Detection of Feedback Contingency Depends on Declarative Systems

Matthew J. Crossley

SRI International

W. Todd Maddox

University of California, Santa Barbara

F. Gregory Ashby

University of Texas at Austin

Abstract

Individuals in drug rehabilitation often experience relapse when returned to the original context of their drug use. Why do treatments generalize so poorly across different contexts? We recently addressed this question in the domain of procedural learning, which is thought to play an important role in forms of addiction, bad habits, and other maladaptive states (Crossley, Ashby, & Maddox, 2013). Our work suggests that feedback contingency, defined as the degree of non-randomness between response and outcome, plays a vital role in controlling a gate that normally prevents procedural knowledge from being modified during interventions. In particular, our results suggested that unlearning is possible only if feedback contingency is high. Here, we ask whether the estimation of feedback contingency depends on declarative mechanisms (e.g., prefrontal networks involved in working memory and executive reasoning). Our rationale is as follows: If feedback contingency is computed by declarative mechanisms, then increasing cognitive load during intervention via the concurrent performance of an additional task should disrupt the accurate estimation of contingency, and thereby prevent the gate on procedural learning from closing. We report the results from an experiment following this logic that suggests that feedback contingency estimation is indeed supported by declarative systems.

Detection of Feedback Contingency Depends on Declarative Systems

Introduction

Drug addiction is often treated in a rehabilitation clinic, and relapse occurs when the patient returns to the original context of their drug use (Higgins, Budney, & Bickel, 1995). How can we engineer treatments that generalize across contexts? Our earlier work (Crossley et al., 2013) showed that feedback contingency – used here to indicate the degree of non-randomness between response and outcome – may play a pivotal role in answering this question.

Crossley et al. (2013) studied procedural modification of behavior in a category-learning task. These experiments contained three phases of equal duration: acquisition, intervention, and test. Participants were trained on the category structure shown in panel C of Figure 1 during the acquisition phase. During the intervention phase, the underlying category structure was not changed, but feedback was manipulated (described in the next paragraph) in an attempt to erase the learning that occurred during the acquisition phase. During the test phase, feedback was returned to 100% veridical (as was the case during acquisition). Participants in the Relearning condition were tested on the original acquisition category-response mappings, whereas participants in the New-Learning condition were tested on a permuted category-response mapping. If the original acquisition-phase learning was preserved through intervention, then test phase performance in the Relearning condition should be better than acquisition-phase performance (i.e., participants should show savings), and test-phase performance in the New-Learning condition should be worse than acquisition-phase performance (due to interference from initial learning).

The changes in feedback used during the intervention phase were chosen by considering classic models of procedural learning (panel A of Figure 2), which extend classic models of associative learning. These models assume that procedural knowledge is learned at cortical-striatal synapses via DA-dependent synaptic plasticity. Consistent with

the classic view that procedural knowledge is robust against forgetting, simply removing feedback during intervention did not cause any substantial performance change (data not shown). Therefore, an intervention phase with completely random feedback was used (i.e., correct feedback was randomly given 25% of the time). Classic models predict that this intervention will replace initially acquired S-R associations with random strengths, thereby leading to identical or attenuated performance during the test phase in both the Relearning and the New-Learning conditions. Contrary to this prediction, random-feedback did not disrupt category knowledge. During the test phase, performance relative to initial acquisition was enhanced in the Relearning condition, and attenuated in the New-Learning condition; both suggesting savings (left panel of Figure 3).

The random feedback intervention result implies that *there is a gating mechanism that protects procedural learning from modification during random feedback.* One candidate biological system for this gating mechanism is the striatal cholinergic interneurons called TANs (for Tonically Active Neurons). It is thought that these neurons mediate cortical-striatal synaptic plasticity via presynaptic inhibition of cortical inputs to the striatum (Figure 2). As their name suggests, the TANs are tonically active in their default state; however, they learn to pause their firing when stimuli that predict reward are encountered. This allows striatal neurons to respond to cortical input, and cortical-striatal learning to take place. The TANs are driven by the parafascicular (Pf) nucleus in the thalamus, which signals salient environmental cues; their pause response occurs only when the Pf-TAN synapse is strong. The learning at both Pf-TAN synapses and cortical-striatal synapses is driven by the DA reinforcement signal.

An updated procedural learning model that includes the TAN gating mechanism has been shown to account for a variety of behavioral and physiological data from simple instrumental conditioning tasks (Ashby & Crossley, 2011a; Crossley, Horvitz, Balsam, & Ashby, 2016), including rapid relearning following extinction. However, the model failed to account for the random feedback results described above. This failure stems from the way

DA activity was modeled. In the model, DA firing reflected a reward prediction error (RPE; which equals the difference between the obtained and expected outcomes), and since random feedback is *by definition* completely unpredictable, the model predicted persistent RPEs and corresponding DA fluctuations during random feedback. These fluctuations prevented the TANs from closing the gate on cortical-striatal plasticity, and thereby permitted random feedback to disrupt initial learning.

In order for the TANs to reliably close the gate and protect cortical-striatal plasticity during random feedback, two conditions must be met:

1. *The model must detect when feedback has become random.* Random feedback has several properties, but our results suggest that the most critical – at least with respect to learning and unlearning – is that random feedback is non-contingent on action. More specifically, when the feedback is random the correlation between response confidence and feedback valence is zero. This correlation, which we refer to as *feedback contingency*, is a critical requirement of successful learning. For example, Ashby and Vucovich (in press) compared learning under high- and low-levels of feedback contingency in two different category-learning tasks – one that recruits declarative memory and one that recruits procedural memory. In both tasks, the high- and low-contingency conditions used exactly the same stimuli, had exactly the same optimal strategies, and optimal accuracy was 80% correct in all conditions. The results were virtually identical in the two tasks. Learning was good when feedback contingency was high, but degrading feedback contingency seemed to abolish all learning in most participants.

