~~Criterial Learning and Feedback Delay: Insights from Computational Models and Behavioral Experiments~~Is there a Criterion in Criterial Learning? Insights from Studying Feedback Delays

Matthew J. Crossley[1, 2], Benjamin O. Pelzer[3], and F. Gregory Ashby [4]
[1]School of Psychological Sciences, Macquarie University, Sydney, Australia
[2]Macquarie University Performance and Expertise Research Centre, Macquarie University, Sydney, Australia
[3]Independent Researcher
[4]Department of Psychological & Brain Sciences, University of California, Santa Barbara

Abstract

The notion of a response criterion is ubiquitous in psychology, yet its cognitive and neural underpinnings remain poorly understood. Two experiments and extensive computational modeling were used to test between two strikingly different interpretations of the criterion. The traditional account is that decisions are made by comparing the stimulus value to a stored value of the criterion. A conceptually different interpretation is that learning instead is a process of associating responses with stimuli, and that the criterion is simply the hypothetical value that separates stimuli associated with contrasting responses. The experiments and modeling tested between these two interpretations by contrasting the effects on criterial learning of feedback delays versus increases in the duration of the intertrial interval in a one-dimensional category-learning task. The empirical results strongly suggested that human criterial learning is sensitive to feedback delay but not to the duration of the intertrial interval. The computational modeling showed that these results are compatible with a stimulus-response learning account, and incompatible with all versions of the stored-criterion account, except for the subset of these models that explicitly assume the criterial updating process is sensitive to feedback delay.

Keywords: response criterion; criterial learning; associative learning; categorization; procedural learning

Introduction

The notion of a response criterion is ubiquitous in psychology. It is a key component of almost all decision models. For example, the hypothesis that even YES-NO detection decisions are determined by comparing the sensory magnitude to a response criterion that is under the observer's control, rather than to a fixed absolute threshold, allowed signal detection theory to supplant classical threshold theory as the dominant model in psychophysics (Green & Swets, 1966). All models that include a response criterion assume its value is learned and can shift if changes are made to instructions or payoffs. So criterial learning is a fundamental component of almost all decision-making models. Despite its importance, however, the cognitive and neural mechanisms that underlie criterial learning remain poorly understood.

This article addresses this shortcoming through a combination of ~~computational modeling and~~ empirical data collection and computational modeling. Specifically, ~~we develop and test three different computational models that make qualitatively different assumptions about how~~ our goal is to test between two general possibilities. One is that the criterion is learned ~~. The models differ in the role they assign to working memory and~~ in ~~whether they treat the response criterion as a fundamental psychological construct, or instead assume that behavior is driven purely by~~ the sense that it is updated trial-by-trial and that its current value is stored in memory. This is the hypothesis implicitly assumed by most decision-making models (e.g., signal-detection theory). The second possibility is that the criterion has no psychological meaning and learning instead is of stimulus-response ~~associations without any criterion guiding responses. These models are then tested in two behavioral experiments. The modeling and empirical focus are on how feedback delays and the length~~ (SR) associations. According to this account, the criterion is simply the sensory (or cognitive) value that separates percepts associated with the contrasting responses. For example, there is evidence that procedural learning works in this way (Ashby & Waldron, 1999).

Two independent variables that seem especially likely to discriminate between these two alternatives are the duration of the intertrial interval (ITI) ~~affect criterial learning. All criterial-learning models assume that updating (i.e., learning) of the criterion occurs during the time interval between feedback presentation and the stimulus presentation that defines the onset of the next trial. So feedback delay~~ and the delay between the response and the feedback. If the value of a criterion is updated following feedback and ~~the length of the ITI are the independent variables that most clearly differentiate the conflicting predictions of criterial-learning models. Furthermore, feedback delays are known to impair some forms of learning (i.e., procedural)~~ then held in memory until needed again, then the longer it must be held in memory, the more time there is for its value to drift. In other words, there should be more criterial drift with long ITIs than with short ITIs, and as a result, increasing the ITI should impair performance. In contrast, if the criterion has no psychological meaning and instead learning is of SR associations, then learning should be impaired by increases in feedback delay. The idea is that positive feedback strengthens and negative feedback weakens recently active synapses that associate a response with the presented stimulus. As a result, increasing the feedback delay will

Correspondence: Matthew J. Crossley, PhD, School of Psychological Sciences, Macquarie University, Australian Hearing Hub, 16 University Ave, Macquarie University, NSW 2109, Australia. Email: matthew.crossley@mq.edu.au

weaken the trace of recently active synapses, and therefore reduce this type of SR learning. This model is supported by many previous reports that feedback delays as short as 3 s impair procedural learning much more than ~~others (i.e., learning that relies on declarative memory), so feedback-delay manipulations offer a powerful method of disambiguating the nature of~~ declarative learning that depends on executive attention and working memory (Dunn et al., 2012; Ell et al., 2009; Maddox et al., 2003; Maddox & Ing, 2005; Worthy et al., 2013) .

Our general approach is as follows. First, we establish an empirical database by describing the results of two experiments that report the results of manipulating ITI and feedback delay on criterial learning. Although a variety of different tasks could be used to study criterial learning~~(Ell et al., 2009; Maddox et al., 2003; Maddox & Ing, 2005).~~

~~Although the models could be tested in any task that depends on criterial learning~~, the two experiments we describe used a one-dimensional category-learning task. In such tasks, stimuli vary across trials on two or more dimensions – one relevant to categorization and one or more that are irrelevant. The observer's goal is to identify the relevant dimension and learn the response criterion that maximizes accuracy. This task has been used in hundreds of studies, and all current models of performance in this task emphasize the role of criterial learning.

In the first ~~behavioral~~ experiment, participants were explicitly instructed about the relevant dimension, thereby isolating criterial learning from rule selection and switching processes. The results showed that ~~short feedback delays impaired learning compared to immediate feedback. The second behavioural experiment used stimuli composed of binary features – where~~ increasing the feedback delay impaired learning but increasing the ITI did not. Experiment 2 used stimuli that varied on six binary-valued dimensions and participants were not told which of these dimensions was relevant to the categorization decision. The observer's task was to discover the single relevant dimension, but after this dimension was identified, no criterial learning ~~is needed – and did not instruct participants about the~~ was required since there were two categories and the stimuli had exactly two values on the relevant dimension. Thus, ~~this experiment~~ Experiment 2 isolated rule selection and switching from criterial learning. The results ~~found that increasing~~ showed that performance was unaffected either by increases in feedback delay or ITI~~did not affect performance~~. These findings therefore suggest that feedback delays impact criterial learning but not the discovery of the relevant stimulus dimension. ~~More broadly, they indicate~~

The results of Experiments 1 and 2 seem to suggest that criterial learning ~~may recruit~~ is a form of procedural learning, even in tasks that seem to rely on explicit, rule-based processes~~.~~

~~We developed three computational models, each representing different architectures of criterial learning, and examined their sensitivity to the duration of ITIs and feedback delay. Two of the models assume~~ , and therefore that the criterion is ~~stored in memory and that response decisions are made by comparing the current stimulus-driven percept to this stored referent. Both of these models further posit that the internal representation of the criterion is updated following errors using a simple gradient-descent rule. The third model assumes that no criterion is stored nor used to generate responses. Instead, reinforcement learning is used to form stimulus-response associations and these associations drive responding without any appeal to the notion of a criterion.~~

~~For each of the three models, denote the time of stimulus presentation on the current trial by T_S, the response time by T_R, the time when feedback is displayed by T_F, and the time when the stimulus that begins the next trial is displayed by T_{S+}. Then note that the feedback delay equals $t_{FD} = T_F - T_R$ and the duration of the ITI equals $t_{ITI} = T_{S+} - T_F$.~~

~~The time-dependent drift model assumes that the observer constructs a criterion, holds this value in working memory, and then makes response decisions by comparing the percept to the criterion. The model also assumes that both the memory representation of the criterion and the perceived value of the stimulus gradually drift over time. The model further assumes that the extent of this drift increases with time, both within a trial and between consecutive trials.~~

~~Let $x_n(t)$~~ an emergent construct with no psychological meaning. Before accepting this conclusion however, it is important to ask whether some type of criterion-learning model can account for these results. To investigate this question, we developed a variety of computational models of criterial learning of two different types. Eight different models assumed the criterion is learned and updated trial-by-trial. These were created by factorially combining three different binary assumptions: the percept drifts randomly over time (yes or no), the criterion drifts randomly over time (yes or no), and ~~$c_n(t)$ denote the values of the percept and the criterion, respectively, on trial n at time t. Then the decision rule on trial n is:~~

$$\text{Respond } R_n = \begin{cases} \text{A}, & \text{if } x_n(\mathrm{T_S}) \le c_n(\mathrm{T_S}) \\ \text{B}, & \text{if } x_n(\mathrm{T_S}) > c_n(\mathrm{T_S}). \end{cases}$$

~~If positive feedback is received, then the value of the criterion remains unchanged for the next trial (except for drift – see Equation 3). If negative feedback is received, then the criterion is modified according to the standard model (Sutton & Barto, 1998):~~

$$c_n(\mathrm{T_F} + \Delta_t) = c_n(\mathrm{T_F}) + \alpha[x_n(\mathrm{T_F}) - c_n(\mathrm{T_F})],$$

~~where α is a learning-rate parameter, and Δ_t is the time it takes to complete the updating. It is straightforward to show that the iterative Equation 2 is equivalent to computing a weighted mean (weighted by recency) of the values of all percepts that occur on error trials (e.g., Ashby, 2017). This updating rule will gradually converge on the optimal criterion value. Since this model assumes that criterial learning relies on working memory – and the available evidence suggests that logical reasoning and working memory are unaffected by feedback delays of several seconds (e.g., in one-dimensional rule-based category learning tasks; Ell et al., 2009; Maddox et al., 2003; Maddox &~~ ~~– we assume that the learning process described in Equation 2 is likewise unaffected by feedback delay.~~

~~Although the criterial-learning process described by Equation 2 is not affected by feedback delay, we assume that both the stimulus and the criterion representations drift randomly throughout the duration of time they are maintained in working memory. The representation of the criterion must always be maintained in working memory, whereas drift in the stimulus representation affects performance up until the feedback is presented, but not afterwards.~~

~~We modeled the drift in both the criterion and the percept by adding white noise~~

to their initial values. Specifically, we assumed that for all $t > \mathrm{T_S}$

$$c_n(t) = c_n(\mathrm{T_S}) + \eta_c \epsilon(t),$$

and

$$x_n(t) = x_n(\mathrm{T_S}) + \eta_x \epsilon(t),$$

where $\epsilon(t)$ is white noise and η_c and η_x are parameters that determine the amount of drift over time. This model predicts that at the time of feedback, the response criterion $c_n(\mathrm{T_F})$ will be normally distributed with mean $c_n(\mathrm{T_S})$ and variance $\mathrm{t_{FD}}\eta_c^2$. Similarly, at the time when the stimulus that defines the next trial is presented, the criterion $c_n(\mathrm{T_{S^+}}) = c_{n+1}(\mathrm{T_S})$ is normally distributed with mean $c_n(\mathrm{T_S})$ and variance $(\mathrm{t_{FD}} + \mathrm{t_{ITI}})\eta_c^2$. The predictions for the percept are similar. Specifically, at the time when feedback is presented, the percept $x_n(\mathrm{T_F})$ is normally distributed with mean $x_n(\mathrm{T_S})$ and variance $(\mathrm{T_F} - \mathrm{T_S})\eta_x^2$the updating of the criterion is sensitive to the feedback delay (yes or no).

Like the time-dependent drift model, the delay-sensitive learning model assumes that decisions are based on comparing the current stimulus to a stored referent. However, unlike the time-dependent drift model, the criterion remains stable over time and does not drift. For this reason, the memory system used to store the criterion may be different from working memory. In other words, $c_n(t) = c_n$ for all $t > \mathrm{T_S}$. This model also assumes that the magnitude of error-driven updates to the criterion decreases in proportion to the length of the feedback delay. As a result, when feedback is delayed, the system becomes less responsive to errors, leading to slower learning compared to immediate feedback conditions.

This model also assumes that the percept does not drift over time . Instead, it models perceptual noise as a time-invariant perturbation. Specifically, the delay-sensitive learning model assumes the observer uses the following decision rule:

$$\text{Respond } R_n = \begin{cases} \text{A}, & \text{if } \; x_n(\mathrm{T_S}) + \epsilon_p \le c_n \\ \text{B}, & \text{if } \; x_n(\mathrm{T_S}) + \epsilon_p > c_n, \end{cases}$$

where the time-invariant perceptual noise term ϵ_p is normally distributed with mean 0 and variances σ_p^2. This same value of perceptual noise corrupts all future values of the percept. Therefore, $x_n(t) = x_n(\mathrm{T_S}) + \epsilon_p$ for all $t > \mathrm{T_S}$.

