responses associated with category

membership were switched (i.e., from "approach" to "avoid" and vice versa).[73] For these reasons, the best evidence seems to suggest that inferotemporal cortex does not mediate the learning of new categories. Even so, this visual association area is crucial to the categorization process because it appears to encode a high-level representation of the visual stimulus.

Although the standard view has long been that object recognition is mediated primarily by the ventral visual stream, more recent evidence suggests that the dorsal stream is also proficient at object recognition.[74,75] Inferotemporal cortex projects heavily to the body and tail of the caudate nucleus, whereas much of parietal cortex projects to the putamen. Thus, most of the striatum has access to high-level object representations.

Automaticity

A typical adult makes hundreds of categorization judgments every day. Almost all of these are automatic. When we sit in a chair or pick up a book we are making an automatic categorization judgment. Although adults sometimes make categorization decisions that are not automatic, categorization decisions based on newly acquired knowledge are probably far less common than categorization decisions that are made automatically. Despite this imbalance, initial category learning has been investigated much more extensively than categorization automaticity. There are both practical and theoretical reasons that the published literature has focused on initial learning.

One practical reason is that studies of automaticity require more time, patience, and resources than studies of initial learning (where meaningful data are available from the first trial). For example, to study the neural basis of automaticity Muhammad, Wallis, and Miller had monkeys practice the same RB classification task almost daily for more than a year.[76]

A theoretical challenge when studying automaticity is to identify a point in training at which the categorization behavior has become automatic. This is a difficult problem because many different criteria for identifying automaticity have been proposed and none of these are widely accepted as definitive. Perhaps the most influential criteria were proposed by Schneider and Shiffrin.[77–79] For example, they proposed that a behavior should be considered automatic if it can be executed successfully while the participant is simultaneously engaged in some other secondary task. Another criterion they popularized is that a behavior should be considered automatic if it becomes behaviorally inflexible. For example, if switching the location of the response keys interferes with the expression of the behavior, then it should be considered automatic. These criteria are especially problematic if applied to categorization. For example, as mentioned above, several studies have reported that a dual task requiring working memory massively interferes with initial RB category learning but not with II category learning.[13,14,80] Similarly, several studies have reported that switching the locations of the response keys interferes with initial II performance but not with initial RB performance.[10,11,81,82] Therefore, by the Shiffrin and Schneider criteria II categorization is automatic after the first training session. Such a conclusion is incompatible with intuitive notions of automaticity, because accuracy in II tasks requires several thousand trials to asymptote. [83]

Behavioral Results

As humans gain practice in virtually any skill, they naturally become faster and more accurate. In many laboratory studies of automaticity, asymptotic accuracy is perfect, and as a result the primary focus is on response time (RT). The most widely replicated and best known empirical result in this area is that mean RT decreases as a power function of the amount of practice. Among many other examples, a power-function speedup has been

reported for skills as diverse as cigar rolling and proving geometry theorems.[84,85] Not surprisingly, Nosofsky and Palmeri reported that RT improvements in an II task that used categories of color patches also followed the power law.[86]

A more subtle question is whether there are multiple systems for automatic categorization in the same way that there are multiple systems for initial category learning. In the only known study to address this question Helie, Waldschmidt, and Ashby compared the performance of participants who each had more than 11,000 trials of training either with one of two RB tasks (one simple, one complex) or with an II task.[83] Qualitative differences were evident in the behavioral data up through the third of 20+ experimental sessions, but after session 4 or so, there were no more behavioral differences among any of the conditions. As mentioned above, during initial learning, a dual task interferes with RB learning but not II learning, whereas switching the response keys induces the opposite pattern of results. Helie et al. however, reported that after 20+ sessions of practice both of these differences disappear – that is, RB and II tasks both show no dual-task interference and they both show a response-key switch interference.[78] Thus, although much more work is needed on this question, at present there is no behavioral evidence for separate RB and II automatic categorization systems. Instead, the behavioral data suggest two possibilities. One is that after extensive training, RB and II categorization are both mediated by the same procedural system that mediates initial II category learning. A second possibility is that automatic performance in the two tasks is mediated by some new, but common neural network.

Cognitive Neuroscience Results

Interest in the neural basis of automaticity has a long history, dating back at least to Sherrington (1906), who argued that long periods of practice gradually make skills reflexive.[87] These ideas led to the theory that dominated the 20th century: Novel behaviors require attention and flexible thinking and therefore are dependent on cortex, whereas automatic behaviors require neither of these and so are not mediated primarily by cortex. Instead, it has long been assumed that automatic behaviors are primarily mediated by subcortical structures. For example, in his classic and influential article entitled “In search of the engram,” Lashley (1950) wrote that “it has been widely held that although memory traces are at first formed in the cerebral cortex, they are finally reduced or transferred by long practice to subcortical levels” (p. 466).[88]

Striatum—As mentioned above, perhaps the most ubiquitous cognitive neuroscience finding about category learning is the important role that the striatum seems to play.[22,23] Thus, a natural question to ask is whether the striatum also participates in automatic categorization. Based on their different inputs and outputs, the striatum is often subdivided into associative and sensorimotor regions. Roughly speaking, the associative striatum, which includes all of the caudate nucleus and the anterior putamen, receives input from sensory association areas in the temporal lobes and from prefrontal cortex and sends projections primarily to prefrontal cortex (i.e., via the medial dorsal and ventral anterior thalamic nuclei). By contrast, the sensorimotor striatum, which includes all of the putamen except its most anterior portion, receives input from the parietal lobes and from motor and premotor cortex and sends projections primarily to premotor and motor cortex (i.e., via the ventral lateral thalamic nucleus).