2. *The TANs must close the gate when non-contingent (e.g., random) feedback is detected.* This only occurs when Pf-TAN synapses undergo consistent weakening (see Figure 2B). This was modeled in (Crossley et al., 2013) by assuming that the magnitude of *the DA response is attenuated and biased below baseline when feedback contingency is low.*

With this modification to the DA system, the model not only accounts for savings in relearning after random feedback intervention, but also makes a novel prediction: *Mixed feedback that contains some random and some veridical feedback trials produces an intermediate level of contingency and therefore should prevent the TANs from completely closing the gate on cortical-striatal plasticity, thereby enabling modification of initial learning.* Our results supported this prediction (panel B of Figure 3) – that is, performance during the Test phase was identical in the Relearning and New-Learning conditions, and each was significantly different from initial learning.

One potential confound of this mixed-feedback intervention was that the rate of positive feedback was greater than during the random-feedback intervention (note the different accuracy levels during the intervention phase in panels A and B of Figure 3). To address this confound, Crossley et al. (2013) examined performance with an intervention that included random feedback, but with a positive feedback rate of 40%. This intervention increased the overall positive feedback rate, but since it was entirely random, feedback contingency was still zero. The results from this experiment are shown in panel C of Figure 3, and are virtually identical to our original random feedback intervention (where the positive feedback rate was 25%, see panel A, Figure 3). Thus, mixed feedback intervention may constitute an effective intervention for procedural modification.

Crossley et al. (2013) hypothesized that the gate on procedural learning — and therefore the key to procedural modification — is controlled by the degree of feedback contingency. Even so, we made no predictions about how contingency is computed by the nervous system. This article begins addressing this question – by asking whether feedback contingency is computed via declarative mechanisms (e.g., prefrontal networks involved in working memory and executive reasoning). Our rationale is as follows: If feedback contingency is estimated by declarative mechanisms, then increasing cognitive load during the intervention phase (by requiring participants to simultaneously perform a dual task) should impair the ability of participants to detect a change to random feedback, which

should cause the TANs gate to remain open, thereby allowing random feedback to modify the procedural knowledge that was acquired during initial learning.

With this goal in mind, we performed an experiment that mimicked the design of Crossley et al. (2013), except we added a concurrent numerical Stroop task during key classification trials. Previous research suggests that this dual task interferes with category learning that recruits declarative memory but not with category learning that recruits procedural memory (Waldron & Ashby, 2001). Thus, since our categorization task recruits procedural memory, any effect of the dual task on categorization performance should be due to its effects on contingency estimation, rather than on category learning per se.

In conditions 1 – 3, the first dual-task trial was 50 trials before the onset of intervention, and continued for 100, 200, or 300 trials, respectively. In Condition 4, the first dual-task trial was 50 trials after the onset of intervention, and continued for 250 trials. Comparing conditions 1 – 3 to condition 4 will allow us to assess the importance of disrupting the estimation of feedback contingency during the transition from acquisition to intervention. Condition 5 was a control condition in which no concurrent Stroop task was ever performed.

If feedback contingency estimation depends on declarative mechanisms then two behavioral markers are expected: (1) the dual task should slow the drop in categorization accuracy that occurs with the onset of random feedback; and (2) reacquisition of the original category learning should be slower in the dual task conditions than in the no dual-task control.

Methods

Design

There were four dual-task conditions (Condition 1 – 4) and one no dual-task control condition (Condition 5). The dual-task conditions differed on two dimensions, (1) the number of trials on which the dual task was applied, and (2) whether or not the onset of

the dual task preceded the onset of intervention.

Participants

163 participants were recruited from the University of Texas at Austin undergraduate population. There were 30 participants in Condition 1, 34 participants in Condition 2, 32 participants in Condition 3, 33 participants in Condition 4, and 34 participants in Condition 5. All participants completed the study and received course credit for their participation. All participants had normal or corrected-to-normal vision.

Stimuli and Categories

Stimuli were black lines that varied across trials only in length (pixels) and orientation (degrees counterclockwise rotation from horizontal). The stimuli are illustrated graphically in Figure 1, and were identical to those used by Crossley et al. (2013).

Procedure

Participants in all conditions were told that they were to categorize lines on the basis of their length and orientation, that there were four equally-likely categories, and that high levels of accuracy could be achieved. The experiment included three phases: acquisition (300 trials), intervention (400 trials), and reacquisition (150 trials). During acquisition and reacquisition, feedback was based on the participant's response, whereas feedback was random during the intervention. Participants were given no prior instructions about the phases, and the transition from one phase to another occurred without any warning to the participant.

At the start of each non-Stroop trial, a fixation point was displayed for 1 second and then the stimulus appeared. The stimulus remained on the screen until the participant generated a response by pressing the "Z" key for category "A", the "W" key for category B, the "/" key for category C, or the "P" key for category D. Written instructions informed participants of the category label to button mappings. An "invalid key" message was

displayed if any other button was pressed. The word "Correct" was presented for 1 second if the response was correct or the word "Wrong" was presented for 1 second if the response was incorrect (except during the intervention phase in which feedback was completely random).