The delay-sensitive learning model assumes that longer feedback delays slow criteral learning. Specifically, the model assumes that if negative feedback is received, the criterion is updated as follows:

$$c_{n+1} = c_n + \frac{\alpha}{t_{\mathrm{FD}}}[x_n(\mathrm{T_F}) - c_n].$$

The scale factor $(1/t_{\mathrm{FD}})$ captures the notion that the length of the feedback delay slows the *rate* of procedural learning.

The reinforcement-learning model gradually associates responses with stimuli, and therefore does not include an explicit representation of the response criterion. The model is based on an actor-critic architecture and uses reinforcement learning to form

~~stimulus-response associations (Sutton & Barto, 1998).~~

~~This model assumes that the perceptual representation of each stimulus is defined by the pattern of activation across 25 sensory units that are characterized by overlapping tuning curves. Each unit is maximally excited by one specific stimulus, which we call the unit's preferred stimulus. Specifically, the activation of the i^{th} sensory unit on trial n is given by~~

$$A_i(n) = \exp\left(\frac{-\left[d_{i,S_n}\right]^2}{\sigma^2}\right)$$

~~where d_{i,S_n} is the (Euclidean)distance between the preferred stimulus value of the i^{th} sensory unit and the value of the stimulus that was presented on trial n, and σ is a constant that increases with perceptual noise.~~

~~The model includes two decision or actor units – one associated with each of the two possible responses. Initially, each sensory unit is connected to both actor units with some random connection strength. Let $\omega_{iJ}(n)$ denote the strength of the connection between sensory unit i and actor unit J (for J = A or B) on trial n. Then the activation in actor unit J on trial n, denoted by $V_J(n)$, equals~~

$$V_J(n) = \sum_i \omega_{iJ}(n) A_i(n)$$

~~Responses are generated by the decision rule:~~

$$\text{Respond } R_n = \begin{cases} A, & \text{if } V_A(n) > V_B(n) \\ B, & \text{if } V_A(n) \le V_B(n). \end{cases}$$

~~The connection strengths $\omega_{iA}(n)$ and $\omega_{iB}(n)$ are updated after feedback is received on each trial according to standard reinforcement learning rules (Sutton & Barto, 1998):~~

$$\omega_{iA}(n) = \omega_{iA}(n-1) + \frac{\alpha_{actor}}{t_{\text{FD}}}\delta(n-1)$$

$$\omega_{iB}(n) = \omega_{iB}(n-1) + \frac{\alpha_{actor}}{t_{\text{FD}}}\delta(n-1),$$

~~where α_{actor} is the learning-rate of the actor, and $\delta(n-1)$ is the reward prediction error on trial $n-1$. Note that as in the delay-sensitive learning model, the learning rate is scaled by the inverse of the feedback delay . Much evidence suggests that this type of stimulus-response learning is mediated largely within the striatum, and is facilitated by a dopamine (DA)mediated reinforcement learning signal that is time dependent (e.g., Valentin et al., 2014). Specifically, the dopamine signal generated by positive feedback appears to peak at around 500 ms after feedback and then decay back to baseline levels within 2 or 3 s (Yagishita et al., 2014). As a result, synaptic plasticity at cortical-striatal synapses is attenuated with increasing feedback delays (Yagishita et al., 2014). The scaling of α_{actor} by $1/t_{\text{FD}}$ models this phenomenon.~~

~~The reward prediction error on trial $n-1$ is defined as the value of the obtained reward, denoted by $R(n-1)$ minus the value of the predicted reward, denoted by $P(n-1)$:~~

$$\delta(n-1) = R(n-1) - P(n-1).$$

~~The predicted reward on trial n is determined by the critic via~~

$$P(n+1) = P(n) + \frac{\alpha_{critic}}{t_{\mathrm{FD}}}\delta(n),$$

~~where α_{critic} is the learning rate of the critic. This learning rate is again scaled by the inverse of the feedback delay.~~

~~We investigated how changing the duration of the feedback delay and ITI affect criterial learning for each of these three models. More specifically, we simulated performance of each model in a categorization task that included two categories of stimuli that varied on a single stimulus dimension and that could be categorized perfectly by comparing each stimulus to an appropriate value of the response criterion (We used the same experimental design as in Experiment 1, so see those methods for more details). Our simulations included three experimental conditions: Delayed Feedback, Long ITI, and Control. The Delayed-Feedback condition included a long feedback delay (3.5 s) and a short ITI (0.5 s). The Long-ITI condition matched the total trial duration of the Delayed-Feedback condition but with a short feedback delay (0.5 s) and a long ITI (3.5 s). The Control condition included both a short~~ Four different models assumed that learning is of SR associations and that the criterion therefore has no psychological meaning. These were created by factorially combining two binary assumptions: the percept drifts randomly over time (yes or no) and SR strengthening is sensitive to feedback delay (~~0.5 s)and a short ITI (0.5 s). Each simulation continued for 200 trials or until the model responded correctly for 12 trials in a row. For each set of parameter values, we simulated the model 100 times and averaged the results, yielding one observed performance metric measured in trials-to-criterion for each condition. All scores were then normalized by the largest observed value, and as a result, all axes in the figures that describe the simulation results range from zero to one.~~

~~Our approach was to generate predictions for each of the three experimental conditions across a wide range of parameter values. We then used~~ yes or no). Next, using a technique called parameter-space partitioning ~~(PSP) to evaluate the performance of each model (PSP; Pitt et al., 2006). PSP calculates the proportion of the parameter space where a model makes specific qualitative predictions. We focused on four such predictions. The first was that learning under feedback delay would be slower than in the other two conditions. The second was that learning in the Long ITI condition would be slower than in the other two. The third was that learning in the control condition would be slower than in either of the other two conditions. The fourth was a catch-all category encompassing any other possible patterns of results. For each model, the PSP analysis quantified the proportions of the explored parameter space where the model made these qualitative predictions.~~ (Pitt et al., 2006), we investigated whether each model could account for the qualitative results of Experiments 1 and 2. The results of these analyses reinforced the conclusion that criterial learning may have strong associative underpinnings, even in tasks that appear to rely on explicit, rule-based reasoning.

~~We investigated predictions of this model across a wide range of values for the parameters η_x, η_c, and α. Specifically, in the case of α, we stepped through every value in the interval $[0, 1]$, with a step size of .01. In the case of both η_x, and η_c, we searched over the interval $[0.1, 5]$, with a step size of 0.5. The results are shown in Figure **??**.~~

~~**A:** The proportion of parameter space where the time-dependent drift model predicted each of four qualitative patterns: (1) slower learning under feedback delay, (2) slower learning in the Long-ITI condition, (3) slower learning in the control condition, and (4) any other pattern. **B:** Simulated trials-to-criterion for each condition. All scores were normalized by the largest trials-to-criterion value within a given parameter set, so all axes range from zero to one. **C:** Boxplot of the parameter ranges leading to each PSP pattern. All parameter values were normalized by the largest value in the search range, so the ordinate ranges from zero to one for all parameters. **D:** Scatter plot of the parameter ranges associated with each PSP pattern. *Note:* In all panels, color indicates the PSP pattern.~~

~~Figure **??**A shows that the model essentially always predicts that increasing the feedback delay or the ITI have similar detrimental effects on learning. The model makes this prediction because it assumes that the criterion drifts during both the feedback delay and during the ITI. So the critical variable for the model is the time between the response on trial n and the presentation of the stimulus on trial $n + 1$. The model assumes that the criterion drifts during this entire time, and therefore how this interval is divided between feedback delay and ITI is relatively unimportant to the model predictions.~~

~~We simulated the performance of the delay-sensitive learning across a wide range of parameter values for σ_p and α. Specifically, in the case of α we stepped through every value in the interval $[0, 1]$, with a step size of .1. In the case of σ_p, we searched over the interval $[0.1, 5]$, with a step size of 0.1.~~

~~The results are shown in Figure **??**. Note that virtually all combinations of the parameter values predict that feedback delay impairs criterial learning more than increasing the ITI. This makes sense because Equation shows that the delay-sensitive learning model predicts that increasing the feedback delay (i.e., increasing t_{FD}) will impair learning – for any value of $\alpha > 0$. Large values of σ_p will also impair performance because of distortion to the percept, but this interference will be the same in all conditions.~~

~~**A:** The proportion of parameter space where the delay-sensitive learning model predicted each of four qualitative patterns: (1) slower learning under feedback delay, (2) slower learning in the Long-ITI condition, (3) slower learning in the control condition, and (4) any other pattern. **B:** Simulated trials-to-criterion for each condition. All scores were normalized by the largest trials-to-criterion value within a given parameter set, so all axes range from zero to one. **C:** Boxplot of the parameter ranges leading to each PSP pattern. All parameter values were normalized by the largest value in the search range, so the ordinate ranges from zero to one for all parameters. **D:** Scatter plot of the parameter ranges associated with each PSP pattern. *Note:* In all panels, color indicates the PSP pattern.~~

~~We investigated the effects on performance predicted by the reinforcement-learning model of three parameters – the perceptual-noise variance σ^2 and the actor and critic learning rates (i.e., α_{actor} and α_{critic}, respectively). In the case of both α_{actor} and α_{critic}, we~~

~~stepped through every value in the interval $[0, 0.2]$, with a step size of .01. We constrained our search over these parameters to this interval because reinforcement-learning models are prone to instability at very high learning rates (Sutton & Barto, 1998). As evidence of this, at higher learning rates, the model failed to learn with any consistency – that is, in most cases, it failed to reach the learning criterion (12 correct responses in a row) within the allowable 200 trials. In the case of σ, we searched over the interval $[1, 10]$, with a step size of 1. As with the other models, all scores were normalized by the largest observed value.~~

~~The results are described in Figure 9. Note this model predicts that delayed feedback impairs criterial learning more than a long ITI across a wide volume of parameters settings. As can be seen in Figure 9C, the only exceptions tend to occur when the value of any of the three parameters approaches the maximum end of the range explored. The logic here is similar to the delay-sensitive learning model – that is, both models predict that increasing the feedback delay t_{FD} will impair learning – for all values of the learning rate.~~

~~**A:** The proportion of parameter space where the reinforcement-learning model predicted each of four qualitative patterns: (1) slower learning under feedback delay, (2) slower learning in the Long-ITI condition, (3) slower learning in the control condition, and (4) any other pattern. **B:** Simulated trials-to-criterion for each condition. All scores were normalized by the largest trials-to-criterion value within a given parameter set, so all axes range from zero to one. **C:** Boxplot of the parameter ranges leading to each PSP pattern. All parameter values were normalized by the largest value in the search range, so the ordinate ranges from zero to one for all parameters. **D:** Scatter plot of the parameter ranges associated with each PSP pattern. *Note:* In all panels, color indicates the PSP pattern.~~

~~We investigated three criterial-learning models that make different predictions about how increases in the~~

Experiment 1

Experiment 1 used a one-dimensional category-learning task to investigate how feedback delay and ITI affect ~~learning. The time-dependent drift model predicts that the criterion always drifts, so the critical variable is the sum of the feedback delay and ITI. How this sum is divided into separate feedback delay and ITI time intervals has little effect on the model's predictions. In contrast, the delay-sensitive learning model predicts that feedback delays are necessarily more detrimental to performance than increases in the ITI. Finally, the reinforcement-learning model also predicts that, in general, feedback delays should impair learning more than long ITIs, although this model is somewhat more flexible than the delay-sensitive learning model and can account for a small or null effect of feedback delay within a restricted region of its parameter space. The obvious next question is how human criterial learningis affected by these independent variables. Experiments 2 and 3 were designed to address this question, and therefore also to test the predictions of the three models.~~

~~Experiment 2 investigated how feedback delay and the length of the ITI affect criterial learning in humans. As a result, it also provides rigorous tests of the predictions of the time-dependent drift model, the delay-sensitive learning model, and the reinforcement-learning model.~~

criterial learning. The stimuli in ~~Experiment 2~~ this experiment were circular sine-wave gratings that varied across trials in bar width and bar orientation. These stimuli were divided into two categories according to their value on one of the two dimensions. In other words, the optimal strategy was to set a response criterion on the single relevant dimension, and then choose a categorization response based on whether the value of the presented ~~stimuli~~ stimulus on this dimension was larger or smaller than the criterion value.