More recent evidence suggests that associative and sensorimotor regions of the striatum may play different roles in learning and automaticity.[89] Several studies have reported that the associative striatum is active during initial skill learning and that its activity decreases with extended training.[90–93] In contrast, the sensorimotor striatum often shows the opposite pattern – that is, evidence suggests that the sensorimotor striatum might be needed for the

performance of automatic behaviors and/or for the transition from initial learning to automaticity.[91,94,95] Largely because of these results, there have been several proposals that the development of automaticity involves a gradual transfer of control from the associative to the sensorimotor striatum,[96–98] although the exact details of how such a transfer is mediated have not been described.

It is also important to note that some results seem inconsistent with the hypothesis that the sensorimotor striatum mediates automatic responding. First, using fMRI, Helie, Roeder, and Ashby examined changes in RB categorization at four separate time points as participants practiced on the same category structures for more than 10,000 trials.[99] Striatal activation increased with practice, but the correlation between striatal activation and behavioral performance decreased to zero by the end of training. Thus, this study found no evidence that the striatum contributed to automatic RB categorization. Note that this result also suggests that automatic RB categorization is not mediated by the same procedural system that is critical for early II learning, which as mentioned earlier, is thought to depend heavily on the striatum.

Second, temporary inactivations of sensorimotor regions of the internal segment of the globus pallidus (via injections of the GABA agonist muscimol) should prevent the sensorimotor striatum from influencing cortex because this area is a relay between the striatum and thalamus. Thus, if the striatum mediates the expression of automatic behaviors, then such inactivations should disrupt highly practiced actions. In contrast to this prediction, however, Desmurget and Turner reported that such inactivations did not prevent monkeys from fluidly executing highly practiced motor sequences.[100] As a result, they concluded that circuits through the sensorimotor striatum do not contribute “to motor sequencing or the storage of overlearned serial skills” (p. 7685). Third, several studies have shown that disconnecting the bird homologue of the basal ganglia completely blocks new song learning, but has little effect on the expression of well-learned songs.[101]

The Role of Dopamine—As mentioned earlier, the evidence is good that dopamine plays an important role in early category learning. Even so, other evidence suggests it plays a diminishing role in the expression of automatic behaviors.[102] For example, some human subjects with Parkinson’s disease are able to emit an automatic motor response when presented with a familiar visual cue (e.g., kicking a ball), despite difficulties in initiating novel voluntary movements.[103] As another example, blockade of dopamine D1 receptors in rats strongly disrupts a simple Pavlovian approach response to a sensory cue during early stages of training, but has little or no disruptive effect if extended training is given before the dopamine antagonist is administered.[104,105]

Another important question is whether the role of dopamine is the same in the striatum and cortex. Both regions receive prominent dopamine projections from midbrain dopamine cells. As mentioned above, striatal dopamine is widely thought to serve as the training signal for reinforcement learning at cortical-striatal synapses. This raises the obvious question of the role that cortex might play in reinforcement-mediated category learning. A necessary feature of any reinforcement training signal is high temporal resolution. If the first response is correct then dopamine must be released into the relevant synapses quickly, before the critical traces disappear. But after the correct synapses have been strengthened, it is also essential that excess dopamine be quickly cleared from the synapse. If it is not, and the next response is an error, then the residual dopamine will strengthen inappropriate synapses – namely, those responsible for producing the incorrect response. This would undo the beneficial learning that occurred following correct responses, and thereby prevent skill learning.

Dopamine re-uptake is exceptionally fast within the striatum.[106] In contrast, in frontal cortex, because of low concentrations of the dopamine reuptake molecule DAT, it takes much longer to clear dopamine from synapses.[107–109] For example, the delivery of a single food pellet to a hungry rat elevates dopamine levels in prefrontal cortex for approximately 30 minutes.[110] Ashby, Ennis, and Spiering argued that this poor temporal resolution effectively rules out dopamine as a trial-by-trial reinforcement training signal in cortex.[111] Instead, although dopamine may facilitate cortical LTP, there is much evidence that synaptic plasticity at cortical-cortical synapses follows classical two-factor Hebbian learning rules.[112]

Transfer to Cortex—Many sensory association areas of cortex project directly into premotor cortex. Ashby et al. proposed that these cortical networks, by themselves, are incapable of skill learning because of the absence of reinforcement learning at cortical-cortical synapses.[111] Instead, they proposed that via reinforcement learning, a subcortical path through the striatum learns to activate the correct post-synaptic target in premotor cortex, which allows the appropriate cortical-cortical synapses in the premotor cortex to then be strengthened via Hebbian learning (because the product of pre- and post-synaptic activations will be greatest at the correct synapse). In this way, control is gradually passed from the subcortical path through the basal ganglia to the faster cortical-cortical path. Thus, according to this model, the development of automaticity is a gradual process via which control is passed from subcortical pathways through the basal ganglia to purely cortical networks that connect sensory association areas of cortex with premotor cortex. In other words, rather than to serve as a long-term store of procedural knowledge, a primary function of the basal ganglia may be to train cortical-cortical representations that mediate automaticity. Note that this theory accounts for results showing that automatic behaviors are striatal- and dopamine-independent.

Learning Characteristics of Each System

As mentioned earlier, the overriding research goal in the first generation was to establish that multiple systems exist, rather than to learn as much as possible about any one system. More recently, some studies have focused more exclusively on a single system, with the goal of developing more detailed models of each system. This section reviews some of this work.

Separate Stages of Explicit and Procedural Category Learning

The category shift literature suggests that RB category learning includes two separate stages: a stimulus-to-label stage that associates stimuli and category labels, and a label-to-response stage that associates category labels and responses.[113–117] In a recent study, Maddox et al. reported evidence that II classification is also mediated by two learning stages.[81] After training participants in an II task, Maddox et al. either changed the mappings between stimuli and category labels or between the category labels and the response locations. Both manipulations change the mappings between stimuli and responses, so if there is a single stage of learning then both manipulations should cause similar deficits. In contrast to this prediction, breaking the association between stimulus and category label caused more interference and led to greater recovery than breaking the association between category label and response location.

