Effects of Categorical and Numerical Feedback on Category Learning

Astin C. Cornwall¹, Tyler Davis², Kaileigh A. Byrne³, & Darrell A. Worthy¹

¹Texas A&M University, ²Texas Tech University, ³Clemson University

In press at Cognition
Abstract

Real-world learning signals often come in the form of a continuous range of rewards or punishments, such as receiving more or less money or other reward. However, in laboratory studies, feedback used to examine how humans learn new categories has almost invariably been categorical in nature (i.e. Correct/Incorrect, or A/Not-A). Whether numerical or categorical feedback leads to better learning is an open question. One possibility is that numerical feedback could give more fine-grained information about a category. Alternatively, categorical feedback is more dichotomous, potentially leading to larger error signals. Here we test how feedback impacts category learning by having participants learn to categorize novel line stimuli from either numerical, categorical, or a combination of both types of feedback. Performance was better for categorical relative to the more variable numerical feedback. However, participants were able to learn to effectively categorize from numerical feedback, and providing larger numerical rewards for easier, more representative stimuli was more successful in promoting learning than providing larger rewards for harder to classify stimuli. Simulations and fits of a connectionist model to participants’ performance data suggest that categorical feedback promotes better learning by eliciting larger prediction errors than numerical feedback.

Keywords: Category learning; Reward learning, Feedback; ALCOVE
Effects of Categorical and Numerical Feedback on Category Learning

1. Introduction

“Hot or Cold?” “Healthy or Unhealthy?” Categorization enables us to make judgments about the value of different actions as well as inferences about future events, such as whether or not a jacket is needed based on the weather (Markman & Ross, 2003). The ability to categorize objects and events allows us to generalize much of our knowledge about the world and reduce it to manageable proportions (Barsalou, 1983; Rosch, 1978).

Category learning paradigms make it possible to study how people acquire new categories in a laboratory setting. Typical category learning paradigms have participants learn to classify stimuli into one or more categories, while receiving feedback about whether their classifications are correct or incorrect (Ashby & Maddox, 2005; Markman & Ross, 2003; Medin & Schaffer, 1978). Thus, this type of learning depends on a process of observation, choice, and feedback in order to classify novel stimuli into a discrete number of categories (Ashby & Maddox, 2011; Nosofsky et al., 2019; Smith, 2014). However, category learning research has primarily employed feedback in the form of either discrete outcomes such as ‘Correct/Incorrect’, or similar discrete numerical values (i.e. 0 or 1 point).

In contrast to category learning paradigms, reinforcement learning paradigms often utilize either discrete or continuous numerical rewards. Reinforcement learning tasks are characterized by choosing between a discrete number of choices and iteratively learning which choices are more valuable based on feedback (Daw, O’Doherty, Dayan, Dolan, & Seymour, 2006; Erev & Barron, 2005; Frank & Claus, 2006; Kool, Gershman, & Cushman, 2017; Niv, 2009). While category and reinforcement learning paradigms differ in their framing of ways to influence learning (Radulescu, Niv, & Ballard, 2019), reward feedback is critical to changing future behaviors in both category learning (e.g. Abohamza, Weickert, Ali, & Moustafa, 2019; Daniel & Pollmann, 2010; Moustafa, Gluck, Herzallah, & Myers, 2015) and reinforcement learning (e.g. Montague, King-Casas, & Cohen, 2006; Sutton & Barto, 2018; Thorndike, 1927) paradigms. Collectively, both category and reinforcement learning are reward-dependent processes that shape how individuals learn new information and make future choices. Despite their similarities, the type of feedback that is typically used in these paradigms these paradigms has largely differed; category learning has relied on categorical feedback, but reinforcement learning has relied on both discrete and continuous numerical feedback. Consequently, the role of categorical feedback, as compared to numerical feedback, on category learning outcomes remains a relatively under-studied area of research.

Although category learning paradigms have favored categorical feedback, rather than feedback that encompasses a more continuous range of values, such dichotomous outcomes do not necessarily mirror the graded feedback one may receive in a real-life situation. As an illustration, when making decisions about what to wear based on the weather, we can make a prediction about the temperature, put on some clothes, and walk outside to where we will be able to gauge the efficacy of our prediction. If our choice of clothes was relatively congruent with the outside temperature, we would most likely opt to wear those clothes again in similar weather. Ergo, predictions that are congruent with the category representation will be reinforced (Ross, 2000). However, imagine if we predict that the weather is warm and wear light clothing. If the weather turns out to be a little cool, we might feel a bit of discomfort, but if it is frigid outside, we may find ourselves in a potentially dangerous situation. To prevent a similar occurrence in the future, we would have to understand that our prediction was incorrect by a certain degree and update our representations of weather types accordingly. Thus, category learning tasks may
benefit from the inclusion of feedback that falls on a continuous scale as it may be able to better confer a degree of correctness in the response to each decision.