The experiment isolated criterial learning by (1) explicitly instructing participants ~~on~~ about the relevant stimulus dimension and rule structure (e.g., thick bars are "A", thin bars are "B"), and (2) eliminating all variability along the irrelevant stimulus dimension. In other words, the instructions identified the relevant stimulus dimension, and as a result, the only learning required was criterial learning.

Each participant practiced each of 14 one-dimensional category-learning tasks, or problems, until they responded correctly on 9 of the 10 previous trials, at which point the problem changed. The 14 different category structures are described in Figure 1. The relevant dimension was bar thickness in problems 1 – 7 and bar angle in problems 8 – 14. Each problem included stimuli in two distinct clusters that varied over a restricted range of the relevant dimension. Critically, the optimal criterion value varied from problem-to-problem with respect to its position within the stimulus range. For example, for some problems the optimal criterion was below the midpoint of the range and for other problems it was above the midpoint. Feedback delay and ITI were varied across three conditions. In the Delayed-Feedback condition, feedback was delayed after the observer responded (by 3.5 s) and the ITI was short (500 ms). In the Long-ITI condition, the feedback delay was short (500 ms) and the ITI was long (3.5 s). Finally, in the Control condition, feedback delay and the ITI were both 500 ms.

Method

Apparatus

All experiments were performed in a dimly lit room. Participants sat approximately 24" from a 17" $\times$ 11" monitor running at a resolution of 1680 $\times$ 1050 pixels. Participants made category judgments by pressing the 'd' or 'k' keys on a standard computer keyboard for 'A' or 'B' choices, respectively. Stickers with bold print 'A' or 'B' were placed on the appropriate keys.

Stimuli and Categories

Stimuli were circular sine-wave gratings that varied in bar width and bar orientation, drawn from various 1-dimensional uniform distributions specific to the current category problem. We first defined an arbitrary 2-dimensional $[0 - 100, 0 - 100]$ stimulus space, and then split each dimension of this space into 7 bins of width 14 units each. Each (x, y) pair from this arbitrary stimulus space was converted to a grating according to the nonlinear transformations defined by Treutwein et al. (1989), which roughly equate the salience of each dimension~~(for~~. For details, see also Crossley and Ashby (2015)~~)~~.

The structure of the various criterial-learning tasks is illustrated in Figure 1. Each criterial-learning problem was created by first ~~randomly~~ selecting a relevant dimension~~, and then randomly selecting~~. This was spatial frequency for the problems 1 through 7 and

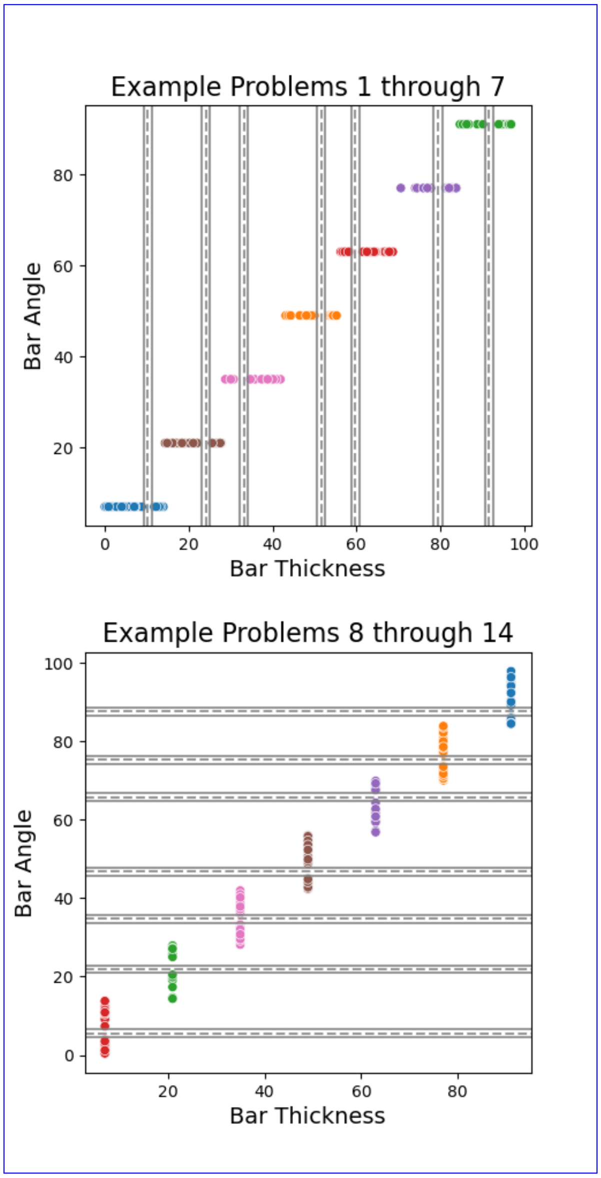

Figure 1

Category sample space. Different colors represent different category problems. Dashed lines are category boundaries (criterion) and the surrounding solid lines mark the no-stimulus region in which no stimuli were sampled.

orientation for problems 8 through 14. We then randomly selected one of the 7 bins defined on that dimension. Each bin was also associated with a corresponding unique value on the irrelevant dimension. We buffered the to-be-learned response criterion by 10% of total bin width on either side with a no-stimulus region. Random uniform samples from the remaining eligible region of each bin were then selected and presented to the participant until 9 correct responses out of any 10 responses in a row advanced the participant to the next problem. Note that every category problem was a simple one-dimensional rule in which optimal accuracy was 100%. Note also that the relative location of the optimal response criterion varied across problems.

~~Category sample space. Different colors represent different category problems. Dashed lines are category boundaries (criterion) and the surrounding solid lines mark the no-stimulus region in which no stimuli were sampled.~~

Procedure

There were three conditions (described in detail in Table 1). In the Delayed-Feedback condition, feedback was delayed 3.5 s after the response and the ITI was 0.5 s. In the Long-ITI condition, feedback was delayed 0.5 s after the response and the ITI was 3.5 s. Finally, in the Control condition, feedback was delayed 0.5 s after the response and the ITI was 0.5 s.

Each participant completed a series of one-dimensional category-learning tasks or problems, which are described in Figure 1. Each problem included stimuli in two distinct clusters that varied over a restricted range of the relevant dimension. Critically, the optimal criterion value varied from problem-to-problem with respect to its position within this range. For example, for some problems the optimal criterion was below the midpoint of the range and for other problems it was above the midpoint.

Participants were explicitly told the relevant dimension for each problem, as well as the generic response mapping (e.g., thick bars = “A”, thin bars = “B”). Figure 2 shows the structure of an example trial, along with an example of a typical category structure. All trials in every condition included a 500 ms fixation cross, a response-terminated stimulus, a circular white-noise mask, corrective feedback, and an inter-trial interval (ITI) that varied according to condition. The text ‘Correct’ was displayed in centered, large green font after correct responses, and the text ‘Incorrect’ was displayed in centered, large red font after incorrect responses.

Participants practiced each problem until they responded correctly on 9 of the previous 10 trials. At this point, the problem changed. Each participant completed as many problems as possible in 512 trials or until they had been in the lab for 60 minutes (including time to acquire consent and give instructions), at which point the session was terminated.

Participants

Fifty-nine participants participated in Experiment ~~2.~~ 1. All were UCSB undergraduates and received course credit for their participation. All had normal or corrected to normal vision. We randomly assigned each participant to one of three conditions (target $N > 16$ per condition based on similar previous research): Delayed Feedback ($N = 20$); Long ITI ($N = 21$), Control ($N = 17$). All participants gave written informed consent

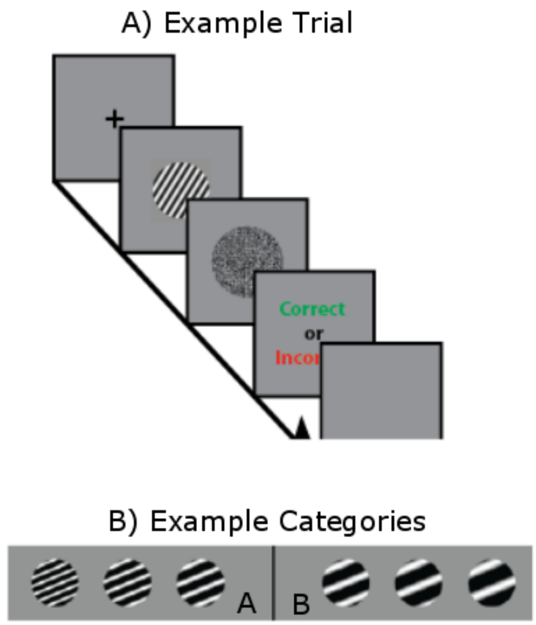

Figure 2

Example trial and category problem. A) Events that occurred on each trial. B) An example of a typical category structure.

Table 1

Durations (in s) of Trial Events in each Condition.

Conditions	Stim	Mask	FB	ITI
Control	RT	0.5	1.0	0.5
Delayed Feedback	RT	3.5	1.0	0.5
Long ITI	RT	0.5	1.0	3.5

before participating in the study. All experimental protocols were approved by the University of California at Santa Barbara Human Subjects Committee in the Office of Research Development and Administration.

Results

~~Figure ?? shows a~~ Both panels of Figure 3 show the same relative frequency histogram of the trials-to-criterion ~~observed across all three~~ collapsed across all participants and conditions. The histogram shows that the majority of participants were able to learn each problem on average in less than 100 trials. However, this histogram also shows that a subset of participants required many more trials to learn each problem. Given that ~~each problem is very simple and that~~ participants were explicitly instructed about the relevant dimension of each problem~~is explicitly instructed to participants, it is~~, the task was relatively

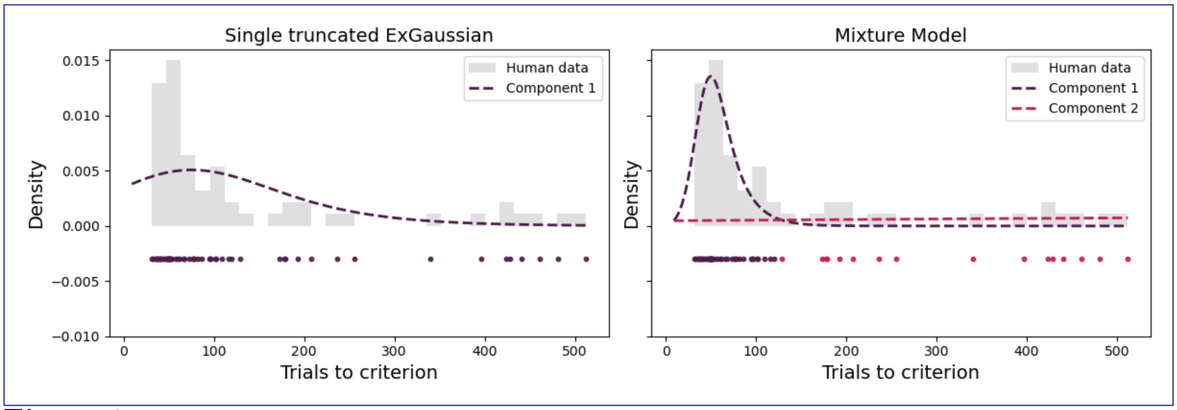

Figure 3

Relative frequency histogram of the trials-to-criterion ~~observed~~ collapsed across all three conditions and participants in Experiment ~~2.~~ 1. The ~~black line are~~ left panel shows the ~~predictions of~~ best-fitting single ExGaussian distribution and the right panel shows the best-fitting ~~one-component Gaussian Mixture Model to these data~~two-component ExGaussian mixture model. In the right panel, ~~whereas~~ the ~~red~~ blue dashed line ~~represents~~ is the ~~best fit of~~ first higher performance component, the ~~two-component Gaussian Mixture Model~~orange dashed line is the second lower performance component. The gray solid line is the mixture distribution that combines these two components. Rug plots along the x-axes show the individual data colored by their assignment to either the higher or lower performance component. The two-component model provided a significantly better fit according to AIC, BIC, and a likelihood ratio test.

simple and as a result, it seems likely that the participants ~~in~~ represented by the tails of this distribution did not pay attention to the instructions, were not motivated to learn the task, were distracted by other factors, or should be considered outliers for some other reason. In our experience, it is not unusual for a small subset of participants to perform very poorly in experiments of this kind. This typically reflects a lack of engagement or misunderstanding of the task instructions rather than an inability to perform the task itself.