COVIS[1] postulates no learning after the cortical-striatal synapses, so it predicts no qualitative performance difference between the stimulus-to-label and label-to-response conditions. Both manipulations should require cortical-striatal relearning. For this reason, COVIS, in its current form, is not consistent with the Maddox et al.[81] results. COVIS might be extended, however, to include a second stage of learning that associates a category label

with a specific response location. Logically, such learning must be downstream from the site of category label learning, which suggests that plausible sites of response learning could be at synapses in the internal segment of the globus pallidus, the ventral anterior or ventral lateral nuclei of the thalamus, or within premotor cortex. Each of these brain regions has been implicated in procedural learning and thus represents plausible loci of such learning.[92]

Switching in the Explicit System

RB category learning depends heavily on executive function. For example, one classic neuropsychological assessment of executive function is the Wisconsin Card Sorting Test, which requires participants to learn a series of RB categorization tasks.[118] The recent literature on the cognitive neuroscience of executive function is immense and a review is well beyond the scope of this article. Interested readers should consult Banich[119] or Jurado and Rosselli[120]. This section touches on some recent findings that are relevant to rule switching in the explicit system.

Theoretical accounts of RB category learning postulate a number of separate processes. For example, COVIS assumes separate rule selection, maintenance, and switching operations.[1] Neuroimaging and neuropsychological results have provided evidence for multiple processes in RB category learning,[21,121–123] and recent work has begun focusing on these sub-processes in more detail. For example, Chiu and Yantis reported fMRI evidence that the mechanism that mediates rule switching is the same as the mechanism used in other types of attentional switching.[124]

Neuropsychological and pharmacological studies have long implicated dopamine in rule switching,[125–128] but recent work also suggests a role for norepinephrine.[129,130] For example, pharmacological manipulations of prefrontal norepinephrine alter rule switching performance in rats.[131,132] Although the mechanism by which norepinephrine supports rule switching is unknown, one possibility is that a lack of positive reinforcement following the rule switch induces exploratory behaviors (e.g., switching away from the present rule) that depend on the ceruleo-cortical norepinephrine system.[133]

A growing body of research suggests that motivational incentives can also influence executive control processes and thus rule switching in the explicit system. In a recent behavioral study, Maddox, Filoteo, Glass and Markman showed that the global and local motivational incentives in a modified Wisconsin Card Sorting Task systematically accentuated rule switching when they matched (i.e., when both the global and local incentives emphasized either maximizing gains or minimizing losses) and systematically attenuated rule switching when there was a mismatch (i.e., when the global incentive was to maximize gains and the local incentive was to minimize losses or vice versa).[134] In a recent fMRI study, Savine and Braver examined the effects of reward incentives on performance and neural activity in a cued task-switching paradigm.[135] They found smaller task-switching costs on incentive trials relative to no incentive trials. In addition, they found that activation in cognitive control networks tracked fluctuations in incentive value.

Positive and Negative Feedback

Feedback is known to play a different role in RB and II learning. A number of recent studies have explored these differences in detail.

A long history of research has shown that in RB category learning, providing feedback about errors allows faster learning than providing feedback about correct responses.[136,137] Two studies have investigated similar questions in II tasks. First, Ashby and O'Brien reported that II category learning requires both types of feedback.[138] Participants in this study who received feedback either only about (some) correct responses or only about (some) errors

generally used explicit, RB strategies. Second, using a semi-supervised training paradigm in which participants received feedback about both correct responses and errors, but only on some trials, Vandist, De Schryver, and Rosseel reported that they found no evidence of any II learning on no-feedback trials.[139]

Plasticity at cortical-striatal synapses is known to be bi-directional,[140,141] and thus current theory would predict that the relevant cortical-striatal synapses should be modified following feedback after both correct and incorrect responses, but should not change much on trials when no feedback is given (i.e., because striatal dopamine levels should not fluctuate much on no-feedback trials). This latter prediction is consistent with the Vandist et al.[139] results.

The COVIS procedural-learning system included only the direct pathway through the basal ganglia[1]. In this pathway synapses are strengthened only if the following three conditions are present: (1) strong presynaptic activation, (2) strong postsynaptic activation (i.e., strong enough to activate NMDA receptors), and (3) dopamine levels above baseline (which occurs following feedback that indicates a correct response). Thus, on positive feedback (correct response) trials, COVIS assumes that synapses for which all three conditions are met will be strengthened, synapses for which the NMDA activation threshold is not met will be weakened, and synapses for which there is no activation will not change. On negative feedback (error response) trials, on the other hand, COVIS assumes that synapses either will be weakened or will remain unchanged. Thus, there is an asymmetry in the COVIS procedural-learning system which predicts that positive feedback should be more effective than negative feedback (because there is a greater dynamic range for increases in dopamine levels in the striatum than for decreases).

Another theoretical proposal, which more evenly balances the effects of positive and negative feedback, is that the direct pathway mediates GO learning, whereas the indirect pathway mediates NO-GO learning.[142] According to this account, the direct pathway primarily uses positive feedback to learn which response to make when a stimulus is presented, while the indirect pathway primarily uses negative feedback to learn which response not to make. Although the indirect pathway has not been formally incorporated into COVIS, it does appear that an augmentation of this sort would allow COVIS to predict learning in positive-only or negative-only conditions.

COVIS assumes that learning should occur in the positive-only condition, and a version that incorporates the indirect pathway would predict learning in the positive- and negative-only conditions. Contrary to both versions of the model, Ashby and O'Brien found that participants in the positive-only and negative-only feedback conditions generally used explicit, RB strategies. One limitation of these studies, and an area for future research, is that the primary effect of changing the feedback in any way may be to change the system that people learn to rely on in the task. In fact, this may be more important than the change in the amount of learning that occurs within each system. For example, in the Ashby and O'Brien study,[138] participants using a one-dimensional rule could achieve an accuracy of 78%, which is only 8% below the accuracy rate of the optimal II strategy. This 8% difference was large enough to induce full-feedback control participants to adopt an II strategy, but it is difficult to rule out the possibility that removing all positive feedback or all negative feedback led participants to rely more heavily on explicit rules, effectively masking learning effects on the procedural system. Future work should examine category structures for which the difference in accuracy between the optimal II and most accurate one-dimensional rule is larger than 8%.