It is, however, currently unknown whether categorical, continuous, or a combination of both types of feedback would promote better category learning. Learning from continuous feedback, as defined by a variable range of numerical values, such as the magnitude of discomfort felt due to the weather or a low amount of payment received for work, is often attributed to the amount of surprise one receives from the outcome based on prior expectation (e.g. Schultz, 2016a; 2017), or the prediction error. Decisions that result in positive prediction errors, where the outcome received was greater than expected, are more likely to be made again as they may become predictive of future rewards. Similarly, repeating the same decisions is less likely when the outcomes, or lack thereof, fall below expectation, leading to negative prediction errors and a reduced likelihood of choosing the same action again. Thus, continuous feedback may facilitate learning by giving participants a sense of how right their choice was. In other words, discrete feedback provides gross-level information, such as “Pass/Fail”, but does not specify how close or far a behavior is from the correct response. Continuous feedback, on the other hand, provides more fine-grained information. For example, a “Fail” score of 69% versus 19% on a test indicates vastly different degrees of revision that would be needed to achieve a passing score. However, because continuous feedback provides a broader range of information, it may take multiple observations before a reliable expectation of reward is learned because ‘correct’ choices are more variable and therefore more ambiguous or uncertain.

Conversely, categorical feedback immediately gives an expectation, in terms of correct or incorrect, but no information is given regarding how correct the response was. Additionally, recent work suggests that prediction error magnitude may have an impact on the rate at which categories are learned (Lohse et al., 2020). From this perspective, categorical feedback may facilitate better learning than continuous feedback as the initial prediction errors will be larger. For example, if categorical feedback is enumerated as ‘1’ for correct and ‘0’ for incorrect as in Ashby et al. (2011), continuous feedback is scaled from 0 to 1, and predicted probabilities of category membership also range from 0 to 1, then categorical feedback will lead to larger prediction errors than continuous feedback (detailed further below). Combining both types of feedback may provide both an immediate expectation and an indication of the magnitude of decision accuracy or inaccuracy. A combined approach therefore provides information about the outcome of one’s decision in both categorical and continuous terms. Thus, it is possible that receiving both forms of feedback may promote better performance on category learning tasks than categorical or continuous feedback alone.

Additionally, category learning in particular may benefit from the inclusion of continuous feedback if that feedback provides information about how representative each stimulus is of its category. Often, in category learning tasks, stimulus classification is defined by a perceptual boundary that distinguishes two categories, such as Category ‘A’ and Category ‘B’. Stimuli that are easier to classify are defined as being more representative of that category; these stimuli are usually farther away from the perceptual boundary. In contrast, stimuli that are more difficult to classify are less representative of that category and are typically closer to the perceptual boundary that divides category members. Continuous feedback may affect learning differently if larger rewards are given for correctly classifying more versus less representative stimuli.

Ultimately, the main difference between discrete and continuous feedback is the type of information that is given in response to a choice. Theoretically, in addition to categorical feedback, both continuous and combined feedback have the potential to promote the learning of
an underlying category structure. However, as we described above, the comparative efficacy of each feedback type is under-researched. To test how each form of feedback impacts learning, we use a computational modeling framework to simulate learning based on the three forms of feedback: categorical only, continuous only, and combined categorical and continuous feedback. In the simulations below, we show that categorical feedback does indeed lead to larger magnitude prediction errors than continuous feedback, and this leads to better predicted learning, when each model is simulated with parameters uniformly drawn from the parameter space (i.e. data uninformed, Wagenmakers et al., 2004). These a priori simulations also predict that when only continuous feedback is given, optimal learning depends on reward magnitude and category representativeness. Specifically, learning is optimized when high-magnitude rewards are given for correctly classifying stimuli that are highly representative of each category compared to less representative stimuli.

In the following sections we first present a basic overview of the task and the reward structure. We then present the formal models of learning from categorical feedback, versus continuous, reward-based feedback, followed by the simulations of each of the models in each experimental condition. We then report results from an experiment with human participants that was designed to evaluate their behavior as compared to our model-based predictions.

1.1. Brief Overview of the Task and Reward Structures

To examine how each type of feedback affects learning, we created a simple line categorization task, in which line stimuli vary in orientation and length, where the feedback given was based on one of two category structures and one of two reward structures for relevant feedback types. In the current study we utilized conjunctive rule (CJ) and information integration (II) category structures, used in prior research (Filoteo et al., 2010), and an easy (ES) and difficult (DI) reward structure where either the easiest or most difficult stimuli to categorize were rewarded more (detailed below). For categorical feedback (Cat), only the category structures were used as participants did not receive rewards. For both the continuous (Rwd) and combined (CatRwd) feedback types, both the category and reward structures were used. Continuous feedback was either 0 points for incorrect, or between 50-100 points for correct, and was scaled to between 0 and 1 in the model, while categorical feedback was ‘correct/incorrect’ and scaled as 1 and 0. The three feedback types and different combinations of the category and reward structures created 10 different conditions in total for this study: two with categorical feedback – CatII and CatCJ; four with continuous feedback – RwdIES, RwdIIDI, RwdCJES, and RwdCJDI; four with both types of feedback – CatRwdIES, CatRwdIIDI, CatRwdCJES, and CatRwdCJDI. Below, we present learning models for each type of feedback, and then simulate each of the ten conditions above.