~~We~~ We identified and removed from further analyses these outlier participants using the following algorithm. First, we modeled the distribution shown in Figure ~~?? using a Gaussian mixture model with two components~~3 using a mixture of two ExGaussian distributions,[1] each truncated to the range of possible trials-to-criterion values (i.e., [9, 512]). The first ~~component captured~~ distribution or component was assumed to model performance of participants with a relatively low mean number of trials-to-criterion, ~~while~~ whereas the second component ~~identified outliers according to the above rationale. We~~ was assumed to identify outlier participants. We estimated parameters of this model using maximum likelihood estimation with proper normalization on the tasks bounded support and compared them via the AIC and BIC goodness-of-fit statistics, and a parametric bootstrap likelihood-ratio test. We then compared the fit of this two-component model

[1] The ExGaussian is the distribution of the sum of two independent random variables – one with an exponential distribution and one with a normal or Gaussian distribution.

to a single-component ~~model using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC)~~ExGaussian model using AIC and BIC, both of which indicated a better fit for the two-component model (~~AIC: 662.81 vs. 735.77; BIC: 673.11 vs. 739.89). Additionally, a likelihood ratio test~~ $\text{AIC}_1 = 692.15$ vs. $\text{AIC}_2 = 652.81$; $\text{BIC}_1 = 698.33$ vs. $\text{BIC}_2 = 667.23$). The bootstrap likelihood-ratio test also confirmed that the two-component model provided a significantly better fit than the one-component model (~~$\chi^2(3) = 78.96$, $p < .001$). Outlier participants were then~~ $D = 42.84$, $p_{\text{boot}} \approx .01$). Next, for each participant, we used the best-fitting two component model to compute the likelihood that their performance was best described by each component. Participants whose data were estimated to most likely belong to the low-performance distribution were classified as outliers and excluded from further analysis. After making these ~~exlcusions~~exclusions, we were left with the following sample sizes: Control condition ($N = 11$); Delayed-Feedback condition ($N = 14$); Long-ITI condition ($N = 17$).

Figure 4A shows the mean trials-to-criterion for all non-excluded participants in each condition of Experiment ~~2.~~ 1. A one-way ANOVA revealed a significant effect of condition ($F(2, 39) = 6.17$, $p < .01$, $\eta^2 = .24$) and planned comparisons revealed that performance in the Delayed-Feedback condition was significantly worse than in either the Control condition ($t(23.00) = 2.70$, $p < .05$, $d = 1.06$) or the Long-ITI condition ($t(20.96) = 2.94$, $p < .01$, $d = 1.10$). Performance in the Control and Long-ITI conditions did not differ significantly from each other (~~$t(17.99) = 0.22$~~$t(17.99) = .22$, $p = .83$~~).~~ , $d = .09$).[2]

Figure 4B shows the mean number of problems solved in each condition of Experiment ~~2.~~ 1. A one-way ANOVA revealed a significant effect of condition ($F(2, 39) = 5.77$, $p < .01$, $\eta^2 = .23$). Planned comparisons revealed that performance in the Delayed-Feedback condition was significantly worse than in the Control condition ($t(18.54) = -3.28$, $p < .01$, $d = -1.21$) and worse than in the Long ITI condition ($t(21.60) = -2.14$, $p < .05$, $d = -.80$). Performance in the Control and Long-ITI conditions did not differ significantly from each other ($t(25.97) = -1.50$, $p = .15$, $d = -.53$).

~~Experiment 3 was designed to reinforce the findings of Experiment~~

Experiment 2

According to a popular model of one-dimensional rule-based category learning (i.e., COVIS; Ashby et al., 1998), successful learning requires at least three separate cognitive processes: selecting a rule for testing, switching from one candidate rule to another (following negative feedback), and criterial learning. Experiment 1 isolated criterial learning from this set. Experiment 2 ~~by using the same design but with stimuli that have binary-valued dimensions, thereby eliminating~~ complements Experiment 1 by using a task that required rule selection and rule switching, but no criterial learning. ~~Unlike~~

Experiment 2 ~~, we did not explicitly instruct participants on the relevant stimulus dimension in Experiment 3. While Experiment 2 isolated criterial learning from rule selection and rule switching, this experiment isolates rule selection and switching from criterial learning . Therefore, Experiment 3 was~~ used colored geometric figures as stimuli, which varied across trials on six binary dimensions (i.e., see Figure 5B). As in Experiment 1, only one dimension was relevant. Unlike Experiment 1, however, participants were given no

[2] Including the excluded outliers in these analyses eliminates the difference between conditions.

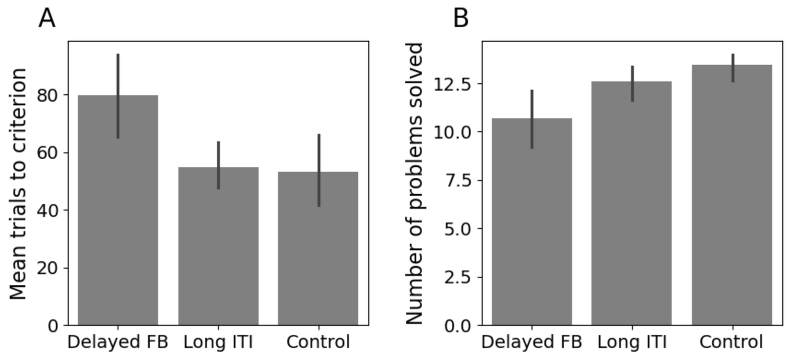

Figure 4

***A**: Mean ~~trials to criterion~~ trials-to-criterion in each condition of Experiment ~~2.~~ 1. Error bars are standard errors of the mean. **B**: Mean number of problems solved in each condition of Experiment ~~2.~~ 1. Error bars are standard errors of the mean.*

instruction as to which dimension was relevant. Therefore, this task required participants to discover the relevant dimension. However, after this dimension was identified, no criterial learning was required because each stimulus always had one of two possible values on the relevant dimension and participants were instructed that there were two categories. Experiment 2 was therefore designed to test whether the impaired learning that we observed in Experiment ~~2~~ 1 when the feedback was delayed can be attributed principally to criterial learning or whether it could be due to some more general category-learning process.

Method

Apparatus

The apparatus was the same as in Experiment ~~2.~~ 1.

Stimuli and Categories

The stimuli consisted of colored geometric figures presented on a colored background. These varied across six binary dimensions: the number of items (either one or two), the size of the items (small or large), the color of the items (yellow or blue), the shape of the items (circle or square), the texture of the background (smooth or rough), and the orientation of the background (horizontal or tilted by 20 degrees). This combination resulted in a total of 64 unique stimuli (2^6). An example trial and stimuli are shown in Figure 5. In all conditions, the order of stimulus presentation was fully randomized for each participant, and the relevant dimension for each problem was selected randomly without replacement from the set of six dimensions.

Procedure

Participants received detailed instructions about the task. However, unlike in Experiment ~~2~~1, they were not informed of the relevant dimension or the correct rule. The trial timing for the Delayed-Feedback and Long-ITI conditions in Experiment ~~3~~ 2 were the same

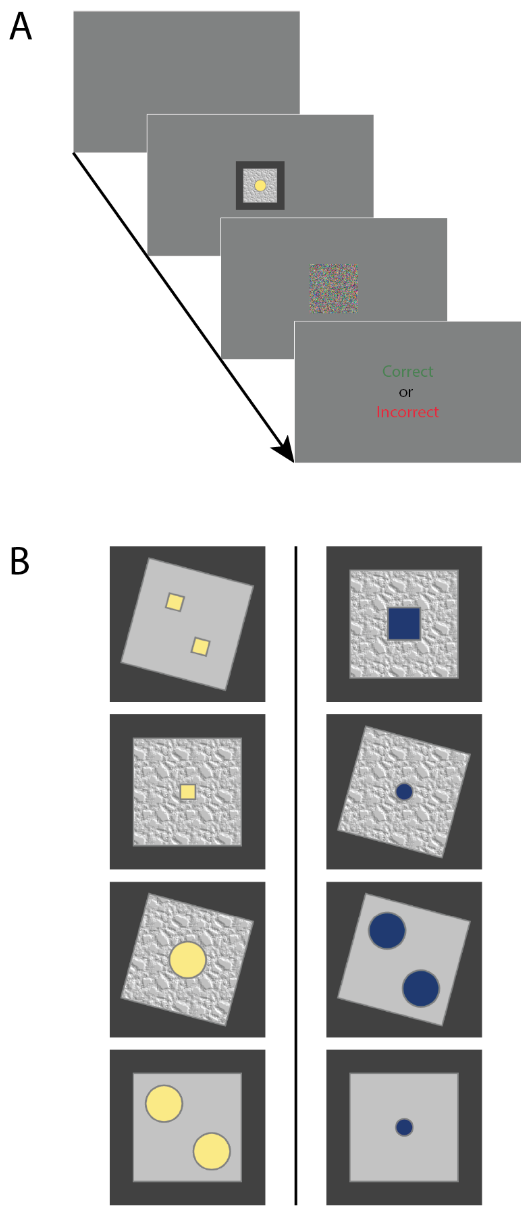

Figure 5

A: *An example trial from Experiment ~~3.~~ 2.* ***B:*** *Eight example stimuli (out of 64 total stimuli) from a possible one-dimensional rule problem. In this case the correct rule is 'A' if the shape color is yellow and 'B' if the shape color is blue. The solid line represents the category boundary.*

as in Experiment ~~2. Experiment 3~~ 1. Experiment 2 did not include a condition that was analogous to the Control condition of Experiment ~~2.~~ 1. Participants practiced each problem until they responded correctly on 12 consecutive trials. At this point, the problem changed. Each participant completed as many problems as possible in 600 trials or until they had been in the lab for 30 minutes (including time to acquire consent and give instructions), at which point the session was terminated.

Participants

Thirty-four participants participated in Experiment ~~3.~~ 2. All were Macquarie University undergraduates and received course credit for their participation. All had normal or corrected to normal vision. We randomly assigned each participant to one of two conditions: Delayed Feedback ($N = 17$) or Long ITI ($N = 17$). All participants gave written informed consent before participating in the study. All experimental protocols were approved by the Macquarie University Human Research Ethics Committee (protocol number: 52020339922086).

Results

Figure ~~?? shows a~~ 6 shows the trials-to-criterion relative frequency histogram ~~of the trials-to-criterion observed in both conditions. The~~ collapsed across participants and conditions. As in Experiment 1, the histogram shows that the majority of participants were able to learn each problem ~~on average in less than~~ in well under 100 trials. Unlike Experiment ~~2,~~ 1, however, there were no highly suspicious outliers in this data set. Nevertheless, for symmetry with Experiment ~~2~~1, we performed the same ~~Gaussian~~ ExGaussian mixture model analysis as used there. The two-component model provided a slightly smaller AIC value (~~317.40 vs . 317.78), but a slightly larger~~ 317.07 vs 325.38), and a slightly smaller BIC value (~~325.03 vs . 320.83~~327.75 vs 329.96). The bootstrap likelihood-ratio test was not significant (~~$\chi^2(3) = 6.38$, $p = .09$). We therefore did not exclude~~ $D = 13.05$, $p_{\text{boot}} \approx .08$). Given the weaker evidence for multiple components in this data set compared to Experiment 1, and the fact that there were no obvious outliers in this data set, we concluded that there was insufficient evidence to justify excluding any participants from further analysis~~in Experiment 2.~~ .

Figure 7A shows the mean trials-to-criterion in each condition of Experiment ~~3.~~ 2. An independent-samples t-test revealed no significant effect of condition (~~$t(32.0) = 0.69$, $p = .50$, $\eta^2 = .01$~~$t(32.0) = 1.37$, $p = .18$, $d = .47$). The mean number of problems solved by each participant is shown in Figure 7B. An independent-samples t-test revealed no significant effect of condition (~~$t(32) = -0.49$, $p = .63$, $\eta^2 = .01$~~$t(32.0) = -1.99$, $p = .06$, $d = -.68$).