Future Directions

The era of work on second generation problems is just beginning and many important topics await systematic investigation. First, future work needs to address the nature of retention and generalization processes in each system. It will be important to understand how categorical knowledge is retained over long periods of time, how it is generalized, and how using that knowledge in one way might interfere with or enhance other uses of the same information. Second, desirable difficulties occur when one training protocol leads to poor learning but better long-term retention than other protocols.[143] This phenomenon has been observed in some motor and verbal learning paradigms and should be studied in categorization. Third, more developmental work is needed to determine how the different categorization systems develop and interact in childhood and in normal aging.[55] Fourth, more work is needed to understand the psychopathological implications of multiple systems. One exciting avenue for future research will be to apply a neuro-rehabilitation approach to category learning in which a normally functioning system is used to "bootstrap" learning or performance in a poorly functioning system. These are just a few of the many exciting avenues of research that will be pursued in the coming years.

Conclusions

Many categorization researchers now accept the large and growing body of behavioral, neuropsychological, and neuroimaging evidence that humans have multiple category learning systems that are functionally distinct at both the neural and cognitive levels. Thus, over the past few years the emphasis has begun to shift to the study of second generation problems – that is, questions that begin with the assumption that humans have multiple systems. This article reviewed much of the existing second generation research.

Acknowledgments

This research was supported by NIH Grant P01 NS044393 and by the U.S. Army Research Office through the Institute for Collaborative Biotechnologies under contract DAAD19-03-D-0004 to FGA. Additional support was provided by NIMH grant MH077708 to WTM.

References

1. Ashby FG, Alfonso-Reese LA, Turken AU, Waldron EM. A neuropsychological theory of multiple systems in category learning. Psychological Review. 1998; 105:442–481. [PubMed: 9697427]
2. Knowlton BJ, Mangels JA, Squire LR. A neostriatal habit learning system in humans. Science. 1996; 273:1399–1402. [PubMed: 8703077]
3. Ashby FG, Maddox WT. Human category learning. Annual Review of Psychology. 2005; 56:149–178.
4. Smith EE, Grossman M. Multiple systems of category learning. Neuroscience and Biobehavioral Reviews. 2008; 32:249–264. [PubMed: 17904637]
5. Ashby FG, O'Brien JB. Category learning and multiple memory systems. Trends in Cognitive Sciences. 2005; 2:83–89. [PubMed: 15668101]
6. Casale MB, Ashby FG. A role for the perceptual representation memory system in category learning. Perception & Psychophysics. 2008; 70:983–999. [PubMed: 18717385]
7. Zeithamova D, Maddox WT, Schnyer DM. Dissociable prototype learning systems: Evidence from brain imaging and behavior. Journal of Neuroscience. 2008; 28:13194–13201. [PubMed: 19052210]
8. Maddox WT, Ashby FG, Bohil CJ. Delayed feedback effects on rule-based and information-integration category learning. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2003; 29:650–662.

9. Maddox WT, Ing AD. Delayed feedback disrupts information-integration but not rule-based category learning. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2005; 31:100–107.
10. Ashby FG, Ell SW, Waldron EM. Procedural learning in perceptual categorization. Memory & Cognition. 2003; 31:1114–1125.
11. Maddox WT, Bohil CJ, Ing AD. Evidence for a procedural-learning-based system in perceptual category learning. Psychonomic Bulletin & Review. 2004; 11:945–952. [PubMed: 15732708]
12. Ashby FG, Maddox WT, Bohil CJ. Observational versus feedback training in rule-based and information-integration category learning. Memory & Cognition. 2002; 30:666–677.
13. Waldron EM, Ashby FG. The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. Psychonomic Bulletin & Review. 2001; 8:168–176. [PubMed: 11340863]
14. Zeithamova D, Maddox WT. Dual task interference in perceptual category learning. Memory & Cognition. 2006; 34:387–398.
15. Maddox WT, Ashby FG, Ing AD, Pickering AD. Disrupting feedback processing interferes with rule-based but not information-integration category learning. Memory & Cognition. 2004; 32:582–591.
16. Maddox, WT.; Filoteo, JV. Modeling visual attention and category learning in amnesiacs, striatal-damaged patients, and normal aging. In: Neufeld, RWJ., editor. Advances in clinical-cognitive science: Formal modeling and assessment of processes and symptoms. Washington, DC: American Psychological Association; 2007. p. 113-146.
17. Dayan, P.; Abbott, LF. Theoretical neuroscience: Computational and mathematical modeling of neural systems. Cambridge, MA: MIT Press; 2001.
18. Doya K. Reinforcement learning: Computational theory and biological mechanisms. HFSP Journal. 2007; 1:30–40. [PubMed: 19404458]
19. Schultz W. Getting formal with dopamine and reward. Neuron. 2002; 36:241–263. [PubMed: 12383780]
20. Filoteo, JV.; Maddox, WT. Category learning in Parkinson's disease. In: Sun, Maio-Kun, editor. Research progress in Alzheimer's disease and dementia. Vol. 3. Hauppauge, NY: Nova Science Publishers, Inc; 2007. p. 2-26.
21. Price A, Filoteo JV, Maddox WT. Rule-based category learning in patients with Parkinson's disease. Neuropsychologia. 2009; 47:1213–1226. [PubMed: 19428385]
22. Ashby FG, Ennis JM. The role of the basal ganglia in category learning. The Psychology of Learning and Motivation. 2006; 46:1–36.
23. Seger CA. How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. Neuroscience and Biobehavioral Reviews. 2008; 32:265–278. [PubMed: 17919725]
24. Ashby FG, Waldron EM. On the nature of implicit categorization. Psychonomic Bulletin & Review. 1999; 6:363–378. [PubMed: 12198775]
25. Ashby FG, Ell SW, Valentin V, Casale MB. FROST: A distributed neurocomputational model of working memory maintenance. Journal of Cognitive Neuroscience. 2005; 17:1728–1743. [PubMed: 16269109]
26. Ashby, FG.; Valentin, VV. Multiple systems of perceptual category learning: Theory and cognitive tests. In: Cohen, H.; Lefebvre, C., editors. Categorization in cognitive science. New York: Elsevier; 2005.
27. Janowsky JS, Shimamura AP, Kritchevsky M, Squire LR. Cognitive impairment following frontal lobe damage and its relevance to human amnesia. Behavioral Neuroscience. 1989; 103:548–560. [PubMed: 2736069]
28. Leng NR, Parkin AJ. Double dissociation of frontal dysfunction in organic amnesia. British Journal of Clinical Psychology. 1988; 27:359–362. [PubMed: 3214689]
29. Hopkins RO, Myers EC, Shohamy D, Grossman S, Gluck M. Impaired probabilistic category learning in hypoxic subjects with hippocampal damage. Neuropsychologia. 2004; 42:524–535. [PubMed: 14728924]