1.2. Learning Models

To gain a better understanding about the potential processes that occur when categorizing stimuli under each type of feedback, we employed three computational models derived from the Attention Learning Covering Map (ALCOVE; Kruschke, 1992) category learning model to both simulate and later fit behavioral data. ALCOVE is a connectionist model which computes the probability that a stimulus belongs to a given category based on the integration of psychological dimensions and the attention allotted to each.

In ALCOVE, category membership is determined by the attention-weighted similarity between stored category representations (exemplars) and the to-be-classified stimulus. In the covering map version of ALCOVE, category representations are a set of nodes that are distributed evenly across the psychological space spanned by a category learning task, instead of
exemplars that have been encountered before. When a new stimulus is observed, the similarity between the stimulus and each node is computed, and these similarity values are aggregated across all nodes to determine the probability that the stimulus belongs in a given category (See Kruschke, 1992 for a full description of the model).

In the current paper, we employ three variants of the covering map version of ALCOVE with a few important deviations. Most notably, our models do not learn the attentional weights of each psychological dimension; rather, our models use a free parameter to estimate the attention that is given to each psychological dimension (length or orientation). In the current paradigm, discussed further below, an optimal observer should not weight the dimensions differently. As such, we fixed the attention weights as attention learning would not be relevant in the current study and models. Our models use ALCOVE’s equations to compute the similarity between a current stimulus and each of the models’ hidden nodes. Each hidden node learns a weight that describes the strength with which its area of psychological space is associated with one category or another. Here, for the covering map, we used a 21 X 21 grid of 441 hidden nodes evenly spread across the two-dimensional stimulus space.

1.2.1. Category-Learner Model.

The category-learner model is identical to ALCOVE (Kruschke, 1992) in the computation of the activation values for each $j^{th}$ node as shown in Equation 1:

$$A_j = \exp \left[ -s \left( \sum_i \alpha_i |h_{j,i} - x_{i,t}|^r \right)^{q/r} \right]$$

(1)

The Activation values ($A$) for each node ($j$) are computed on each trial ($t$) for each psychological dimension ($h$). In the present model, we used only the psychological dimensions of line length and orientation for $h_{1,2j}$. On each trial, the difference between the dimension value of each node ($h_{j,i}$) and the actual observed dimensions ($x_{i,t}$) are modified by an attentional weight ($\alpha_i$) free parameter and specificity constant ($s; s > 0$). Importantly, nodes that are more similar to the current stimulus that must be classified will have higher activation than nodes dissimilar to the current stimulus. Values of 0 or 1 for indicate exclusive attention given to one dimension. In the context of the current model, a value of 1 indicates attention towards stimulus length, and attention to stimulus orientation for values of 0. The similarity metric and gradient values, $r$ and $q$ respectively, are both set to 1 in the current model.

The category-learner model slightly departs from the original ALCOVE model in the calculation of the activation values of the output nodes ($A^{out}_k$):

$$A^{out}_k = \sum_j \frac{A_j}{\sum_j} \cdot w_{jk}$$

(2)

where for each response ($k$), a vector containing the weights between each $j$th node and each $k$ response node ($w_{j,k}$) is multiplied by the normalized activation values ($A_j$) for each node that result from dividing each activation value by the sum of activation across all nodes. The normalization of activation values is a slight modification from the original instantiation of ALCOVE (Kruschke, 1992), and is done here to be consistent with the reward-learner model we present below.

The expected response node values ($w_{j,k}$) for each node in the chosen response are modified on every trial according to Equation 3 below:

$$\Delta w_{j,k} = d (\Psi_t - w_{j,k}) A_j$$

(3)

where $\Psi_t = C_t = \{ \text{Correct}1 \text{, Incorrect}0 \}$

where $d$ is a learning rate parameter, and $\Psi_t$ is a teacher signal that will take on values based on the feedback given in each model. In the category-learner model, $\Psi_t$ will equal $C_t$, a binary
value, which represents categorical feedback. Correct categorizations result in a value of 1, and 0 for incorrect categorizations. Effectively, this means that whenever a category learning trial is correct, the $Ψ_t - w_{j,k}$ error computed in Equation 3 will always be positive and result in an improved association weights between nodes and the chosen category. Conversely, on an incorrect trial, the error value will always be negative and should result in a poorer association between category and hidden layer nodes. As we will see below, these prediction errors will be larger for categorical than for continuous reward-based feedback, which leads to different predictions about how well the underlying categories will be learned from each feedback type.