~~Experiment 2 clearly~~

Discussion of Experiments 1 and 2

Experiment 1 showed that a short feedback delay of only 3.5 s slowed criterial learning. In contrast, increasing the ITI ~~to this same value~~ from 500 ms to 3.5 s had no effect on learning. The task we used was one-dimensional category learning, but we isolated criterial learning by instructing participants about the optimal strategy. Specifically, we told

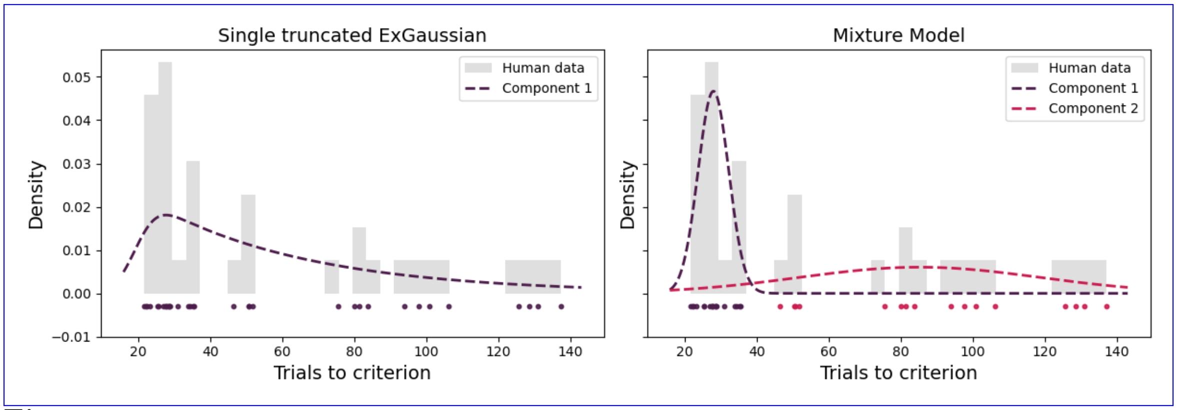

Figure 6

Relative frequency histogram of the trials-to-criterion ~~observed~~ collapsed across all ~~three~~ participants and conditions in Experiment ~~3.~~ 2. The ~~black line represents~~ left panel shows the best ~~fit of the~~ fitting one-component ~~Gaussian Mixture Model~~ExGaussian distribution, and the ~~red line represents~~ right panel shows the ~~best fit~~ two components of the ~~two-component Gaussian Mixture Model~~best-fitting two component ExGaussian mixture model.

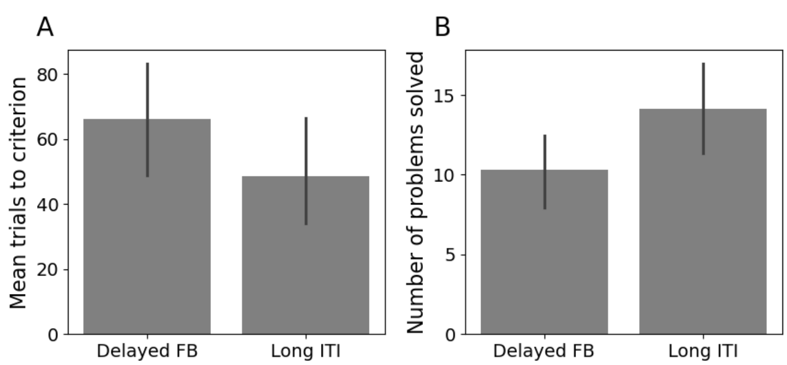

Figure 7

A: Mean trials to criterion in each condition of Experiment ~~3.~~ 2. Error bars are standard errors of the mean. **B**: Mean number of problems solved in each condition of Experiment ~~3.~~ 2. Error bars are standard errors of the mean.

them which stimulus dimension was relevant and that there was no trial-by-trial variability on the irrelevant dimension. Although these results strongly suggest that the feedback delay had its interfering effects on criterial learning, Experiment ~~3~~ 2 was designed to ~~confirm~~ complement this inference. The goal here was to examine the effects of the same feedback delays and ITIs on performance in a one-dimensional category-learning task that did not require any criterial learning, but did require category-learning processes that are thought to mediate rule discovery (e.g., rule selection and switching). If the feedback delay effects observed in Experiment ~~2~~ 1 were acting on some general category-learning skill then we should have seen the same interfering effects of feedback delay in Experiment ~~3.~~ 2. However, if the feedback-delay interference of Experiment ~~2~~ 1 was operating selectively on criterial learning, then it should disappear in Experiment ~~3~~2, since no criterial learning was required. Our results strongly supported this latter prediction. Therefore, Experiments ~~2 and 3~~ 1 and 2 together strongly suggest that criterial learning is impaired by feedback delays and is relatively unaffected by the length of the ITI.

~~Two prior studies investigated similar issues. First, Ell and colleagues reported that feedback delay and also a concurrent memory-scanning task each impaired~~ Whereas the results of Experiment 1 seem clear, the Experiment 2 results are weaker since they suggest a null result. Although it impossible to rule out the possibility that Experiment 2 missed some small effect of feedback delay due to insufficient power, the null results found there are consistent with a wide variety of other evidence. In particular, a variety of other studies have reported that feedback delays of up to 10 s do not significantly slow one-dimensional rule-based ~~category learning (Ell et al., 2009). Based on these results, they hypothesized that criterial learning may be tied to working-memory capacity and therefore to explicit cognitive mechanisms. However,~~ learning (Dunn et al., 2012; Ell et al., 2009; Maddox et al., 2003; Maddox & Ing, 2005; Worthy et al., 2013). In addition, many studies have reported evidence that one-dimensional rule-based learning recruits working memory and executive attention, but not procedural learning (for reviews, see, e.g., Ashby & Valentin, 2017; Ashby et al., 2020). So the results of Experiment 2 are consistent with all of these studies. On the other hand, several of these feedback-delay studies did not use binary-valued stimulus dimensions, so some criterial learning was ~~confounded with rule selection and task difficulty in their design. Furthermore, Ell et al. (2009) found that feedbackdelay only impaired learning when working memory demand was high – that is, when participants had to learn more than one response criterion for optimal performance. In contrast, we found that feedback delay impairs learning even when working memory demands are trivial (participants never have to keep in mind more than one criterion)~~required. If so, then why did they all report no effect of feedback delay? It is important to note that criterial learning was not the focus of any of these studies, and as a result, in all of these previous studies the optimal criterion was always set exactly midway between the category prototypes, which makes criterial learning trivial. For example, under these conditions, criterial learning might not even require feedback. The unsupervised category-learning experiments reported by Ashby et al. (1999) provide strong support for this because all of their rule-based participants learned the correct criterion (which was midway between the category means), even though the task was completely unsupervised.

We know of only one previous study that reported impaired one-dimensional rule-based category learning when feedback was delayed (Ell et al., 2009). However, in this study, a feedback delay only impaired performance in a condition with four contrasting categories that required learning three different response criteria. So this task was considerably more complex than any of the other feedback-delay studies, which all used two contrasting categories that required the learning of only one response criterion.

~~Second~~One other study with similar goals to ours studied how manipulations of category base rates affect criterial learning in one-dimensional rule-based categorization. Specifically, Bohil and Wismer (2014) reported that ~~the effects of unequal base rates on~~ changes in criterion placement in ~~rule-based category learning were diminished under delayed feedback and also under an observational trainingprotocol that is also thought to selectively impair basal ganglia-dependent associative-learning mechanisms~~response to changes in base rate were diminished when feedback was delayed (and also by observational training). These results are consistent with our findings. However, because they were obtained by manipulating base rates, it is difficult to conclude that the delay affected criterial learning, *per se*. For example, consider a simple two-stage model in which the first stage learns the base rates and the second stage uses what the first stage learned to adjust the response criterion appropriately. The Bohil and Wismer (2014) results are also consistent with the hypothesis that the feedback delay impaired the first of these stages but not the second. Our results strongly suggest that feedback delays impair criterial learning (and therefore the second of these hypothetical stages).

~~We developed three novel models of criterial learning that were designed to explore how feedback delay and ITI duration affect criterial learning . Two of these models assumed that decisions are made~~

New Computational Models of Criterial Learning

The results of Experiments 1 and 2 seem to suggest that criterial learning is a form of procedural learning, which previous research suggests is of SR associations (Ashby & Waldron, 1999; Casale et al., 2012). If so, then the criterion is simply the sensory, perceptual, or cognitive value that separates representations associated with contrasting responses. However, before accepting this conclusion, it is important to ask whether our results might also be compatible with an account in which the criterion is held is some form of working or short-term memory and updated trial by trial.

To address this question, we developed computational models of criterial learning within two broad architectures and examined their sensitivity to ITI duration and feedback delay. The architectures differed in whether they assume that the response criterion is explicitly represented in memory. In the first architecture, the criterion is stored in memory, and responses are generated by comparing the current percept ~~to a stored referent, or criterion , and that the remembered value of this criterion is updated following error feedback via a gradient-descent learning rule. The first~~ with this stored referent. The criterion is updated after error feedback using a gradient-descent rule. Models in this class can exhibit sensitivity to ITI and feedback delay in three distinct ways: (1) perceptual drift, where stimulus representations change over time, (2) criterial drift, where the stored criterion shifts over time, and (3) updating that is sensitive to feedback delays, where the magnitude of criterion updates decreases with longer feedback delays. In the second architecture, no explicit criterion is stored. Instead, reinforcement learning forms direct stimulus–response associations that drive performance. Sensitivity to ITI and feedback delay arise from perceptual drift and/or delay-sensitive updating, but criterial drift is not possible since no criterion is represented.

For each of these models, ~~the time-dependent drift model, posits that both the criterion and the percept~~ denote the time of stimulus presentation on the current trial by $\mathrm{T_S}$, the response time by $\mathrm{T_R}$, the time when feedback is displayed by $\mathrm{T_F}$, and the time when the stimulus that begins the next trial is displayed by $\mathrm{T_{S+}}$. Then note that the feedback delay equals $\mathrm{t_{FD}} = \mathrm{T_F} - \mathrm{T_R}$ and the duration of the ITI equals $\mathrm{t_{ITI}} = \mathrm{T_{S+}} - \mathrm{T_F}$.

Models that assume a stored criterion

These models assume that the observer constructs a response criterion, maintains it in memory, and makes decisions by comparing the percept to this criterion. They may further posit that both the stored criterion and the perceived stimulus value drift over time, ~~leading to similar impairments when feedback is delayed or the ITI is increased. The second model,~~ with the extent of drift increasing both within a trial and across consecutive trials. In addition, they may assume that the learning rate for updating the criterion decreases as the feedback delay increases.

Let $x_n(t)$ and $c_n(t)$ denote the values of the percept and the criterion, respectively, on trial n at time t. Then the decision rule on trial n is:

$$\text{Respond } R_n = \begin{cases} \mathrm{A}, & \text{if} \quad x_n(\mathrm{T_S}) \leq c_n(\mathrm{T_S}) \\ \mathrm{B}, & \text{if} \quad x_n(\mathrm{T_S}) > c_n(\mathrm{T_S}). \end{cases} \tag{1}$$

Note that this equation indicates that the response is made immediately upon stimulus presentation. We therefore are not explicitly modeling the time-extended cognitive processes that govern evidence accumulation and the corresponding response time.

If positive feedback is received, then the value of the criterion remains unchanged for the next trial (except for drift – see equation 3). The rationale here is that if the response is correct, then the observer has effectively gained zero information about how their criterion should be modified. If negative feedback is received, then the criterion is modified according to the standard model (Sutton & Barto, 1998):

$$c_n(\mathrm{T_F} + \Delta_t) = c_n(\mathrm{T_F}) + \frac{\alpha}{t_{FD}}[x_n(\mathrm{T_F}) - c_n(\mathrm{T_F})], \tag{2}$$

where α is a learning-rate parameter, and Δ_t is the time it takes to complete the updating. It is straightforward to show that the ~~delay-sensitive learning model , assumes that although neither the percept nor the criterion drift over time, optimal criterial learning requires immediate feedback~~ iterative equation 2 is equivalent to computing a weighted mean (weighted by recency) of the values of all percepts that occur on error trials (e.g., Ashby, 2017). This updating rule will gradually converge on the optimal criterion value (i.e., the value of the criterion that maximizes accuracy).

The term α/t_{FD} captures the notion that the magnitude of the update decreases as the feedback delay increases. As a result, ~~this model predicts that increasing the feedback~~

~~delay will slow the rate of criterial learning . The third model, the reinforcement-learning model , assumes~~ when feedback is delayed, the system becomes less responsive to errors, leading to slower learning compared to immediate feedback conditions. However, since this model is consistent with the idea that criterial learning ~~arises through the gradual formation of stimulus-response associations, rather than via the construction of a response criterion. This model also assumes that the rate at which these associations are learned is slowed by feedback delays.~~ relies on working memory – and the available evidence suggests that logical reasoning and working memory are unaffected by feedback delays of several seconds (e.g., in one-dimensional rule-based category learning tasks; Ell et al., 2009; Maddox et al., 2003; Maddox & I – the learning process described in equation 2 likewise may be unaffected by feedback delay. We therefore also investigated a version of the model in which the term α/t_{FD} is replaced by α, making the learning rate independent of feedback delay.