30. Ashby FG, Crossley MJ. Interactions between declarative and procedural-learning categorization systems. Neurobiology of Learning and Memory. 2010; 94:1–12. [PubMed: 20304078]
31. Erickson MA. Executive attention and task switching in category learning: Evidence for stimulus-dependent representation. Memory & Cognition. 2008; 36:749–761.
32. Cincotta CM, Seger CA. Dissociation between striatal regions while learning to categorize via observation and via feedback. Journal of Cognitive Neuroscience. 2007; 19:249–265. [PubMed: 17280514]
33. Moody TD, Bookheimer ZV, Knowlton BJ. An implicit learning task activates medial temporal lobe in patients with Parkinson's disease. Behavioral Neuroscience. 2004; 118:438–442. [PubMed: 15113271]
34. Nomura EM, Maddox WT, Filoteo JV, Ing AD, Gitelman DR, Parrish TB, et al. Neural correlates of rule-based and information-integration visual category learning. Cerebral Cortex. 2007; 17:37–43. [PubMed: 16436685]
35. Poldrack RA, Prabhakaran V, Seger CA, Gabrieli JDE. Striatal activation during acquisition of a cognitive skill. Neuropsychology. 1999; 13:564–574. [PubMed: 10527065]
36. Poldrack RA, Clark J, Pare-Blagoev EJ, Shohamy D, Moyano JC, Myers C, Gluck MA. Interactive memory systems in the human brain. Nature. 2001; 414:546–550. [PubMed: 11734855]
37. Seger CA, Cincotta CM. Dynamics of frontal, striatal, and hippocampal systems in rule learning. Cerebral Cortex. 2006; 16:1546–1555. [PubMed: 16373455]
38. Dagher A, Owen AM, Boecker H, Brooks DJ. The role of the striatum and hippocampus in planning: A PET activation study in Parkinson's disease. Brain. 2001; 124:1020–1032. [PubMed: 11335704]
39. Jenkins IH, Brooks DJ, Nixon PD, Frackowiak RSJ, Passingham RE. Motor sequence learning: A study with positron emission tomography. Journal of Neuroscience. 1994; 14:3775–3790. [PubMed: 8207487]
40. Poldrack RA, Gabrieli JD. Characterizing the neural mechanisms of skill learning and repetition priming: Evidence from mirror reading. Brain. 2001; 124:67–82. [PubMed: 11133788]
41. Mitchell JA, Hall G. Caudate-putamen lesions in the rat may impair or potentiate maze learning depending upon availability of stimulus cues and relevance of response cues. Quarterly Journal of Experimental Psychology. 1988; 40:243–258. [PubMed: 3175038]
42. O'Keefe, J.; Nadel, L. The hippocampus as a cognitive map. Oxford, UK: Oxford University Press; 1978.
43. Schroeder JA, Wingard J, Packard MG. Post-training reversible inactivation of the dorsal hippocampus reveals interference between multiple memory systems. Hippocampus. 2002; 12:280–284. [PubMed: 12000124]
44. Packard MG, McGaugh JL. Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. Neurobiology of Learning and Memory. 1996; 65:65–72. [PubMed: 8673408]
45. Foerde K, Knowlton BJ, Poldrack RA. Modulation of competing memory systems by distraction. Proceedings of the National Academy of Sciences of the USA. 2006; 103:11778–11783. [PubMed: 16868087]
46. Foerde K, Poldrack RA, Knowlton BJ. Secondary-task effects on classification learning. Memory & Cognition. 2007; 35:864–874.
47. Nambu A, Tokuno H, Hamada I, Kita H, Imanishi M, Akazawa T, Ikeuchi Y, Hasegawa N. Excitatory cortical inputs to pallidal neurons via the subthalamic nucleus in the monkey. Journal of Neurophysiology. 2000; 84:289–300. [PubMed: 10899204]
48. Nambu A, Tokuno H, Takada M. Functional significance of the cortico-subthalamo-pallidal 'hyperdirect' pathway. Neuroscience Research. 2002; 43:111–117. [PubMed: 12067746]
49. Hikosaka O, Isoda M. Switching from automatic to controlled behavior: cortico-basal ganglia mechanisms. Trends in Cognitive Sciences. 2010; 14:154–161. [PubMed: 20181509]
50. Joel D, Weiner I. The connections of the primate subthalamic nucleus: Indirect pathways and the open-interconnected scheme of basal ganglia– thalamocortical circuitry. Brain Research Reviews. 1997; 23:62–78. [PubMed: 9063587]