We used values of 0 and 1 for categorical feedback for two reasons: to adhere to methods used in prior modeling of category learning (e.g. Ashby et al., 2011), and to portray incorrect categorizations as neutral information in the models rather than punishment, as punishment has produced differences when learning from rule-based and information integration category structures in the literature (e.g. Freedberg et al., 2017; Ashby & O’Brien, 2007).

1.2.2. Reward-Learner Model.

The reward-learner model was designed to learn from continuously valued reward information only. This model is functionally identical to the category-learner model; however, categorical feedback, used in Equation 3, is no longer used. In place of the binary categorical feedback ($C_t$), we used scaled, continuous, reward information ($R_t; r \in (0,100)$) as the teacher signal $Ψ_t$. The calculation of the scaled reward value can be seen in Equation 4:

$$Ψ_t = R_t = \frac{r_t - r_{\text{min}}}{r_{\text{max}} - r_{\text{min}}}$$

(4)

where $r_t$ is the actual reward value that is observed on any particular trial. In Equation 4, to calculate the scaled reward value ($R_t$), the minimum and maximum reward values observed across all trials are recorded, denoted by $r_{\text{min}}$ and $r_{\text{max}}$ respectively, and used in normalizing the current observed reward as shown in Equation 4. This function ensures that the observed reward values are scaled to a value between 0 and 1, but also constrained to the range of known outcomes up to that time point in the task or simulation. Once the reward values of 0 and 100 are observed, the scaled reward value calculation reduces to $r_t/100$. For consistency with the category-learner model, and to avoid any potential confounds, we opted to not include negative rewards, or punishment, in this model as well.

While this model is able to use the reward information to determine how correct the response is on each trial, one drawback is that this model will likely lead to smaller positive prediction errors and more frequent negative prediction errors. The reasoning is that when a correct response is made, the reward value will likely be fairly close in magnitude to the expected outcome. This will either produce a small positive or negative prediction error. When an incorrect response is given, it will always produce a large negative prediction error due to incorrect outcomes always resulting in 0 points. As we noted for the category-learner model, this will most likely have an impact on the rate at which the best responses are learned for each group of stimuli.

1.2.3. CatRwd-Learner Model.

The final model variant we used was a hybrid catrwd-learner model which simply weights how much information from both types of feedback is used in updating. The model itself is structurally identical to the reward-learner model aside from a modification of the teacher signal computation in Equation 4. This modification is detailed in Equation 5 below:

$$Ψ_t = qR_t + (1 - q)C_t$$

(5)

where $q \in (0,1)$ is a free parameter which represents the weight given to the scaled reward value ($R_t$: as calculated in Eq. 4), with $(1-q)$ representing the weight to categorical feedback...
information \((C_i)\). Importantly, depending on the value of the \(q\) parameter, the model will make increasing similar predictions to either the category- or reward-learner model as \(q\) approaches 0 or 1, respectively.

Since this model incorporates both categorical and reward information, the resulting prediction errors on each trial should be equal to or greater than the reward-learner and less than or equal to the category-learner model in terms of absolute value. The ability to equally weight both types of feedback in this model should overcome the issue of producing relatively small positive prediction errors that is predicted for the reward-learner by way of having inherently larger \(R_t\) values. For example, if a person equally weighted both types of feedback, made a correct choice and received a reward value of 50, a possible \(R_t\) value would be .75 as compared to a value of .5 for the reward-learner. However, as a comparison, the category-learner would produce a value of 1 given the same information. By imagining different expected values, one can begin to see what prediction errors each model would hypothetically produce. Prediction errors will always be the largest in absolute magnitude for the category-learner model because the reward amount is the maximum possible and is further away from the predicted probability of category membership. Importantly, the catrwd-learner is flexible in that if one type of feedback is weighted more heavily than the other, the predictions of the catrwd-learner will begin to reflect the predictions of the model (Category-Learner or Reward-Learner) with which it shares the over-weighted feedback type with.

The predicted response \((K)\) for any given trial and model is denoted by a probability value computed using Equation 6:

\[
Pr(K) = \frac{\exp(\phi A_{K}^{out})}{\sum_{k} \exp(\phi A_{k}^{out})}
\]

where \((\phi; \phi \in (0,5))\) is an inverse temperature parameter and \(k\) represents individual response options. Low inverse temperature parameter values typically lead to random decision-making, whereas higher parameter values indicate more consistent responses for the most probable category predicted by the model.

1.3. Simulation Methods and Predictions

To explore the general predictions for each of the previously described models, we simulated 500 datasets for each combination of the three models and ten task conditions. Each of the datasets represents simulated behavior for an individual participant with a unique set of parameters. The parameter representing attention given to the psychological dimensions of length and orientation \((\alpha)\) was set to 0.5 across all simulations to reflect equal attention to both dimensions; the specificity parameter \((s)\) was set to 0.05 as this value provides reasonable similarity gradients with these category structures, and these values were also used in the reward structure (detailed below); the learning rate \((d)\) and feedback weight \((q)\) parameters were independently drawn from a uniform distribution, \(U(0,1)\), for each simulated dataset; the inverse temperature parameter \((\phi)\), used in simulations, was drawn from \(U(0,5)\). For consistency in comparison, the same 500 simulated participants/sets of parameters were used by each model for each of the ten experiment conditions. To accommodate the randomized stimulus presentation within the task, each simulated participant naively completed each of the four condition/model combinations 100 times and had the 100 sets of data aggregated by individual trials.