This model further assumes that both the stimulus and the criterion representations may drift randomly throughout the duration of time they are maintained in working memory. The representation of the criterion must always be maintained in working memory, whereas drift in the stimulus representation affects performance up until the feedback is presented, but not afterwards (i.e., because the equation 2 updating rule depends on the value of the current stimulus representation). We modeled the drift in both the criterion and the percept by adding white noise to their initial values. Specifically, we assumed that for all $t > \mathrm{T_S}$

$$c_n(t) = c_n(\mathrm{T_S}) + \eta_c \epsilon(t), \tag{3}$$

and

$$x_n(t) = x_n(\mathrm{T_S}) + \eta_x \epsilon(t), \tag{4}$$

where $\epsilon(t)$ is white noise and η_c and η_x are parameters that determine the amount of drift over time. Note that these parameters are permitted to be equal to zero, which is equivalent to assuming that the criterion and/or the percept do not drift over time at all.

~~Simulations of these models in a task that was structurally identical to Experiment 2 revealed that the time-dependent drift model typically predicts that the most important variable for criterial learning is the total amount of time~~ This model predicts that between the ~~response and the stimulus presentation that defines the next trial. Where the feedback is presented within this interval is relatively unimportant. In contrast, both the delay-sensitive learning model and~~ time of stimulus presentation and feedback, the response criterion will drift randomly. The amount of drift during this time will be normally distributed with mean zero and variance $\mathrm{t_{FD}}\eta_c^2$. Similarly, the amount of criterial drift after the updating that follows feedback and the presentation of the stimulus on the next trial will also be normally distributed with mean zero, but now with variance $(\mathrm{t_{ITI}})\eta_c^2$. Predictions for the percept are similar. Specifically, the ~~reinforcement-learning model consistently predict that feedback delays should impair performance more than a long ITI~~amount of drift between stimulus presentation and feedback will be normally distributed with mean zero and variance $\mathrm{t_{FD}}\eta_x^2$.

~~The experiments were designed to test these predictions. Our results strongly suggested that human criterial learning is sensitive to feedback delay but not to the ITI duration. This suggests that~~

Models that assume no criterion

A qualitatively different class of models, based on an actor-critic architecture (Sutton & Barto, 1998), assume that learning is a process of associating each stimulus with a response via reinforcement learning. These models do not include an explicit representation of the response criterion. Rather the criterion is simply the mental representation that separates percepts associated with contrasting responses.

In this class of models, the perceptual representation of each stimulus is modeled as a pattern of activation across 25 sensory units that are each characterized by an overlapping tuning curve. Each unit is maximally excited by one specific stimulus, which we call the unit's preferred stimulus. Specifically, the activation of the i^{th} sensory unit on trial n at time t is given by

$$A_i(n,t) = \exp\left(\frac{-\left[d_{i,S_n(t)}\right]^2}{\sigma^2}\right) \tag{5}$$

where $d_{i,S_n(t)}$ is the (Euclidean) distance between the preferred stimulus value of the i^{th} sensory unit and the value of the stimulus that was presented on trial n at time t, and σ is a constant that increases with perceptual noise.

Because this class of models does not store an explicit response criterion, it cannot incorporate criterial drift. It can, however, incorporate perceptual drift. In this case, the value of the ~~delay-sensitive learning model and the reinforcement-learning model provide a better account of human behavior across wide ranges of their respective parameter spaces than the time-dependent drift model~~current stimulus used to compute sensory unit activation (Equation 5) is assumed to drift over time according to

$$S_n(t) = S_n(\mathrm{T_S}) + \eta_S \epsilon(t), \tag{6}$$

where η_S scales the magnitude of the drift and $\epsilon(t)$ is a noise process.

The ~~time-dependent drift model was based on the idea that criterial learning might be entirely supported by working memory. The assumption of drift in this model stems from the understanding that maintaining items in working memory is inherently challenging. The longer an item is held, the more likely it is to deteriorate. Therefore, the model predicts that performance should deteriorate as the time between the response on~~ model includes two decision or actor units – one associated with each of the two possible responses. Initially, each sensory unit is connected to both actor units with some random connection strength. Let $\omega_{iJ}(n)$ denote the strength of the connection between sensory unit i and actor unit J (for J = A or B) on trial n~~and stimulus presentation on trial $n+1$ increases, regardless of how this interval is divided between feedback delay~~. Then the activation in actor unit J on trial n at time t, denoted by $V_J(n,t)$, equals

$$V_J(n,t) = \sum_i \omega_{iJ}(n) A_i(n,t). \tag{7}$$

A response is selected by comparing the activations in competing actor units immediately after stimulus presentation, that is, at time $t = \mathrm{T_S}$, using the following the decision rule:

$$\text{Respond } R_n = \begin{cases} A, & \text{if } V_A(n,\mathrm{T_S}) > V_B(n,\mathrm{T_S}) \\ B, & \text{if } V_A(n,\mathrm{T_S}) \leq V_B(n,\mathrm{T_S}). \end{cases} \tag{8}$$

The connection strengths $\omega_{iA}(n)$ and ~~ITI. The results of Experiments 1 and 2 strongly disconfirm this prediction.~~

~~In contrast, the delay-sensitive learning model was inspired by the idea that a response criterion might be encoded in a more stable memory system. One appealing candidate for this function is the cerebellum,~~ $\omega_{iB}(n)$ are updated after feedback is received on each trial according to standard reinforcement learning rules (Sutton & Barto, 1998). If the feedback is positive, then for J = A or B:

$$\omega_{i\mathrm{J}}(n) = \omega_{i\mathrm{J}}(n-1) + \frac{\alpha_{actor}}{t_{\mathrm{FD}}} A_i(n,\mathrm{T_F}) V_{\mathrm{J}}(n,\mathrm{T_F}) \delta(n-1)[1-\omega_{i\mathrm{J}}(n-1)], \tag{9}$$

where ~~synaptic plasticity has been shown to follow a gradient descent learning rule and also to be sensitive to feedback timing (Brudner et al., 2016; Held et al., 1966; Honda et al., 2012; Kitazawa et al., 1995; Kitazawa & Yin, 2002). Finally,~~ α_{actor} is the learning-rate of the ~~reinforcement-learning model draws on the idea that criterial learning could emerge through~~ actor, and $\delta(n-1)$ is the reward prediction error on trial $n-1$. The last term prevents the weight from exceeding 1. Note that the increase in synaptic strength is scaled by the product of the pre- and postsynaptic activation values. Also note that, as in equation 2, the learning rate is scaled by the inverse of the feedback delay. If the feedback is negative, then

$$\omega_{i\mathrm{J}}(n) = \omega_{i\mathrm{J}}(n-1) + \frac{\alpha_{actor}}{t_{\mathrm{FD}}} A_i(n,\mathrm{T_F}) V_{\mathrm{J}}(n,\mathrm{T_F}) \delta(n-1) \omega_{i\mathrm{J}}(n-1). \tag{10}$$

Now the last term prevents the weight from falling below 0. Note that on error trials, $\delta(n-1)$ will be negative, so equation 10 decreases the synaptic weight.

The reward prediction error on trial $n-1$ is defined as the value of the obtained reward, denoted by $R(n-1)$ minus the value of the predicted reward, denoted by $P(n-1)$:

$$\delta(n-1) = R(n-1) - P(n-1). \tag{11}$$

The predicted reward on trial n is determined by the critic via

$$P(n) = P(n-1) + \frac{\alpha_{critic}}{t_{\mathrm{FD}}} \delta(n-1), \tag{12}$$

where α_{critic} is the learning rate of the critic. Note that this learning rate is again scaled by the inverse of the feedback delay.

Much evidence suggests that this type of stimulus-response ~~association learning driven by dopamine-dependent synaptic plasticity in the striatum. This process aligns with established theories of category learning that posit a key role to this learning process in~~

~~the basal ganglia Ashby et al., 1998.~~ learning is mediated largely within the striatum, and is facilitated by a dopamine-mediated reinforcement learning signal that is time dependent (e.g., Valentin et al., 2014). Specifically, the dopamine signal generated by positive feedback appears to peak at around 500 ms after feedback and then decay back to baseline levels within 2 or 3 s (Yagishita et al., 2014). As a result, synaptic plasticity at cortical-striatal synapses is attenuated with increasing feedback delays (Yagishita et al., 2014). The scaling of α_{actor} and α_{critic} by $1/t_{\mathrm{FD}}$ is one way to model this phenomenon. However, it is possible that stimulus-response association learning is mediated by other brain systems that are less sensitive to feedback delay. For this reason, we also investigated a version of the model in which the updating equations are independent of feedback delay. In this version, the term $\alpha_{actor}/t_{\mathrm{FD}}$ is replaced by α_{actor} in equations 9 and 10 and the term $\alpha_{critic}/t_{\mathrm{FD}}$ is replaced by α_{critic} in equation 12.

~~Our results do not strongly favor the delay-sensitive learning model over the reinforcement learning model or vice versa. The critical disagreement between these accounts is whether the criterion has any real psychological meaning. In the delay-sensitive learning model it does~~

Simulation Results

We investigated how changing the duration of the feedback delay and the ITI affect criterial learning for each class of model. More specifically, we simulated performance of each model in each of the three conditions of the Experiment 1 categorization task. Each simulation continued for 200 trials or until the model responded correctly for 12 trials in a row. For each set of parameter values, we simulated the model 100 times and averaged the results, yielding one observed performance metric measured in mean trials-to-criterion for each condition.

Our approach was to generate predictions for each of the three experimental conditions across a wide range of parameter values. We then used parameter-space partitioning (PSP) to evaluate the performance of each model (Pitt et al., 2006). Traditional computational modeling identifies parameter values that allow the model to provide the best possible fit to the observed data. This traditional approach therefore treats the single observed data set being fit as a gold standard. Our goal instead was to investigate the range of possible predictions of each model. We were interested in questions such as: can a model predict any possible outcome or is it constrained to always predict, for example, that feedback delays impair learning? PSP was designed to answer such questions.

Classical PSP is defined for non-stochastic models and is often implemented by exploring the parameter space via a Markov chain Monte Carlo algorithm. In contrast, we used brute force search over a grid of plausible parameter values. Because our models are stochastic, at each sampled set of parameter values we generated multiple simulations and classified the results into one of a small set of qualitative outcome patterns. PSP calculates the proportion of the parameter space where the model makes specific predetermined qualitative predictions. We focused on the following four such predictions. 1) Delay impaired: Performance is worst in the Feedback-Delay condition. 2) Long ITI impaired: Performance is worst in the Long-ITI condition. 3) Short ITI impaired: Performance is worst in the Control condition. 4) Other: Two or more conditions are tied for worst performance. For each model, the PSP analysis quantified the proportions of the explored parameter space where the model made each of these qualitative predictions.

Models that assume a stored criterion

We investigated predictions of these models across a wide range of values for the parameters η_x, ~~and thus this model is psychologically similar to the time-dependent drift model – the main disagreement being about whether the representation~~ η_c, and α. Specifically, in the case of α, we stepped through every value in the interval $[0, .2]$, with a step size of .01. In the case of both η_x, and η_c, we searched over the interval $[0, 5]$, with a step size of .1. We also explored the binary case of whether or not the update rate α was scaled by the feedback delay.

Figure 8 shows the results. Panel A lists the proportions of parameter space over which the version of the model that is sensitive to feedback delay predicts each of the four qualitatively different outcomes. Note that this model predicts that the poorest performance will be in the Feedback Delay condition for all values of its parameters. At first glance, this result seems surprising because it means that there are no parameter values that make the model more sensitive to ITI than to feedback delay. This makes sense though because increasing the ITI has no effect on the mean criterion value – the only effect is to increase the criterion value's variance. In contrast, increasing the feedback delay affects the mean criterion value because it slows the criterion updating process from converging on the optimal value.

Panel C of Figure 8 shows the same results as in panel A, except for the version of the model in which the criterion updating process is insensitive to feedback delay. Note that this version of the model predicts that performance must be worst in the Long-ITI condition – again for all values of its parameters. This result also shows that no amount of perceptual or criterial drift allows this version of the model to account for the Experiment 1 results.