Stroop trials began with a fixation point that was displayed for 1 second. The category stimulus and the Stroop stimuli (numbers flanking the category stimulus) were displayed simultaneously. After 200 ms the Stroop stimuli were replaced by white rectangles which remained on the screen until they made a category response. Responses emitted before the Stroop stimuli were replaced by white rectangles were not accepted. Feedback about the category response was given immediately in the same fashion as on non-Stroop trials. The word "value" or "size" then appeared on the screen prompting participants to indicate which side contained the numerically larger or the physically larger number. Participants pressed the "F" key to choose the number on the left or the "J" key to choose the number on the right. The word "Correct" was then again presented for 1 second if the response to the Stroop task was correct or the word "Wrong" was presented for 1 second if the response was incorrect. See Figure 1 for example trials both including and excluding the Stroop component. The Stroop task was included on trials 251-350 in condition 1, 251-450 in condition 2, 251-550 in condition 3 and 400-650 in condition 4.

Participants were instructed to try their hardest on both task components but to prioritize performance on the Stroop task. Both the category-learning task and the Stroop task were explained to participants prior to beginning the experiment, and on screen messages warned them when the Stroop component would begin, and again when it would end. These messages read, "You will now perform both the categorization task and the paired numbers task simultaneously. Keep trying your hardest!" and "You have now finished the section with the paired numbers task. You will now be shown only the line categorization task. Keep trying your hardest." 85% of Stroop trials the numerically larger number was physically smaller. The proportion of Stroop trials that prompted "size" or

"value" was split 50/50. Accuracy on the numerical Stroop task was indicated at the top of the screen when they received feedback regarding their performance on the concurrent task on each trial. This score was displayed in green if it was above 80% and red if it was below 80%. Note that when we refer to the "dual-task", we are referring to the Stroop task just described.

<div align="center">

Results

</div>

Numerical Stroop Accuracy

Figure 4 shows histograms characterizing mean dual-task performance seperately for each condition. Overall, mean accuracy on the dual-task was very good, with mean proportion correct at 0.88 in Condition 1, 0.87 in Condition 2, 0.84 in Condition 3, and 0.82 in Condition 4.

Classification Accuracy

Figure 5 shows the mean accuracy in each block of 25 trials across the duration of the experiment. Recall that if feedback contingency is estimated via declarative mechanisms, then (1) dual-task trials should slow the change in classification performance during intervention, and (2) dual-task conditions should show reduced savings relative to the no dual-task control. We see evidence for both features in our data.

Acquisition. Conditions 1 − 5 are identical for the first 250 trials (10 blocks) of acquisition (before dual-task onset), and so we expect performance during these blocks to be the same across conditions. This is clearly the case by visual inspection of Figure 5, and is supported by the results of a 5 Condition × 10 Block repeated-measures ANOVA. The main effect of Condition was nonsignificant $[F(4, 1620) = 1.89, p = 0.11, \Omega^2 = 0.00]$, and so was the Condition × Block interaction $[F(4, 1620) = 1.08, p = 0.37, \Omega^2 = 0.00]$. However, there was a significant effect of Block $[F(1, 1620) = 191.00, p < 0.001, \Omega^2 = 0.10]$, reflecting improvement across the acquisition phase.

Intervention. If the estimation of feedback contingency depends on declarative mechanisms, then we expect change in performance during intervention to be slowed during the simultaneous performance of the dual task. This is clearly seen in Figure 5, and is supported by the results of a 5 condition × 14 block repeated-measures ANOVA. A significant effect of Condition $[F(4, 2598) = 9.74, p = 0.00, \Omega^2 = 0.01]$ reflected an overall difference in intervention performance in dual-task conditions relative to the no dual-task control. A significant effect of Block $[F(1, 2598) = 166.65, p < 0.001, \Omega^2 = 0.06]$ reflected the change in classification performance across intervention seen in all conditions. The Condition × Block interaction was also significant $[F(4, 2598) = 14.64, p < 0.001, \Omega^2 = 0.02]$, reflecting the slower change in performance in the dual-task conditions relative to the no dual-task control.

The directional interpretation of the omnibus tests is supported by several planned comparisons on the overall mean accuracies during the intervention phase. First, intervention accuracy in all dual-task conditions was significantly different from intervention accuracy in the no dual-task control [condition 1 > condition 5: $t(940) = 2.78$, $p < .05, d = 0.25$; condition 2 > condition 5: $t(1046) = 5.31, p < .01, d = 0.87$; condition 3 > condition 5: $t(962) = 5.45, p < .01, d = 0.96$; condition 4 > condition 5: $t(996) = 2.99$, $p < .01, d = 0.28$]. Second, the dual task slowed the accuracy drop during intervention only in conditions in which it was first introduced during the acquisition phase [condition 2 > condition 4: $t(1063) = 1.94, p = 0.05, d = 0.12$; condition 3 > condition 4: $t(1037) = 2.24$, $p = 0.03, d = 0.16$], although this difference was not observed in condition 1 (shortest dual-task exposure) [condition 1 > condition 4: $t(1006) = -0.23, p = 0.82, d = 0.00$].

Savings: Group Level. If the computation of feedback contingency depends on declarative systems, then we expect the dual-task conditions to exhibit less savings than the no dual-task control – that is, we expect reacquisition of the original categories to be slower under dual-task conditions. This is apparent via visual inspection of Figure 5 (e.g., the red lines are always below the blue line), and also of Figure 6, which shows the savings

observed for every participant in every condition as a box plot with individual participant data points superimposed. As a comparison, Figure 6 also shows the results from the random-feedback intervention condition of Crossley et al. (2013), which was similar in design to our no dual-task control condition (although the Crossley et al. condition included fewer intervention trials).

Figure 7 shows the savings during each 25-trial block, where savings is defined as reacquisition accuracy minus acquisition accuracy on the block in the same ordinal order (e.g., first reacquisition block accuracy minus first acquisition block accuracy, etc.). Several features of these data are worth noting. First, savings in the no dual-task control condition is uniformly greater than in any other condition. Second, there is little difference in savings in any dual-task condition. Third, in every condition, the curves are decreasing, and fourth, in every case, savings become negative, indicating that asymptotic accuracy during reacquisition was lower than during acquisition in every condition.