Choices on each simulated trial were determined by comparing a number drawn from a \(U(0,1)\) distribution to the probability value computed in Equation 6 for choice 1. If the random number was less than or equal to the probability of Choice 1, Choice 1 was chosen. Otherwise, Choice 2 was chosen. Each trial had an objective optimal response - either making a correct categorization or choosing the most rewarding option. The proportion of optimal responses were
computed for each model and experimental condition. These proportions were averaged across all respective simulations and can be seen graphically in Figure 1.

Figure 1: Plot of simulated data by condition and model. Higher values denote a greater proportion of optimal choices as predicted by the respective model. CJDI refers to a conjunctive rule with a difficult reward structure; CJES refers to a conjunctive rule with an easy reward structure; IIDI refers to an information integration rule with a difficult reward structure, and IIES refers to an information integration rule with an easy reward structure. Error bars represent the standard error of the mean.

Overall, based on a uniform sample, the reward-learner model made fewer optimal responses than the category-learner and catreward-learner models. The poorer performance for the reward-learner model was most evident in the DI reward type conditions where the largest rewards were given for correctly classifying the most difficult items. Simulations using the catreward-learner model showed a similar trend, albeit of lesser magnitude. Differences between category structures amongst all models are relatively small but more pronounced in the category-
learner model.

![Figure 2: Plot of the raw and absolute predictions errors. Values were computed by taking the sum of prediction errors on each trial for each simulated participant. CIDI refers to a conjunctive rule with a difficult reward structure; CJES refers to a conjunctive rule with an easy reward structure; IIDI refers to an information integration rule with a difficult reward structure, and IIES refers to an information integration rule with an easy reward structure. Error bars show the standard error of the mean.](image)

As mentioned previously, a mechanistic explanation for the differences found within each model may be attributable to differences in both the magnitude and sign of the summed prediction errors each task structure produces over all 400 trials. Using the same simulated datasets, we extracted and summed the prediction error values on each trial for each simulated participant, and for each model and experimental condition. The mean raw and absolute prediction error values can be seen in Figure 2 above. Importantly, these simulated prediction errors are the summed errors over the course of the entire task. Trial-level prediction errors, however, as demonstrated by prior research in the literature (Schultz, 2016b; Schultz et al., 2017), should decrease in magnitude and variance as the task is learned.

The prediction errors for the category-learner model are the greatest in magnitude, followed by the catrwd-learner and the reward-learner models. Consistent with our reasoning, the magnitude of prediction errors seems to predict performance in each experimental condition if we compare the prediction error magnitude shown in Figure 2 to the proportion of best choices made in Figure 1. Thus, the poor predicted performance for the DI conditions with reward feedback is due to lower magnitude prediction errors, and, in particular, lower positive prediction errors. This leads to less updating of connection weights from hidden nodes to response nodes in Equation 3, and therefore poorer or slower learning of the underlying categories or best responses.

As shown by the simulation output, models with the largest summed prediction errors were associated with best predicted performance. The simulations also suggest that we should expect to find a difference between feedback types in both overall performance and summed prediction errors. Accordingly, categorical feedback is predicted to elicit the best performance in all conditions, and continuous feedback is expected to produce the worst learning performance. However, despite these predictions, it is unknown which of these models most corresponds with
human behavior. To determine which of our three models best reflects human behavior, we conducted an experiment in which participants completed a line categorization task analogous to the simulated task.

2. Method
2.1. Participants

Students from Texas A&M University participated in the study in partial fulfillment of an Introductory Psychology course requirement. We created a ten-group experimental design, as mentioned above, which consisted of 8 groups who were given reward information, and 2 groups that were only given categorical feedback. These 10 groups are the same as those presented above in the simulations.

Our goal was to collect data from 80 participants for each condition. This provides over 80% power to detect even small effects ($\eta^2_p = .02$). We also wanted relatively large sample sizes to avoid having to draw conclusions from insufficient data which could lead to Type 1 errors, and to offer precision of estimation for Bayesian and model-based analyses. In total, we recruited 760 participants and randomly assigned each to one of the following conditions. For reward feedback, the sample sizes per condition were as follows: 78 in the CJDI condition, 77 in the CJD condition, 79 in the IIDI condition, 78 in the IIES condition. For catrwd feedback, the sample sizes were: 84 in the CJD-I condition, 79 in the WES condition, 80 in the IIDI condition, 78 in the IIES condition. Finally, the sample sizes for the categorical feedback conditions were: 62 in the CJ condition, 64 for the II condition. These two conditions were run as comparison conditions at a later date than the other eight conditions and had slightly smaller sample sizes due to being run near the end of the data collection period for an academic term.