51. Parent A, Hazrati LN. Functional anatomy of the basal ganglia. II. The place of the subthalamic nucleus and external pallidum in basal ganglia circuitry. Brain Research Reviews. 1995; 20:128–154. [PubMed: 7711765]
52. Aron AR, Behrens TE, Smith S, Frank MJ, Poldrack RA. Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. Journal of Neuroscience. 2007; 27:3743–3752. [PubMed: 17409238]
53. Aron AR, Poldrack RA. Cortical and subcortical contributions to stop signal response inhibition: Role of the subthalamic nucleus. Journal of Neuroscience. 2006; 26:2424–2433. [PubMed: 16510720]
54. Mostofsky SH, Simmonds DJ. Response inhibition and response selection: Two sides of the same coin. Journal of Cognitive Neuroscience. 2008; 97:751–761. [PubMed: 18201122]
55. Maddox WT, Pacheco J, Reeves M, Zhu B, Schnyer DM. Rule-based and information-integration category learning in normal aging. Neuropsychologia. 2010; 48:2998–3008. [PubMed: 20547171]
56. Schnyer DM, Maddox WT, Ell S, Davis S, Pacheco J, Verfaellie M. Prefrontal contributions to rule-based and information-integration category learning. Neuropsychologia. 2009; 47:2995–3006. [PubMed: 19643119]
57. Damasio AR, Damasio H, Van Hoesen GW. Prosopagnosia: Anatomic basis and behavioral mechanisms. Neurology. 1982; 32:331–341. [PubMed: 7199655]
58. Warrington EK, Shallice T. Category specific semantic impairments. Brain. 1984; 107:829–854. [PubMed: 6206910]
59. Humphreys GW, Forde EME. Hierarchies, similarity, and interactivity in object recognition: "Category-specific" neuropsychological deficits. Behavioral and Brain Sciences. 2001; 24:453–509. [PubMed: 11682799]
60. Joseph JE. Functional neuroimaging studies of category specificity in object recognition: A critical review and meta-analysis. Cognitive, Affective, & Behavioral Neuroscience. 2001; 1:119–136.
61. Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC. Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. Nature Neuroscience. 1999; 2:568–573.
62. Op de Beeck HP, Baker CI, DiCarlo JJ, Kanwisher NG. Discrimination training alters object representations in human extrastriate cortex. Journal of Neuroscience. 2006; 26:13025–13036. [PubMed: 17167092]
63. Weisberg J, van Turennout M, Martin A. A neural system for learning about object function. Cerebral Cortex. 2007; 17:513–521. [PubMed: 16581980]
64. Freedman DJ, Miller EK. Neural mechanisms of visual categorization: Insights from neurophysiology. Neuroscience and Biobehavioral Reviews. 2008; 32:311–329. [PubMed: 17950874]
65. Sigala N. Visual categorization and the inferior temporal cortex. Behavioural Brain Research. 2004; 149:1–7. [PubMed: 14739004]
66. Sigala N, Logothetis NK. Visual categorization shapes feature selectitivity in the primate temporal cortex. Nature. 2002 January 17.415:318–320. [PubMed: 11797008]
67. Nosofsky RM. Attention, similarity, and the identification-categorization relationship. Journal of Experimental Psychology: General. 1986; 115:39–57. [PubMed: 2937873]
68. Freedman DJ, Riesenhuber M, Poggio T, Miller EK. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. The Journal of Neuroscience. 2003; 23:5235–5246. [PubMed: 12832548]
69. Jiang X, Bradley E, Rini RA, Zeffiro T, VanMeter J, Riesenhuber M. Categorization training results in shape- and category-selective human neural plasticity. Neuron. 2007; 53:891–903. [PubMed: 17359923]
70. Op de Beeck H, Wagemans J, Vogels R. Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. Nature Neuroscience. 2001; 4:1244–1252.
71. Thomas E, Van Hulle MM, Vogels R. Encoding of categories by noncategory-specific neurons in the inferior temporal cortex. Journal of Cognitive Neuroscience. 2001; 13:190–200. [PubMed: 11244545]