These observations are supported via the results of a 5 condition $\times$ 6 block ANOVA. There was a significant effect of Condition $[F(1, 974) = 5.77, p < 0.05, \Omega^2 = 0.01]$, indicating less savings in the dual-task conditions relative to the no dual-task condition. There was a significant effect of Block $[F(1, 974) = 22.55, p < 0.001, \Omega^2 = 0.02]$, indicating that savings in all conditions was most prominent during early blocks and gradually decreased (see Figure 7). The Condition $\times$ Block interaction was n.s. $[F(1, 974) = 0.001, p = 0.97, \Omega^2 = 0.00]$.

The directional interpretation of these omnibus tests is supported by several planned comparisons. First, Condition 2 savings were significantly less than in the no dual-task control condition $[t(63) = 2.09, p = 0.04, d = 0.55]$, whereas savings in the other dual-task conditions were all marginally less than in the control condition [condition 1 < control: $t(46) = -1.24, p = 0.22, d = 0.22$; condition 3 < control: $t(57) = -1.63, p = 0.11, d = 0.35$; condition 4 < control: $t(61) = -1.67, p = 0.10, d = 0.36$]. Second, there were no significant

differences in savings between any dual-task conditions [1] [condition 1 > condition 2: $t(53) = 0.37, p = 0.71, d = 0.02$; condition 1 > condition 3: $t(56) = 0.12, p = 0.90, d = 0.00$; condition 1 > condition 4: $t(53) = 0.08, p = 0.94, d = 0.00$; condition 2 > condition 3: $t(63) = -0.29, p = 0.78, d = 0.01$; condition 2 > condition 4: $t(65) = -0.36, p = 0.72, d = 0.02$; condition 3 > condition 4: $t(62) = -0.06, p = 0.96, d = 0.00$].

Participant-Level Savings. Figure 6 shows large individual differences – with many participants in all conditions expressing both positive and negative savings. To examine these differences in more detail, we performed an exploratory analysis that examines what factors might predict savings in individual participants. In particular, we performed a multiple regression with three predictors: (1) mean dual-task accuracy, (2) mean intervention accuracy, and (3) mean acquisition accuracy. Dual-task performance might reflect the degree to which the computation of feedback contingency is impaired, and thereby predict savings. Mean intervention accuracy may reflect the degree to which procedural knowledge is being expressed, and therefore is vulnerable to modification. Mean acquisition accuracy may predict savings in that the stronger the initial learning, the more robust to intervention it may be, and the more likely it should be manifest as savings in relearning.

Figure 8 plots savings against each of these predictor variables. Note that data was pooled across the four dual-task conditions for this analysis. The regression revealed a significant negative coefficient for mean acquisition performance [$\beta = -0.23$, $t(125) = -2.17, p < .05$], showing that the better a participant did during acquisition, the less savings they were likely to express. The coefficient for mean intervention performance was positive and nearly significant [$\beta = .30, t(125) = 1.79, p = .07$], showing that there was a trend for higher intervention accuracy to predict better savings. The coefficient for mean dual-task performance was not significantly different from zero [$\beta = .09, t(125) = .87$,

---

[1] Note that the following p-values should all be corrected for multiple comparisons. However, any such correction would only increase each p value, and therefore would not change our nonsignificance conclusions.

$p = .39$]. The overall model, however, was not a good predictor of savings, accounting for only 5% of the total variance in the data [$r^2 = .05, F(3, 125) = 2.29, p = .08$].

Discussion

Summary

Feedback contingency, which we define as the correlation between response confidence and feedback valence, can be manipulated experimentally by varying the randomness of feedback. Our current and previous results (Crossley et al., 2013) suggest that such manipulations may be key to flexibly modifying procedural memories. To our knowledge, this article reports results from the first behavioral experiments that investigate the cognitive and neural mechanisms that estimate feedback contingency. Specifically, our goal was to determine whether prefrontal-based declarative memory mechanisms mediate contingency estimation. If they do, then a dual task that depends on working memory and executive function should make it more difficult for participants to recognize the sudden onset of random feedback. In our experiments, behavioral signatures of this difficulty would include (1) a slowed decrease in classification accuracy during intervention, and (2) decreased savings in relearning relative to a no dual-task control. Our results were consistent with both of these predictions.

Weakness of Raw Savings

Perhaps our most conspicuous and unanticipated result was the lack of large and robust savings in any condition. Even in the no dual-task control, savings is only clearly expressed during early blocks of reacquisition – that is, during the first two blocks of reacquisition, performance was better than during initial acquisition, but during the last few blocks it was worse. We speculate that this is due to the duration of the intervention phase, which was a full 100 trials longer than in previous studies (Crossley et al., 2013; Crossley, Ashby, & Maddox, 2014). Does this finding suggest that longer interventions are

a possible key to true unlearning? Several features of our data argue against this hypothesis. First, participants in the no dual-task control condition showed considerable savings during the first 50 trials (i.e., 2 blocks) of reacquisition (see Figure 7). If the longer intervention caused unlearning then such savings should not have occurred. Second, this initial savings is reversed to an interference during the last half of the reacquisition phase. True unlearning predicts zero savings throughout reacquisition. The most parsimonious account of the negative savings that occurred in every condition during the latter half of the reacquisition phase may be participant fatigue.