2.2. Experimental Task

Participants were shown one of 400 unique line stimuli on each trial in a randomized order. The line varied in both its length and orientation, and participants could make one of two responses to indicate what category they thought the line was in. Depending on what between-subjects condition participants were in, they received one of three types of feedback after making each choice: continuous reward-based feedback (reward condition) that varied between 0 and 100 points, categorical feedback (category condition), or both types of feedback (catrwd condition).

For the category and catrwd feedback conditions, participants categorized the lines into two categories of A or B. They were told whether they were correct for each choice, and participants in the catrwd condition also received either 0 points for an incorrect classification, or between 50-100 points for a correct classification. For the reward feedback conditions, participants were not told anything about the stimuli belonging to categories. Instead, they were told that on each trial they would pick from either option 1 or 2, and that the line on the screen would aid them in predicting which option would result in a reward, similar to the procedure detailed in Kahnt, Park, Burke, & Tobler (2012). To mimic real-life scenarios where an underlying category structure must be learned from non-categorical feedback, participants would then receive either 0 points or between 50-100 points; participants were not told whether they were ‘correct’ or not, but they could draw this inference from whether they received points or not on each trial. Thus, the major distinction between the Reward and the other two conditions is that for the category and catrwd conditions participants were explicitly told to categorize the line stimuli, whereas for the reward feedback condition participants are told to use the line as a reference for which option will be more rewarding. Ultimately, the tasks only differed in framing (Radulescu, et al., 2019) with the tasks with categorical feedback following typical category
learning procedures, and the reward-only task following the procedure of a basic reinforcement learning or decision-making task.

2.2.1. Category Structures.

The category structures briefly described prior were composed of lines which differed in the two perceptual dimensions of length and orientation which have been used in prior categorization work (Filoteo et al., 2010). The Conjunctive Rule (CJ) structure is, putatively, a readily verbalizable rule-based structure where the optimal bounds are orthogonal to the axes of the stimulus dimensions (Ashby et al., 1998; Ashby & Maddox, 2005). By contrast, the optimal bound for the Information Integration (II) structure is oblique to the perceptual dimensions, making it difficult to verbalize an optimal rule. However, our goal in the present work was not to focus on differences between rule-based and information-integration category learning which have been thoroughly studied (Ashby et al., 2019; Carpenter et al., 2016; Donkin et al., 2015; Ell et al., 2006; Maddox et al., 2004; Zaki & Kleinschmidt, 2014), but instead to simply use two commonly utilized category structures to examine how categorical information interacts with reward information and determine the value that both categorical and continuous reward feedback have when categorizing stimuli. A depiction of each category structure can be seen in Figure 3a below.

![Figure 3](https://via.placeholder.com/150)

Figure 3: Plots of stimuli used in each condition by the psychological dimensions of length and orientation. A.) Plot of the two category structures used in the current study: Conjunctive rule and Information Integration. Each individual colored dot on the plot represents a single stimulus reflective of both psychological dimensions. B.) Plot of reward values given for each stimulus in the DI reward conditions; separated by both category (CJ or II). C.) Plots of reward values given for each stimulus in the ES reward conditions. While the stimuli dots on each plot remain consistent within category structures, the potential reward value for a particular dot differs dependent on reward structure.

2.2.2. Reward Structures.

The reward structures used by each of the reward and catrwd feedback conditions were derived from the probability of each stimulus being in the correct category based on the relative similarity of each stimulus to all the exemplars from each category. Specifically, we used the Generalized Context Model (GCM; Nosofsky, 1986) to calculate the classification probability for each stimulus by comparing its similarity to exemplars from its category against its similarity
to exemplars from the other category. Classification probabilities were calculated by assuming equal weighting to each stimulus dimension and using a sensitivity parameter ($\phi$) value of .05, a value that was determined in a semi-arbitrary manner in order to generate reasonable probability gradients for the category structures used in our task.

In the reward and catrwd feedback type conditions, participants were given between 50 and 100 points for correct responses and 0 points for incorrect responses. In the ES condition, the points were a direct function of the classification probability of the stimulus for the correct category from the GCM model. The probability that the current stimulus is in the correct category, according to the model, was simply multiplied by 100 points. In the DI condition points given for correct classification of response/category $k$ were determined by the following transformation: $100 \times (1 - \text{Probability}(k)) + 50$. Thus, in the ES condition, more points were given for correctly classifying the easier, high probability stimuli, whereas in the DI condition more points were given for correctly classifying the more difficult, low probability stimuli. The category structures with rewards associated for correct classifications of each stimulus are shown in Figure 3b-c above.

2.2.3. Trial Procedure.

Each participant completed an experimental task on a computer in a laboratory environment after signing an Institutional Review Board-approved consent form. The instructions and stimuli were presented onscreen using Matlab and PsychToolbox version 2.54. Participants were told that they would be shown images on a screen, which consisted of white lines that varied in length and orientation. Depending on the assigned feedback type, there were slight differences in the task as detailed below and in Figure 4.