72. Vogels R. Complex visual images by rhesus monkeys. Part 2: Single-cell study. European Journal of Neuroscience. 1999; 11:1239–1255. [PubMed: 10103119]
73. Rolls ET, Judge SJ, Sanghera M. Activity of neurons in the inferotemporal cortex of the alert monkey. Brain Research. 1977; 130:229–238. [PubMed: 406968]
74. Konen CS, Kastner S. Two hierarchically organized neural systems for object information in human visual cortex. Nature Neuroscience. 2008; 11:224–231.
75. Sereno AB, Amador SC. Attention and memory related responses of neurons in the lateral intraparietal area during shape and spatial delayed match-to-sample tasks. Journal of Neurophysiology. 2006; 95:1078–1098. [PubMed: 16221750]
76. Muhammad R, Wallis JD, Miller EK. A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum. Journal of Cognitive Neuroscience. 2006; 18:1–16. [PubMed: 16417678]
77. Schneider W, Shiffrin RM. Controlled and automatic human information processing: I. Detection, search, and attention. Psychological Review. 1977; 84:1–66.
78. Shiffrin RM, Schneider W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review. 1977; 84:127–190.
79. Schneider W, Chein JM. Controlled and automatic processing: Behavior, theory, and biological mechanism. Cognitive Science. 2003; 27:525–559.
80. Zeithamova D, Maddox WT. The role of visuo-spatial and verbal working memory in perceptual category learning. Memory & Cognition. 2007; 35:1380–1398.
81. Maddox WT, Glass BD, O'Brien JB, Filoteo JV, Ashby FG. Category label and response location shifts in category learning. Psychological Research. 2010; 74:219–236. [PubMed: 19471959]
82. Spiering BJ, Ashby FG. Response processes in information-integration category learning. Neurobiology of Learning and Memory. 2008; 90:330–338. [PubMed: 18550397]
83. Hélie S, Waldschmidt JG, Ashby FG. Automaticity in rule-based and information-integration categorization. Attention, Perception, & Psychophysics. 2010; 72:1013–1031.
84. Crossman ERFW. A theory of the acquisition of speed-skill. Ergonomics. 1959; 2:153–166.
85. Neves, DM.; Anderson, JR. Knowledge compilation: Mechanisms for the automatization of cognitive skills. In: Anderson, JR., editor. Cognitive skills and their acquisition. Hillsdale, NJ: Erlbaum; 1981. p. 57-84.
86. Nosofsky RM, Palmeri TJ. An exemplar-based random walk model of speeded classification. Psychological Review. 1997; 104:266–300. [PubMed: 9127583]
87. Sherrington CS. Observations on the scratch reflex in the spinal dog. Journal of Physiology. 1906; 34:1–50. [PubMed: 16992835]
88. Lashley, KS. Society of Experimental Biology Symposium. Vol. 4. Cambridge, England: Cambridge University Press; 1950. In search of the engram; p. 454-480.
89. Ashby FG, Turner BO, Horvitz JC. Cortical and basal ganglia contributions to habit learning and automaticity. Trends in Cognitive Sciences. 2010; 14:191–232. [PubMed: 20363177]
90. Lehéricy S, Benali H, Van de Moortele PF, Pélégrini-Issac M, Waechter T, Ugurbil K, Doyon J. Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. Proceedings of the National Academy of Sciences, USA. 2005; 102:12566–12571.
91. Miyachi S, Hikosaka O, Lu X. Differential activation of monkey striatal neurons in the early and late stages of procedural learning. Experimental Brain Research. 2002; 146:122–126.
92. Poldrack RA, Sabb FW, Foerde K, Tom SM, Asarnow RF, Bookheimer SY, Knowlton BJ. The neural correlates of motor skill automaticity. Journal of Neuroscience. 2005; 25:5356–5364. [PubMed: 15930384]
93. Wu T, Kansaku K, Hallett M. How self-initiated memorized movements become automatic: A functional MRI study. Journal of Neurophysiology. 2004; 91:1690–1698. [PubMed: 14645385]
94. Miyachi S, Hikosaka O, Miyashita K, Karaki Z, Rand MK. Differential roles of monkey striatum in learning of sequential hand movement. Experimental Brain Research. 1997; 115:1–5.
95. Yin HH, Knowlton BJ, Balleine BW. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. European Journal of Neuroscience. 2004; 19:181–189. [PubMed: 14750976]

96. Belin D, Jonkman S, Dickinson A, Robbins TW, Everitt BJ. Parallel and interactive learning processes within the basal ganglia: Relevance for the understanding of addiction. Behavioral Brain Research. 2009; 199:89–102.

97. Costa RM. Plastic corticostriatal circuits for action learning: What's dopamine got to do with it? Annals of the New York Academy of Sciences. 2007; 1104:172–191. [PubMed: 17435119]

98. Yin HH, Knowlton BJ. The role of the basal ganglia in habit formation. Nature Reviews Neuroscience. 2006; 7:464–476.

99. Hélie S, Roeder JL, Ashby FG. Evidence for cortical automaticity in rule-based categorization. Journal of Neuroscience. 2010; 30:14225–14234. [PubMed: 20962243]

100. Desmurget M, Turner RS. Motor sequences and the basal ganglia: Kinematics, not habits. Journal of Neuroscience. 2010; 30:7685–7690. [PubMed: 20519543]

101. Doupe AJ, Perkel DJ, Reiner A, Stern EA. Birdbrains could teach basal ganglia research a new song. Trends in Neurosciences. 2005; 28:353–363. [PubMed: 15935486]

102. Horvitz JC, Choi WY, Morvan C, Eyny Y, Balsam PD. A "good parent" function of dopamine: Transient modulation of learning and performance during early stages of training. Annals of the New York Academy of Sciences. 2007; 1104:270–288. [PubMed: 17360799]

103. Asmus F, Huber H, Gasser T, Schöls L. Kick and rush: Paradoxical kinesia in Parkinson disease. Neurology. 2008; 71:695. [PubMed: 18725599]

104. Bespalov AY, Harich S, Jongen-Rêlo AL, van Gaalen MM, Gross G. AMPA receptor antagonists reverse effects of extended habit training on signaled food approach responding in rats. Psychopharmacology. 2007; 195:11–18. [PubMed: 17634927]

105. Choi WY, Balsam PD, Horvitz JC. Extended habit training reduces dopamine mediation of appetitive response expression. Journal of Neuroscience. 2005; 25:6729–6733. [PubMed: 16033882]

106. Cragg SJ, Rice ME, Greenfield SA. Heterogeneity of electrically evoked dopamine release and reuptake in substantia nigra, ventral tegmental area, and striatum. Journal of Neurophysiology. 1997; 77:863–873. [PubMed: 9065855]

107. Seamans JK, Yang CR. The principal features and mechanisms of dopamine modulation in the prefrontal cortex. Progress in Neurobiology. 2004; 74:1–57. [PubMed: 15381316]

108. Seamans, JK.; Robbins, TW. Dopamine modulation of prefrontal cortex and cognitive function. In: Neve, KA., editor. The dopamine receptors. 2. New York: Springer; 2009. p. 373-398.

109. Tzschentke TM. Pharmacology and behavioral pharmacology of the mesocortical dopamine system. Progress in Neurobiology. 2001; 63:241–320. [PubMed: 11115727]

110. Feenstra MG, Botterblom MH. Rapid sampling of extracellular dopamine in the rat prefrontal cortex during food consumption, handling and exposure to novelty. Brain Research. 1996; 742:17–24. [PubMed: 9117391]

111. Ashby FG, Ennis JM, Spiering BJ. A neurobiological theory of automaticity in perceptual categorization. Psychological Review. 2007; 114:632–656. [PubMed: 17638499]

112. Feldman DE. Synaptic mechanisms for plasticity in neocortex. Annual Review of Neuroscience. 2009; 32:33–55.

113. Kendler HH, Kendler TS. Vertical and horizontal processes in problem solving. Psychological Review. 1962; 69:1–16. [PubMed: 14455127]

114. Kendler, HH.; Kendler, TS. Mediation and conceptual behavior. In: Spence, KWSJT., editor. The Psychology of learning and motivation. Vol. 2. New York: Academic Press; 1968. p. 197-244.