At first glance, one might therefore wonder whether all our results simply reflect that categorization with a dual-task is more tiring than categorization without a dual task. In other words, perhaps dual-task participants were more fatigued than no dual-task control participants at the beginning of the reacquisition phase, and this extra fatigue was the primary reason they did not show savings, rather than because the intervention disrupted their acquisition memory traces. However, a closer examination argues against this hypothesis. First, and most importantly, all three conditions in which the dual-task overlapped the transition from acquisition to intervention showed a slower decrease in intervention accuracy than in either the no dual-task control condition or in Condition 4, where the dual task did not begin until later during intervention (see Figure 5). Thus, one important signature of unlearning appeared in Conditions 1 – 3 early in the experiment (i.e., well before the beginning of reacquisition), before fatigue should have been a factor. Second, if participant fatigue was the driving factor, then there should be differences among dual-task conditions, which we did not observe. Specifically, Condition 1 had the fewest dual-task trials and Conditions 3 and 4 had the most. So the fatigue hypothesis should predict poorer reacquisition performance in Conditions 3 and 4 than in Condition 1. Figure 7 disconfirms this prediction. Third, we performed a linear regression (on the pooled data from all five experimental conditions) that tried to predict savings as a function of overall experiment time. The idea here is that the longer the experiment took,

the more susceptible the participant should be to fatigue, and the less savings they should express. The coefficient for total experiment time was not significantly different from zero $[\beta = 7.5 \times 10^{-6}, t(161) = .93, p = .35]$, and did not predict a significant proportion of variance in savings $[r^2 = .005, F(1, 161) = .87, p = .35]$.

Category Learning as a Procedural Skill

A natural question for readers unfamiliar with the category-learning literature is whether our behavioral paradigm is a good choice for studying procedural behaviors. In other words, how can a task with such simple motor demands (e.g., push a button) possibly recruit procedural networks that are strongly tied to motor processes? In fact, the empirical evidence is strong that performance improvements in the classification task used here are mediated via procedural learning and memory. A large database of evidence suggests that humans have multiple, qualitatively distinct category-learning systems (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby & Maddox, 2005; Erickson & Kruschke, 1998), and according to this view, procedural memory is used to form many-to-one stimulus-to-response mappings (S-R associations), whereas declarative memory is used to apply rules and test explicit hypotheses about category membership.

The majority of this evidence comes from prior research with rule-based (RB) and information-integration (II) category-learning tasks (Helie, Roeder, & Ashby, 2010; Nomura et al., 2007; Soto, Waldschmidt, Helie, & Ashby, 2013; Waldschmidt & Ashby, 2011). In RB tasks, the categories can be learned via an explicit hypothesis-testing procedure (Ashby et al., 1998). In the simplest variant, only one dimension is relevant (e.g., bar width), and the task is to discover this dimension and then map the different dimensional values to the relevant categories. In II tasks, accuracy is maximized only if information from two or more stimulus dimensions is integrated perceptually at a pre-decisional stage (Ashby & Gott, 1988). In most cases, the optimal strategy in II tasks is difficult or impossible to describe verbally (Ashby et al., 1998). Verbal rules may be (and

sometimes are) applied, but they lead to suboptimal performance. The task used here (and illustrated in Figure 1) was an II category-learning task.

At least 25 different behavioral dissociations tie II learning to procedural memory (and RB learning to declarative memory; for reviews, see Ashby & Maddox, 2005, 2010; Ashby & Valentin, in press). For example, one behavioral signature of procedural learning is that because of its motor component, switching the locations of the response keys interferes with performance (Willingham, Wells, Farrell, & Stemwedel, 2000). In agreement with this result, switching the locations of the response keys interferes with II categorization performance, even when the task only includes two categories[2] (Ashby, Ell, & Waldron, 2003; Maddox, Bohil, & Ing, 2004; Spiering & Ashby, 2008).

This hypothesis is further supported by a variety of investigations into the neural underpinnings of successful II and RB learning. Specifically, success in RB tasks depends on a broad neural network that includes the prefrontal cortex (PFC), anterior cingulate, the head of the caudate nucleus, and medial temporal lobe structures—regions that are also frequently associated with declarative memory and executive attention (Brown & Marsden, 1988; Filoteo, Maddox, Song, et al., 2007; Muhammad, Wallis, & Miller, 2006; Seger & Cincotta, 2006). Success in II tasks, on the other hand, depends on regions that have been implicated in procedural memory, including the striatum, premotor cortex, and the associated sensorimotor basal ganglia loop (Ashby & Ennis, 2006; Filoteo, Maddox, Salmon, & Song, 2005; Knowlton, Mangels, & Squire, 1996; Nomura et al., 2007). This network is consistent with the idea that S-R associations are built at cortical-striatal synapses via dopamine-dependent reinforcement learning (Ashby & Crossley, 2011b; Houk, Adams, & Barto, 1995; Joel, Niv, & Ruppin, 2002).

---

[2]In contrast, the same button switch does not interfere with RB performance in tasks where the RB categories are created by simply rotating the II categories by 45°

Therapeutic Relevance

The old adage of "it's like riding a bike" is a surprisingly accurate description of procedural knowledge, reflecting its remarkable retention over years without practice. Paradigms designed to study procedural learning in the lab have echoed this adage, reporting savings in learning up to a year after training (Romano, Howard, & Howard, 2010; Turner, 2012). However, the stability of procedural memory comes at the cost of remarkable inflexibility. For example, changing any stimulus or response parameter that was present during training can prove catastrophic to performance (Rozanov, Keren, & Karni, 2010; Dienes & Berry, 1997). While resilience and inflexibility are desirable traits when a useful skill has been sufficiently learned, they can also lead to persistent maladaptive behaviors that have serious negative consequences, and in some cases may prove detrimental to a person's health (e.g., drug abuse). Unfortunately, neither the potential for modification of procedural knowledge, nor a method to do so, are well understood.