Models that assume no criterion

We investigated predictions of the SR-learning model by varying four parameters; namely, the perceptual-noise variance σ^2, the stimulus drift parameter η_S, and the actor and critic learning rates (i.e., α_{actor} and α_{critic}, respectively). In the case of σ, we searched over the interval $[1, 10]$, with a step size of 3. For η_S, we searched over the interval $[0, 1]$, with a step size of .1. For both α_{actor} and α_{critic}, we stepped through every value in the interval $[0, .2]$, with a step size of .02. We constrained our search over these parameters to this interval because reinforcement-learning models are prone to instability at very high learning rates (Sutton & Barto, 1998). As evidence of this, at higher learning rates, the model failed to learn with any consistency – that is, in most cases, it failed to reach the learning criterion (12 correct responses in a row) within the allowable 200 trials. We also explored the binary case of whether or not the update rates α_{actor} and α_{critic} were scaled by the feedback delay. As with the other models, all scores were normalized by the largest observed value.

The results are described in Figure 9. Panel A lists the proportions of parameter space over which the feedback-delay sensitive version of the model predicts each qualitatively different outcome. Note that across the vast majority of its parameter space, this model

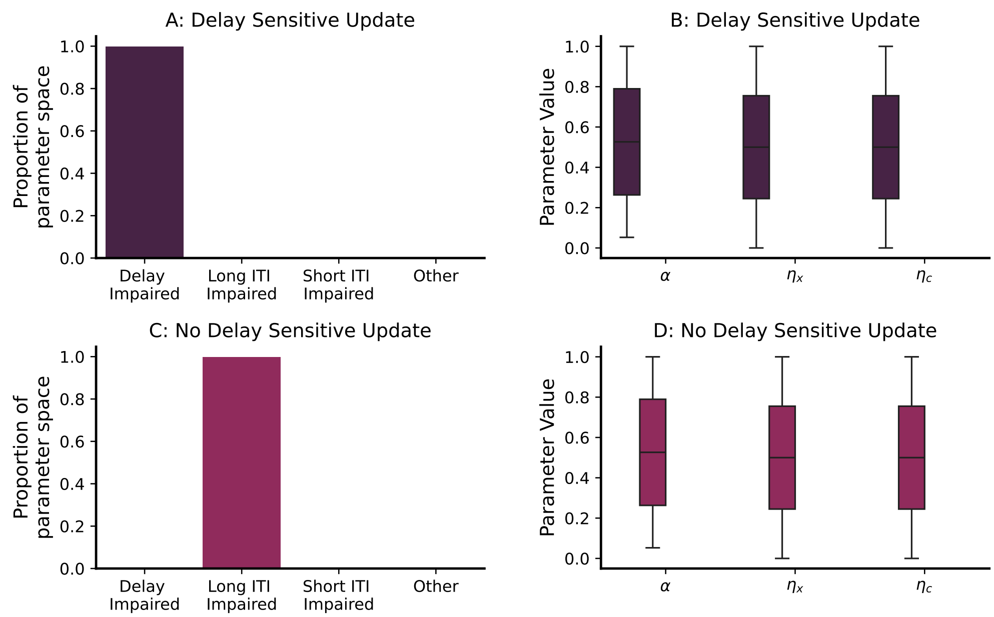

Figure 8

Simulation results for models that assume a stored criterion. ***A:*** *The proportions of parameter space where each of the four qualitative outcomes is predicted by the versions of the model that assume criterion updating is sensitive to feedback delay.* ***B:*** *Boxplot showing the parameter ranges that produced the panel A results. All parameter values were normalized by the largest value in the search range, so the ordinate ranges from zero to one for all parameters.* ***C:*** *The proportions of parameter space where the models that assume criterion updating is insensitive to feedback delay predict each of the four qualitative outcomes.* ***D:*** *Boxplot showing the parameter ranges that produced the panel C results. Note: In all panels, color indicates the PSP pattern.*

predicts that an increase in feedback delay will impair performance more than an increase in ITI.

Figure 9C shows the proportions of parameter space over which the models that assume synaptic updating is insensitive to feedback delay predict each of the four qualitative outcomes. Surprisingly, this version of the model also almost always predicts the poorest performance in the Feedback-Delay condition. What is going on here? First, note that even if $\alpha_{actor}/t_{\mathrm{FD}}$ and $\alpha_{critic}/t_{\mathrm{FD}}$ in equations 9, 10, and 12 are replaced by α_{actor} and α_{critic}, respectively, then equations 9 and 10 are still sensitive to feedback delay because the pre- and postsynaptic activations in equations 9 and 10 are computed at the time of feedback – that is, at time $\mathrm{T_F}$. So the longer that the feedback is delayed, the more time there is for the stimulus representation to drift from its true value. As a result, this model predicts that delayed feedback will impair learning, even though the amount of updating on each trial does not depend on feedback delay. Second, note that all versions of this model are

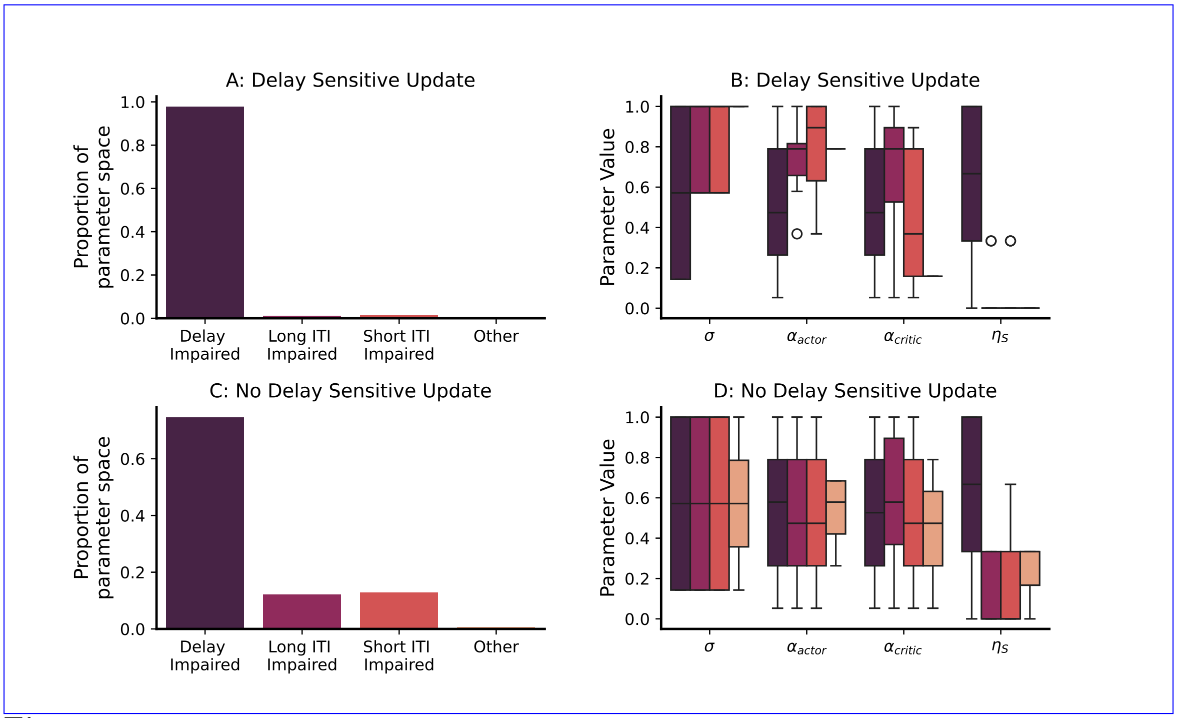

Figure 9

Simulation results for models that assumed stimulus-response learning and therefore no stored criterion. ***A:*** *The proportions of parameter space where each of the four qualitative outcomes is predicted by the versions of the model that assume synaptic updating is sensitive to feedback delay.* ***B:*** *Boxplot showing the parameter ranges that produced the panel A results. All parameter values were normalized by the largest value in the search range, so the ordinate ranges from zero to one for all parameters.* ***C:*** *The proportions of parameter space where the models that assume synaptic updating is insensitive to feedback delay predict each of the four qualitative outcomes.* ***D:*** *Boxplot showing the parameter ranges that produced the panel C results. Note: In all panels, color indicates the PSP pattern.*

generally insensitive to the length of the ITI. The stored-criterion models are adversely affected by a long ITI because the criterion drifts during this time and large drift means that on the next trial a decision will be made with a suboptimal value of the criterion~~is stable and if the updating of the criterion is~~ . But there is no criterion to drift in the SR-learning models. The stimulus will continue to drift during the ITI in these models, but with no consequence, because after the synaptic updating that occurs with presentation of the feedback, the stimulus representation on the current trial is no longer needed. So all versions of the SR-learning model are more sensitive to feedback delay ~~.~~ than to the duration of the ITI. In support of this interpretation, note from panel D of Figure 9 that the few versions of the model in which synaptic updating is insensitive to feedback delay that predict poorest performance in the Long ITI or Control conditions all include almost no stimulus drift (i.e., η_S is small), which eliminates the one remaining component of the model that makes it sensitive to feedback delay.

~~The reinforcement-learning model makes very different psychological assumptions. This model assumes that behaviors described as “criterial learning ” are actually mediated by the learning of stimulus-response associations~~

Summary of modeling results

We examined two classes of criterial-learning models that make different predictions about the effects of feedback delay and ITI. The first class assumes that the criterion is stored in memory and compared against incoming stimuli to generate responses. These models predict poorest performance in the Feedback-Delay condition only when the update rule is explicitly delay-sensitive. Otherwise, they make the strong prediction that performance must be worst in the Long-ITI condition. The second class of models assumes no stored criterion, with responding driven instead by reinforcement learning of SR associations. Almost all versions of this general model make the strong a priori prediction that performance must be worst in the Feedback-Delay condition – even those versions that predict no effect of feedback delay on the strength of synaptic updating.

General Discussion

The notion of a response criterion is ubiquitous in psychological models of decision making. Despite its theoretical importance, relatively little is known about its perceptual and cognitive basis. This article focused on two strikingly different interpretations of the criterion. The first, which seems to be assumed implicitly by most decision-making models – including for example, signal detection theory – is that decisions are made by comparing the stimulus value to a stored value of the criterion, and that this stored value is updated trial-by-trial after the feedback is presented. A second interpretation, which is motivated by results from the procedural-learning literature, is that learning instead is a process of associating responses with stimuli. According to this account, ~~the criterion has no internal representationand therefore no psychological meaning. Testing~~ there is no response criterion – at least not one with any mental representation. The value that the first class of models would refer to as the response criterion is simply the hypothetical value that separates stimuli associated with contrasting responses.

As a first step in testing between these two very different ~~accounts of criterial learning should be a focus of future research.~~ interpretations, we explored, both empirically and theoretically, how feedback delays and increases in the ITI affect criterial learning in a one-dimensional category-learning task. Our empirical results strongly suggested that human criterial learning is sensitive to feedback delay but not to the ITI duration. To investigate the theoretical implications of this result, we examined predictions of two qualitatively different types of computational model. This analysis showed that SR-learning models almost always predict that feedback delays should impair learning more than long ITIs, even versions of the model in which the amount of synaptic updating is insensitive to feedback delay. In contrast, a broad class of models that assumed the criterion is learned and updated trial-by-trial predict that increasing the ITI should have greater effect than delaying feedback, except for the subset of these models that explicitly assume the criterial updating process is sensitive to feedback delay.