115. Kruschke JK. Dimensional relevance shifts in category learning. Connection Science. 1996; 8:225–247.

116. Sanders B. Factors affecting reversal and nonreversal shifts in rats and children. Journal of Coparative & Physiological Psychology. 1971; 74:192–202.

117. Wills AJ, Noury M, Moberly NJ, Newport M. Formation of category representations. Memory & Cognition. 2006; 34:17–27.

118. Heaton, RK. A manual for the Wisconsin Card Sorting Test. Odessa, FL: Psychological Assessment Resources; 1981.

119. Banich MT. Executive function: The search for an integrated account. Current Directions in Psychological Science. 2009; 18:89–94.
120. Jurado MB, Rosselli M. The elusive nature of executive functions: A review of our current understanding. Neuropsychology Review. 2007; 17:213–233. [PubMed: 17786559]
121. Kehagia AA, Cools R, Barker RA, Robbins TW. Switching between abstract rules reflects disease severity but not dopaminergic status in Parkinson's disease. Neuropsychologia. 2009; 47:1117–1127. [PubMed: 19166864]
122. Monchi O, Petrides M, Petre V, Worsley K, Dagher A. Wisconsin Card Sorting Revisited: Distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. Journal of Neuroscience. 2001; 21:7733–7741. [PubMed: 11567063]
123. Tachibana K, Suzuki K, Mori E, Miura N, Kawashima R, Horie K, Sato S, Tanji J, Mushiake H. Neural activity in the human brain signals logical rule identification. Journal of Neurophysiology. 2009; 102:1526–1537. [PubMed: 19553481]
124. Chiu YC, Yantis S. A domain-independent source of cognitive control for task sets: shifting spatial attention and switching categorization rules. Journal of Neuroscience. 2009; 29:3930–3938. [PubMed: 19321789]
125. Cools R, Lewis SJ, Clark L, Barker RA, Robbins TW. L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. Neuropsychopharmacology. 2007; 32:180–189. [PubMed: 16841074]
126. Frank MJ, O'Reilly RC. A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. Behavioral Neuroscience. 2006; 120:497–517. [PubMed: 16768602]
127. Monchi O, Petrides M, Doyon J, Postuma RB, Worsley K, Dagher A. Neural bases of set-shifting deficits in Parkinson's disease. Journal of Neuroscience. 2004; 24:702–710. [PubMed: 14736856]
128. Price AL. Explicit category learning in Parkinson's disease: deficits related to impaired rule generation and selection processes. Neuropsychology. 2006; 20:249–257. [PubMed: 16594785]
129. Lapiz MD, Bondi CO, Morilak DA. Chronic treatment with desipramine improves cognitive performance of rats in an attentional set-shifting test. Neuropsychopharmacology. 2007; 32:1000–1010. [PubMed: 17077810]
130. Newman LA, Darling J, McGaughy J. Atomoxetine reverses attentional deficits produced by noradrenergic deafferentation of medial prefrontal cortex. Psychopharmacology. 2008; 200:39–50. [PubMed: 18568443]
131. Lapiz MD, Morilak DA. Noradrenergic modulation of cognitive function in rat medial prefrontal cortex as measured by attentional set shifting capability. Neuroscience. 2006; 137:1039–1049. [PubMed: 16298081]
132. Tait DS, Brown VJ, Farovik A, Theobald DE, Dalley JW, Robbins TW. Lesions of the dorsal noradrenergic bundle impair attentional set-shifting in the rat. European Journal of Neuroscience. 2007; 25:3719–3724. [PubMed: 17610591]
133. Aston-Jones G, Rajkowski J, Cohen J. Role of locus coeruleus in attention and behavioral flexibility. Biological Psychiatry. 1999; 46:1309–1320. [PubMed: 10560036]
134. Maddox WT, Filoteo JV, Glass BD, Markman AB. Regulatory match effects on a modified Wisconsin Card Sort Task. Journal of the International Neuropsychological Society. 2010; 16:352–359. [PubMed: 20128935]
135. Savine AC, Braver TS. Motivated cognitive control: Reward incentives modulate preparatory neural activity during task-switching. Journal of Neuroscience. 2010; 30:10294–10305. [PubMed: 20685974]
136. Buss AH, Buss EH. The effect of verbal reinforcement combinations on conceptual learning. Journal of Experimental Psychology. 1956; 52:283–287. [PubMed: 13367352]
137. Meyer WJ, Offenbach SI. Effectiveness of reward and punishment as a function of task complexity. Journal of Comparative and Physiological Psychology. 1962; 55:532–534.
138. Ashby FG, O'Brien JB. The effects of positive versus negative feedback on information-integration category learning. Percept Psychophys. 2007; 69(6):865–878. [PubMed: 18018967]

139. Vandist K, De Schryver M, Rosseel Y. Semisupervised category learning: The impact of feedback in learning the information-integration task. Attention, Perception, & Psychophysics. 2009; 71:328–341.

140. Calabresi P, Picconi B, Tozzi A, Di Filippo M. Dopamine-mediated regulation of corticostriatal synaptic plasticity. Trends in Neurosciences. 2007; 30:211–219. [PubMed: 17367873]

141. Wickens JR, Reynolds JNJ, Hyland BI. Neural mechanisms of reward-related motor learning. Current Opinion in Neurobiology. 2003; 13:685–690. [PubMed: 14662369]

142. Frank MJ. Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and non-medicated Parkinsonism. Journal of Cognitive Neuroscience. 2005; 17:51–72. [PubMed: 15701239]

143. Bjork, RA. Memory and metamemory considerations in the training of human beings. In: Metcalfe, J.; Shimamura, A., editors. Metacognition: Knowing about knowing. Cambridge, MA: MIT Press; 1994. p. 185-205.