Our previous research identified the interplay between the striatal cholinergic interneurons and the midbrain dopamine system in controlling the eligibility of procedural knowledge for modification. Directly targeting this network for improved interventions is unfortunately challenging, due to the difficulty of manipulating and measuring subcortical networks. Here, we looked for more easily accessible cortical substrates that may control the striatal mechanism. Our results indicate that prefrontal networks likely do play an important role in controlling the estimation of feedback contingency, and therefore may provide an accessible cortical target for electrical or magnetic intervention.

References

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), 442–481.

Ashby, F. G., & Crossley, M. J. (2011a). A computational model of how cholinergic
interneurons protect striatal-dependent learning. *Journal of Cognitive Neuroscience*,
23(6), 1549–1566. Retrieved 2016-09-15, from
`http://www.mitpressjournals.org/doi/abs/10.1162/jocn.2010.21523`

Ashby, F. G., & Crossley, M. J. (2011b). 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*, 1–36.

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., & Maddox, W. T. (2005). Human category learning. *Annual Review of
Psychology*, *56*, 149-178.

Ashby, F. G., & Maddox, W. T. (2010). Human category learning 2.0. *Annals of the New
York Academy of Sciences*, *1224*, 147-161.

Ashby, F. G., & Valentin, V. V. (in press). Multiple systems of perceptual category
learning: Theory and cognitive tests. In H. Cohen & C. Lefebvre (Eds.), *Handbook of
categorization in cognitive science, second edition* (p. ?). New York: Elsevier.

Ashby, F. G., & Vucovich, L. E. (in press). The role of feedback contingency in perceptual
category learning. *Journal of Experimental Psychology: Learning, Memory, and*

Cognition, *?*(?), ?

Brown, R. G., & Marsden, C. D. (1988). Internal versus external cues and the control of attention in parkinson's disease. *Brain*, *111*(2), 323-345.

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. Retrieved 2016-09-18, from `http://psycnet.apa.org/journals/xge/142/3/710/`

Crossley, M. J., Ashby, F. G., & Maddox, W. T. (2014). Context-dependent savings in procedural category learning. *Brain and Cognition*, *92*, 1–10.

Crossley, M. J., Horvitz, J. C., Balsam, P. D., & Ashby, F. G. (2016). Expanding the role of striatal cholinergic interneurons and the midbrain dopamine system in appetitive instrumental conditioning. *Journal of neurophysiology*, *115*(1), 240–254. Retrieved 2016-09-18, from `http://jn.physiology.org/content/115/1/240.abstract`

Dienes, Z., & Berry, D. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin & Review*, *4*(1), 3–23.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*(2), 107-140.

Filoteo, J. V., Maddox, W. T., Salmon, D. P., & Song, D. D. (2005). Information-integration category learning in patients with striatal dysfunction. *Neuropsychology*, *19*(2), 212-222.

Filoteo, J. V., Maddox, W. T., Song, D. D., et al. (2007). Characterizing rule-based category learning deficits in patients with parkinson's disease. *Neuropsychologia*, *45*(2), 305–320.

Helie, S., Roeder, J. L., & Ashby, F. G. (2010). Evidence for cortical automaticity in rule-based categorization. *The Journal of Neuroscience*, *30*(42), 14225-14234.

Houk, J., Adams, J., & Barto, A. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In J. L. Houk J. C. Davis &

D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 249–270). Cambridge, MA: MIT Press.

Joel, D., Niv, Y., & Ruppin, E. (2002). ActorâĂŞcritic models of the basal ganglia: New anatomical and computational perspectives. *Neural networks*, *15*(4), 535–547. Retrieved 2016-09-18, from
`http://www.sciencedirect.com/science/article/pii/S0893608002000473`

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science*, *273*(5280), 1399–1402.

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.

Muhammad, R., Wallis, J. D., & Miller, E. K. (2006). A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum. *Journal of Cognitive Neuroscience*, *18*(6), 974-989.

Nomura, E., Maddox, W., Filoteo, J., Ing, A., Gitelman, D., Parrish, T., . . . Reber, P. (2007). Neural correlates of rule-based and information-integration visual category learning. *Cerebral Cortex*, *17*(1), 37-43.

Romano, J. C., Howard, J. H., & Howard, D. V. (2010). One-year retention of general and sequence-specific skills in a probabilistic, serial reaction time task. *Memory*, *18*(4), 427–41. Retrieved from
`http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2894701&tool=pmcentrez&re`
doi: 10.1080/09658211003742680

Rozanov, S., Keren, O., & Karni, A. (2010, may). The specificity of memory for a highly trained finger movement sequence: Change the ending, change all. *Brain Research*, *1331*, 80–7. Retrieved from `http://www.ncbi.nlm.nih.gov/pubmed/20298683` doi: 10.1016/j.brainres.2010.03.019

Seger, C. A., & Cincotta, C. M. (2006). Dynamics of frontal, striatal, and hippocampal

systems during rule learning. *Cerebral Cortex*, *16*(11), 1546-1555.

Soto, F. A., Waldschmidt, J. G., Helie, S., & Ashby, F. G. (2013). Brain activity across the development of automatic categorization: A comparison of categorization tasks using multi-voxel pattern analysis. *NeuroImage*, *71*, 284–297.

Spiering, B. J., & Ashby, F. G. (2008). Response processes in information–integration category learning. *Neurobiology of Learning and Memory*, *90*(2), 330-338.

Turner, B. O. (2012). *Long-term retention in the cortical and subcortical procedural category-learning systems.* University of California, Santa Barbara. Retrieved 2016-02-10, from `http://gradworks.umi.com/35/45/3545135.html`

Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychonomic Bulletin & Review*, *8*(1), 168–176.