2.2.3.1. Categorical Feedback. Participants were asked to categorize each line on the screen into either Category 1 or 2 and were asked to respond with either the ‘z’ and ‘/?’ keyboard keys. Upon selection, participants were explicitly told if they were correct or not with text strings of ‘CORRECT’ and ‘INCORRECT’ colored green and red respectively before the next trial began after a 2s delay for feedback time. Sample trial screens showing both the stimulus presentation and feedback can be seen in Figure 4.

2.2.3.2. Reward Feedback. Participants were told that the line on the screen would aid them in predicting which of the two options would be the most rewarding on that trial using the same letter keys used when given categorical feedback. Upon choosing an option, the screen would display 0 or 50-100 points and continue to the next trial. As previously mentioned, no negative reward values, or punishments, were included. Importantly, no mention of categories or classification was present in this condition. Participants given catrwd feedback were shown experiment screens identical to the categorical feedback, but reward information was also given. Like the categorical feedback participants, they were asked to categorize each line on the screen into either Category 1 or 2. Upon selection, participants would be shown both the categorical and reward feedback simultaneously.

2.2.3.3. Trial Blocks. The experiment consisted of four 100-trial blocks. A break screen separated each 100-trial block in an effort to reduce fatigue. These screens would display progress information and the number of trials that have been completed. Reward feedback participants were only told that they had completed X number of trials out of 400, and to keep
trying to earn as many points as possible. In the category and catrwd feedback groups, participants were told what percentage of the previous 100 trials were correctly categorized. The primary dependent variable for both experiments was the proportion of optimal responses made across trial blocks and across all trials.

2.3 Data Analysis. To analyze categorization accuracy we fit mixed effects linear regression models using R’s brms package (Bürkner, 2017). We report the estimated linear regression coefficient, denoted as b, for each fixed effects predictor, along with its 95% credible interval (CI). This means that, given our data, the true value of the parameter is encompassed within an interval of the posterior distribution with a 0.95 probability (Nalborczyk et al., 2019). In our current analyses, a parameter value of 0 would indicate that a factor had no meaningful effect in the model. As such, we interpret any credible interval that contains 0 as evidence that the true value of the parameter has at least some probable value of not impacting the model, or as ‘non-significant,’ and any 95% credible interval that does not include 0 as evidence of a non-zero effect, or as ‘significant.’

For some analyses we report follow up contrasts using Bayesian t-tests. These were computed using JASP software (jasp-stats.org), using the default priors. The Bayes Factors ($BF_{10}$) for the t-test represents the odds that the alternative hypothesis is more supported than the null, given the data. Bayes factors between 3-10 denote moderate support for the alternative
hypothesis, Bayes factors between 10-30, 30-100, or greater than 100 represent strong, very strong, and extreme support for the alternative (Wagenmakers et al., 2018).

3. Results
3.1. Behavioral Results
3.1.1. Learning Over Time.

We first analyzed the data for the catrwd and reward feedback conditions, and did not include data from the categorical feedback conditions. To compare the rates at which participants learned the task across conditions, we fit a mixed effects logistic regression model using R’s brms package (Bürkner, 2017). The first model predicted the optimal responses on each trial from the interaction between the two feedback types (catrwd and reward), the two category structures (II and CJ), the two reward structures (ES and DI), and block number (four 100-trial blocks centered). Block number was entered as a random slope for each participant, in addition to a random intercept.

Figure 5: Plot of best option chosen proportions across (A) all 400 trials and (B) 100-trial blocks for each feedback type and category/reward structure condition. CJDI refers to a conjunctive rule with a difficult reward structure; CIES refers to a conjunctive rule with an easy reward structure; IID1 refers to an information integration rule with a difficult reward structure, and IIES refers to an information integration rule with an easy reward structure. Error bars shown depict the standard error of the mean.

Figure 5 plots the average proportion of optimal choices for participants in each condition. Taking into account the overlap between categories, the expected classification performance for an optimal observer, with perfect knowledge of the optimal category bound, was 0.7997 for II, and 0.7895 for CJ. Similar to our simulations, participants in the reference group (catrwd feedback) conditions made more optimal responses than participants in the reward feedback alone conditions (b = -0.09, CI = [-0.13, -0.07]). Also, in line with the trend detailed by the simulations, we observed a main effect of category structure (b = 0.04, CI = [0.01, 0.07]) with poorer performance in conditions using the conjunctive rule category structure than the information integration structure. While we did not specifically analyze the simulated performance by trial block, we found a main effect of block (0.04, [0.03, 0.05]) suggesting that learning occurred across trial blocks.
Additionally, we observed a Feedback Type X Category Structure X Block interaction \((b = 0.02, \text{CI} = [0.01, 0.04])\), a Feedback Type X Block interaction \((b = -0.03, \text{CI} = [-0.04, -0.02])\), and an interaction between Feedback Type and Reward Structure \((b = 0.05, \text{CI} = [0.01, 0.09])\). Plots of the interactions can be seen in Figure 6 below. The Feedback Type X Block interaction is due to a difference in learning slope between the catrwd and reward feedback types in which learning improves less across blocks for participants with reward feedback versus catrwd feedback (Figure 6a).