~~We are not aware of any data that directly addresses the question of whether a response criterion is a fundamental psychological construct, as the delay-sensitive learning model suggests, or whether it is unnecessary, as assumed by the reinforcement learning model. Even so, there are some results in the~~ In summary, our results suggest that if a criterion is learned and stored in memory, then its trial-by-trial updating must be sensitive to feedback delay. How likely is this model? If a criterion is learned and stored in memory, then one obvious possibility is that the storage is in working memory. After all, in the present experiments at least, the current value of the criterion would need to be accessed and updated every few seconds – properties that are commonly attributed to working memory (e.g., Baddeley, 2010; Oberauer, 2002). The problem with this account is that it is widely accepted that working memory is largely insensitive to feedback delay. For example, rule-based category-learning ~~literature that support the interpretation of the reinforcement-learning model. In information-integration (II) category-learning~~ tasks ~~, the optimal strategy is similarity-based, and difficult or impossible to describe verbally (e.g., Ashby & Valentin, 2018). When the stimuli vary on two dimensions, the stimuli from contrasting categories can be partitioned by a decision bound that is conceptually similar to a response criterion. In both cases, all stimuli on one side are associated with one response, and all stimuli on the other side are associated with the contrasting response~~that are known to recruit working memory and executive attention are generally unaffected by feedback delays of up to 10 s (Dunn et al., 2012; Ell et al., 2009; Maddox et al., 2003; Maddox & Ing, 2005; Worthy et al., 2013) . Whereas working memory is generally attributed to circuits centered in prefrontal cortex (e.g., Ashby et al., 2005; Funahashi, 2017), an alternative possibility is that the criterion is stored in the cerebellum. For example, the cerebellum has been associated with working-memory tasks (e.g., Ashida et al., 2019), and synaptic plasticity there follows a gradient descent learning rule that is sensitive to feedback timing (Brudner et al., 2016; Held et al., 1966; Honda et al., 2012; Kitazawa et al., 1995; Kitazawa & Yin, 2002) . Although our results can not rule out this possibility, there is virtually no previous literature linking criterial learning to the cerebellum, so much more empirical and theoretical work would be needed to establish this as a viable model.

Whereas the stored-criterion model struggles to account for our results, the SR-learning account is simple and straightforward. In particular, almost all versions of this model predict our results – even those versions that assume the strength of synaptic updating is unaffected by feedback delays. Furthermore, ~~in agreement with Experiment 2, II category learning is impaired by short feedback delays (Maddox et al., 2003; Maddox & Ing, 2005). The analogous question in II category learning is whether the~~ there is a well-accepted neural account of this model. First, there is abundant evidence that the learning of arbitrary SR associations is mediated primarily within the striatum (e.g., Horvitz, 2009; Packard & McGaugh, 1992). Second, there is also abundant evidence that striatal-mediated learning is impaired with feedback delays as short as 2.5 or 3 s (Maddox et al., 2003; Maddox & Ing, 2005; Yagishita et al., 2014).

We are not aware of any other attempts to address directly the question of whether or not a response criterion is a fundamental psychological construct. Even so, there are some results in the category-learning literature that support a mixed view. The multivariate generalization of a response criterion is a decision bound that separates multidimensional stimuli belonging to contrasting categories (e.g., Ashby & Gott, 1988). The multivariate analogue of the stored-criterion hypothesis is that the decision bound is learned ~~directly or whether it is simply the set of points~~ and stored trial-by-trial in some form of memory and that categorization decisions are made by noting which side of the decision bound the current stimulus is on. A second possibility though is that category-learning in these tasks is of SR associations, in which case the decision bound has no psychological meaning and instead is just the set of percepts that divide the perceptual space into ~~contrasting response regions . In fact, the evidence strongly supports this latter interpretation~~ regions associated with contrasting responses. Although more research is needed on this question, the current evidence strongly suggests that both hypotheses are correct, but that they apply to different types of categorization tasks.

With rule-based tasks in which the optimal strategy is a one-dimensional rule (like the tasks used here), the evidence strongly favors the decision-bound-is-learned hypothesis, whereas with information-integration tasks in which the optimal strategy is impossible to describe verbally and therefore requires some sort of similarity-based response strategy, the evidence strongly favors the SR-learning hypothesis (Ashby & Waldron, 1999; Casale et al., 2012). For example, if ~~the~~ a decision bound is learned, ~~then~~ it should be possible to ~~apply~~ use this bound to ~~novel stimuli.~~ respond to novel stimuli that fall in some nearby, but untrained region of perceptual space. In contrast, SR learning builds associations between specific stimuli and specific responses and would not generalize to novel stimuli – at least, not if the novel stimuli are perceptually distinct. With rule-based categories, ~~this is easy for participants, but with II categories~~generalization to novel stimuli is nearly perfect – that is, responses to novel stimuli are almost perfectly predicted by the one-dimensional decision bound that best describes the participant's training performance (Casale et al., 2012). However, with information-integration categories, there is no evidence ~~that the response strategy that participants learn can be generalized to novel stimuli (Casale et al., 2012) – a result that strongly supports the hypothesis that the decision bound has no psychological meaning~~of any generalization. Instead, performance on novel stimuli is at chance (Casale et al., 2012).

~~Our results clearly demonstrate that criterial learning~~ Rule-based category learning depends on working memory and executive attention and is relatively unaffected by feedback delays as long as 10 s, whereas information-integration category learning recruits procedural learning and is ~~impaired by delayed feedback~~ impaired by feedback delays as short as 2.5 s (for a review of the scores of studies supporting this hypothesis, see, e.g., Ashby, in press) . Our results suggest that criterial learning shares the same sensitivity to feedback delay as information-integration category learning. Therefore, our results support the hypothesis that criterial learning, ~~and not by extending the intertrial interval. These results are consistent with the hypothesis that criterial learning~~ like information-integration category learning, is a form of procedural learning – that is, criterial learning is a form of ~~basal-ganglia mediated associative learning , and are inconsistent with hypotheses that criterial learning is a working-memory-based process. Thus, our results provide a critical constraint on future models of rule-based classification and decision making, and possibly also on more general accounts of criterion setting, such as in signal detection theory.~~

~~Previous studies have failed to find that feedback delays impair rule-based category learning , and on the face of it, this seems to contradict our finding that feedback delays~~

~~impair criterial learning. However, all earlier RB studies that looked for effects of a feedback delay, either used binary-valued stimulus dimension and so no criterial learningwas needed, or else set the response criterion exactly midway between the category prototypes, which makes criterial learning trivial. For example, under these conditions, criterial learning might not even require feedback. The unsupervised category-learning experiments reported by Ashby et al. (1999) provide strong support for this because all of their rule-based participants learned the correct criterion (which was midway between the category means), even though the task was completely unsupervised. Moreover, all previous studies failed to isolate criterial learning , so even if there was an effect of feedback delay on criterial learning, it could have been masked by larger effects caused by other rule-learning processes~~SR associative learning that is mediated primarily in the basal ganglia. Similarly, our results are inconsistent with the hypothesis that criterial learning depends in any significant way on working memory and/or executive attention.

Criterial learning is among the most classic and ubiquitous of all cognitive skills. For example, signal-detection theory teaches that it is the central form of learning in an enormous range of decision-making tasks – everything from simple YES-NO detection of a weak signal to assessing the guilt or innocence of a defendant in a jury trial. Our results suggest that even in simple rule-based tasks, criterial learning seems to be subserved, at least in part, by associative mechanisms. Most current theories tend to classify tasks as either executive function (e.g., rule-based category learning) or procedural (e.g., mirror tracing). Our results suggest that such classification schemes might oversimplify how humans perform these tasks, and therefore that much more work is needed to understand how different learning and memory systems interact.

Transparency and Openness

All data have been made publicly available in the following GitHub repository: https://github.com/crossley/crit_learn_delay

Author Notes

Preparation of this article was supported by Public Health Service Grant MH2R01-063760.

References

Ashby, F. G. (2017). Computational cognitive neuroscience. In W. Batchelder, H. Colonius, E. Dzhafarov, & J. Myung (Eds.), *New handbook of mathematical psychology, Volume 2*. MIT Press.

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. (in press). *Statistical decision theory in perception and cognition: Signal detection & general recognition theories*. MIT Press.

Ashby, F. G., Ell, S. W., Valentin, V. V., & Casale, M. B. (2005). FROST: A distributed neurocomputational model of working memory maintenance. *Journal of Cognitive Neuroscience*, *17*(11), 1728–1743.

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., Queller, S., & Berretty, P. M. (1999). On the dominance of unidimensional rules in unsupervised categorization. *Attention, Perception, & Psychophysics*, *61*(6), 1178–1199.

Ashby, F. G., Smith, J. D., & Rosedahl, L. A. (2020). Dissociations between rule-based and information-integration categorization are not caused by differences in task difficulty. *Memory & Cognition*, *48*(4), 541–552.

Ashby, F. G., & Valentin, V. V. (2017). Multiple systems of perceptual category learning: Theory and cognitive tests. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science, Second Edition* (pp. 157–188). New York: Elsevier.

Ashby, F. G., & Valentin, V. V. (2018). The categorization experiment: Experimental design and data analysis. In E. J. Wagenmakers & J. T. Wixted (Eds.), *Stevens' handbook of experimental psychology and cognitive neuroscience, Fourth edition, Volume five: Methodology* (pp. 307–347). Wiley.

Ashby, F. G., & Waldron, E. M. (1999). On the nature of implicit categorization. *Psychonomic Bulletin & Review*, *6*(3), 363–378.

Ashida, R., Cerminara, N. L., Edwards, R. J., Apps, R., & Brooks, J. C. (2019). Sensorimotor, language, and working memory representation within the human cerebellum. *Human Brain Mapping*, *40*(16), 4732–4747.

Baddeley, A. (2010). Working memory. *Current Biology*, *20*(4), R136–R140.

Bohil, C. J., & Wismer, A. J. (2014). Implicit learning mediates base rate acquisition in perceptual categorization. *Psychonomic Bulletin & Review*, *22*, 1–8.

Brudner, S. N., Kethidi, N., Graeupner, D., Ivry, R. B., & Taylor, J. A. (2016). Delayed feedback during sensorimotor learning selectively disrupts adaptation but not strategy use. *Journal of Neurophysiology*, *115*(3), 1499–1511.

Casale, M. B., Roeder, J. L., & Ashby, F. G. (2012). Analogical transfer in perceptual categorization. *Memory & Cognition*, *40*, 434–449.

Crossley, M. J., & Ashby, F. G. (2015). Procedural learning during declarative control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(5), 1388–1403.

Dunn, J. C., Newell, B. R., & Kalish, M. L. (2012). The effect of feedback delay and feedback type on perceptual category learning: The limits of multiple systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(4), 840–859.

Ell, S. W., Ing, A. D., & Maddox, W. T. (2009). Critrial noise effects on rule-based category learning: The impact of delayed feedback. *Attention, Perception, & Psychophysics*, *71*(6), 1263–1275.

Funahashi, S. (2017). Working memory in the prefrontal cortex. *Brain Sciences*, *7*(5), 49.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley.

Held, R., Efstathiou, A., & Greene, M. (1966). Adaptation to displaced and delayed visual feedback from the hand. *Journal of Experimental Psychology*, *72*(6), 887–891.

Honda, T., Hirashima, M., & Nozaki, D. (2012). Adaptation to visual feedback delay influences visuomotor learning (B. Brembs, Ed.). *PLoS ONE*, *7*(5), e37900.

Horvitz, J. C. (2009). Stimulus–response and response–outcome learning mechanisms in the striatum. *Behavioural Brain Research*, *199*(1), 129–140.

Kitazawa, S., Kohno, T., & Uka, T. (1995). Effects of delayed visual information on the rate and amount of prism adaptation in the human. *The Journal of Neuroscience*, *15*(11), 7644–7652.

Kitazawa, S., & Yin, P.-B. (2002). Prism adaptation with delayed visual error signals in the monkey. *Experimental Brain Research*, *144*(2), 258–261.

Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2003). Delayed feedback effects on rule-based and information-integration category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 650–662.

Maddox, W. T., & Ing, D. A. (2005). Delayed feedback disrupts the procedural-learning system but not the hypothesis-testing system in perceptual category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 100–107.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 411–421.

Packard, M. G., & McGaugh, J. L. (1992). Double dissociation of fornix and caudate nucleus lesions on acquisition of two water maze tasks: Further evidence for multiple memory systems. *Behavioral Neuroscience*, *106*(3), 439–436.

Pitt, M. A., Kim, W., Navarro, D. J., & Myung, J. I. (2006). Global model analysis by parameter space partitioning. *Psychological Review*, *113*(1), 57–83.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT Press.

Treutwein, B., Rentschler, I., & Caelli, T. (1989). Perceptual spatial frequency-orientation surface: Psychophysics and line element theory. *Biological Cybernetics*, *60*(4), 285–295.

Valentin, V. V., Maddox, W. T., & Ashby, F. G. (2014). A computational model of the temporal dynamics of plasticity in procedural learning: Sensitivity to feedback timing. *Frontiers in Psychology*, *5*, 643.

Worthy, D. A., Markman, A. B., & Maddox, W. T. (2013). Feedback and stimulus-offset timing effects in perceptual category learning. *Brain and Cognition*, *81*(2), 283–293.

Yagishita, S., Hayashi-Takagi, A., Ellis-Davies, G. C., Urakubo, H., Ishii, S., & Kasai, H. (2014). A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science*, *345*(6204), 1616–1620.