Waldschmidt, J. G., & Ashby, F. G. (2011). Cortical and striatal contributions to automaticity in information-integration categorization. *Neuroimage*, *56*(3), 1791–1802.

Willingham, D. B., Wells, L. A., Farrell, J. M., & Stemwedel, M. E. (2000). Implicit motor sequence learning is represented in response locations. *Memory & Cognition*, *28*(3), 366–375.
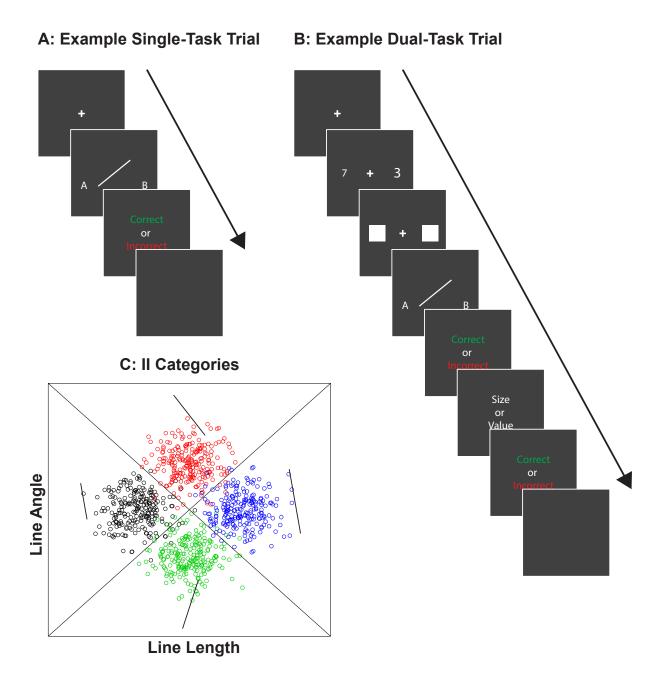
A: Example Single-Task Trial   **B: Example Dual-Task Trial**

C: II Categories

Figure 1. **A:** An Example trial during single-task conditions. **B:** An example trial during dual-task conditions. **C:** The II categories used during the acquisition phase of Crossley et al. (2013).
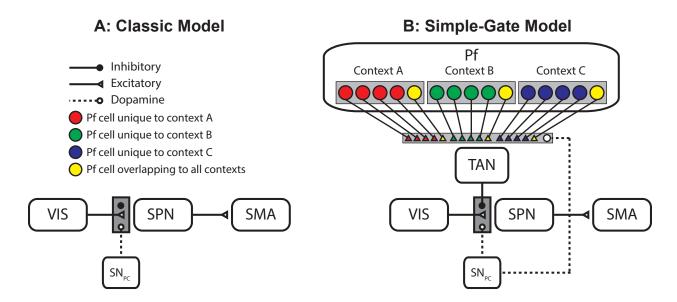
Figure 2. **A: The Classic Model.** A classic model of procedural learning based on a greatly simplified representation of the "direct pathway" through the basal ganglia. S-R associations are learned at cortical-striatal synapses, which are modified via dopamine-dependent reinforcement learning. The likelihood of repeating actions that lead to *unexpected* positive outcomes is gradually increased, and the likelihood of repeating actions that lead to *unexpected* negative outcomes is gradually decreased. **B: The TANs Model.** The classic model of procedural learning with the addition of a context-specific Pf-TAN pathway. This pathway acts as a gate on cortical-striatal synaptic plasticity, permitting or preventing the learning and expression of procedural knowledge. (MSN - medium spiny neuron of the striatum. D1 - Direct pathway MSN expressing the D1 DA receptor. D2 - Indirect pathway MSN expressing the D2 DA receptor. SMA - Supplementary Motor Area. SNpc - substantia nigra pars compacta. Pf - parafascicular nucleus of the thalamus. VIS - visual cortex)
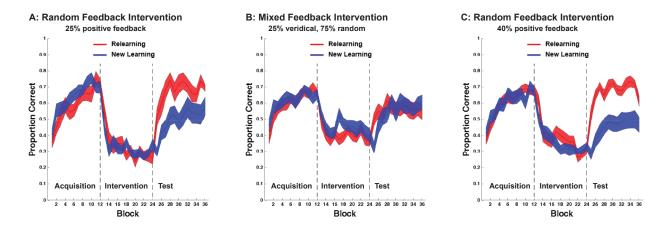
Figure 3. Behavioral results with different interventions. **A:** Random feedback intervention with 25% positive feedback. Accuracy drops to near chance during intervention, but is reacquired faster than original learning in the Relearning condition (red). In contrast, a lasting interference is observed in the New Learning condition (blue). Both results are consistent with the hypothesis that initial learning was not overwritten by random feedback. **B:** Mixed feedback intervention. Accuracy drops during intervention – though not to chance (i.e., 25%) – but subsequent learning proceeds at approximately the same rate and to the same extent as initial learning when either the original category-response mappings (red) or new category-response mappings (blue) are introduced. These results are consistent with the hypothesis that initial learning was overwritten during the intervention. **C:** Random feedback intervention with 40% positive feedback. Results are qualitatively identical to random feedback intervention with 25% positive feedback, implying that the mixed feedback results were driven by feedback contingency and not by positive feedback.
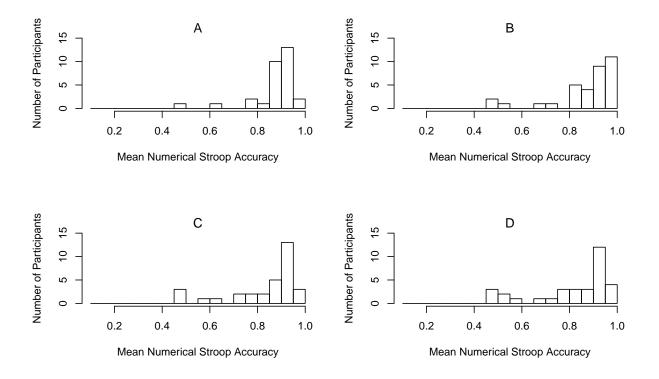
Figure 4. Histograms showing distribution of mean Numerical Stroop accuracy seperately for each condition. **A:** Condition 1. **B:** Condition 2. **C:** Condition 3. **D:** Condition 4.
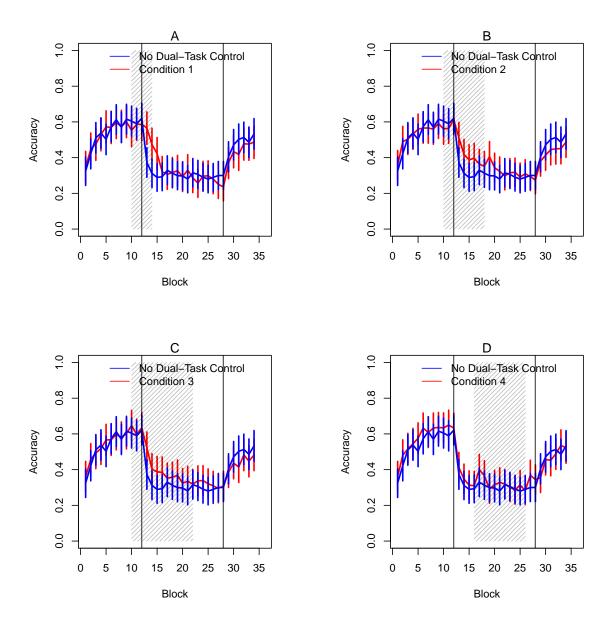
Figure 5. Mean accuracy per 25 trial block. The blue line in each panel is Condition 5 (no dual-task control). The hatch marks indicate dual-task trials. The key features are (1) dual-task slows the change in classification strategy (seen in this plot as "accuracy" decline), and (2) the dual-task conditions show less savings than the no dual-task control. There is no obvious dose-dependent effect of the dual task, nor is there an obvious difference between dual-task conditions. **A:** Condition 1 (dual-task applied on trial 251 through trial 350). **B:** Condition 2 (dual-task applied on trial 251 through trial 450). **C:** Condition 3 (dual-task applied on trial 251 through trial 550). **D:** Condition 4 (dual-task applied on trial 351 through trial 650). Error bars are SEM.
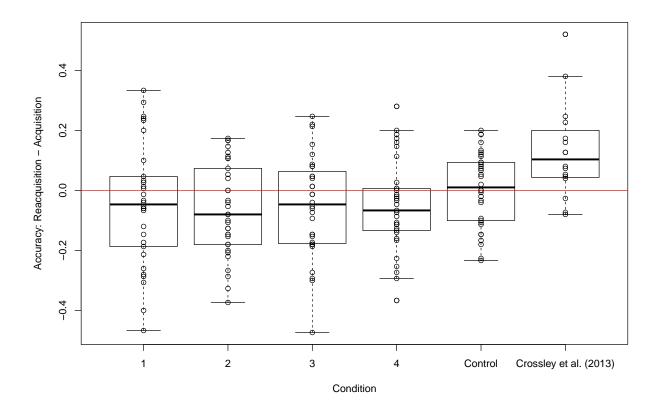
Figure 6. savings (mean of all 150 reacquisition trials - mean of the first 150 acquisition trials) in all conditions of the present experiment, and also including data from the random feedback condition of crossley et al. (2013), which shows the savings observed in a no dual-task control condition with only 300 trials of intervention. Each circle corresponds to a single participant.
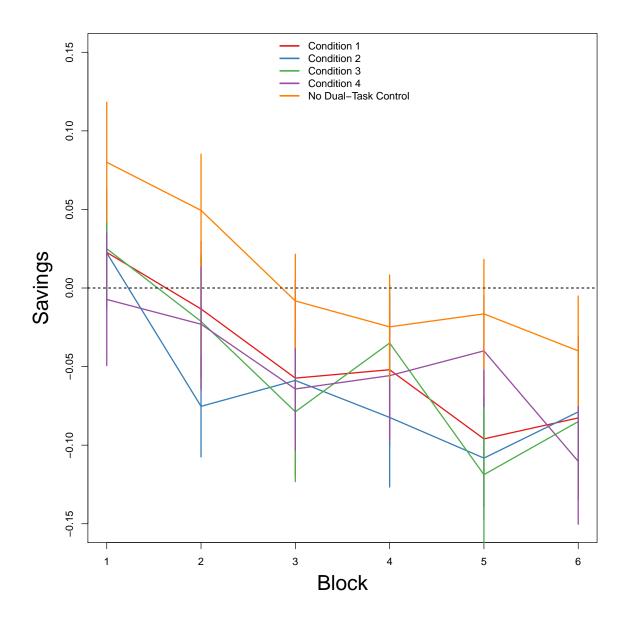
Figure 7. Savings (reacquisition - acquisition) per 25 trial block. Error bars are SEM.
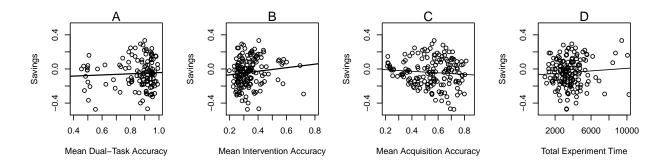
Figure 8. Regression analyses to examine possible predictor variables for savings. A: Mean dual-task accuracy. B: Mean Intervention accuracy. C: Mean acquisition accuracy. D: Total experiment time.