![Figure 6: A.) Plot detailing the differing learning slopes between reward and catrwd feedback. B.) Plot of the Feedback X Reward Structure interaction showing the interaction is likely influenced by the reward feedback group. C.) Plot of the interaction between block and category structure. Error bars detail the 95% credible interval for each mean value.](image)

The Reward Structure X Feedback Type interaction is of particular interest. In Figure 5 above, there is very little difference in performance between ES and DI reward structures when given catrwd feedback, but an apparent advantage for ES conditions when given reward feedback. To further explore each of these interactions, we regressed the above full model, with the feedback type parameter excluded, on the catrwd and reward feedback data independently.

For the catrwd feedback data, we found only the main effects of category structure \((b = 0.04, \text{CI} = [0.01, 0.07])\) and block \((b = 0.04, \text{CI} = [0.03, 0.05])\). The difference in performance between category structures was expected based on the simulation results. For the reward feedback conditions, we found main effects of category \((b = 0.08, \text{CI} = [0.05, 0.11])\) and reward \((b = 0.04, \text{CI} = [0.01, 0.08])\) structures, and a Category Structure X Block interaction \((b = 0.02, \text{CI} = [0.01, 0.03])\). In Figure 5b above, we can see that halfway through the task, performance begins to decrease in the CJ conditions whereas performance continues to increase to the expected levels in the II conditions. Since this behavior is not seen in the CatRwdCJ conditions, the lack of categorical information is a possible cause for this difference. These results show that the Reward Structure X Feedback Type and three-way Feedback Type X Category Structure X Block interactions observed in the full model were due to the reward structure effect and Category X Block interaction which were present in only the reward feedback data. To determine if catrwd and reward feedback types differed from categorical feedback alone, we ran an additional model that collapsed the catrwd and reward feedback data across reward structures. This model predicted the average proportion of optimal responses from each type of feedback (Category, CatRwd, Reward), category structure, and block number. As with the full model above, block and participant number were used as a random slope and intercept respectively.
Setting categorical feedback as the ‘Feedback Type’ reference group, there were differences between reward feedback ($b = -0.05, CI = [-0.08, -0.02]$), but not catrwd feedback ($b = 0.02, CI = [-0.01, 0.04]$), and also an effect of block ($b = 0.03, CI = [0.03, 0.04]$). These results suggest that categorical and catrwd feedback elicit similar performance and are somewhat surprising based on our simulation predictions. The simulations predicted that the catrwd feedback would show poorer performance than categorical feedback on the assumption that both forms of feedback were equally weighted. Additionally, the results provide evidence against our opposing predictions that the inclusion of reward would increase performance in the task. However, these results show that performance on both the catrwd and categorical feedback conditions were similar.

Finally, we ran two additional models to directly compare categorical feedback data to reward feedback data. To do this, we split the reward feedback data by reward structure (RwdES and RwdDI) and compared each set to the total category feedback data. This model simply predicted the optimal response from both types of feedback and gave a random intercept to each participant. Similar to previous comparisons of the two feedback types, the RwdDI participants showed poorer performance overall when compared to category feedback participants ($b = -0.07, CI = [-0.10, -0.04]$). Interestingly, however, the difference in performance between RwdES and categorical feedback participants was lower, but not significantly lower ($b = -0.02, CI = [-0.05, 0.00]$). This suggests that when the easiest stimuli are the most rewarding, learning from continuously valued-rewards is only slightly worse than learning from categorical feedback alone, and that the overall difference between feedback types stems from the poor performance in the DI conditions where the largest rewards were given to the stimuli that were the hardest to classify.

For exploratory purposes, since there was only a marginal difference in accuracy between the category feedback conditions and the RwdES conditions, we also compared the RwdES data to both the CatRwdES and CatRwdDI data using the same model above. Interestingly, the RwdES data differed from both CatRwdES ($b = -0.04, CI = [-0.06, -0.01]$) and CatRwdDI conditions ($b = -0.04, CI = [-0.06, -0.01]$). Since the differences between RwdES and the CatRwd groups were more pronounced than the difference between the RwdES and the categorical feedback group, this would imply that learning performance is improved to some extent when reward information is included alongside categorical information. However, we also used the above model to compare the CatRwdES and CatRwdDI data to the category feedback data. Similarly to an effect in a prior model, neither CatRwd dataset significantly differed from the category feedback data.

### 3.1.2. Response Time

We also conducted exploratory analyses on the average response times (RT) for participants in each condition. Similar to the analyses we conducted to examine learning over time, we used a Bayesian mixed-effects regression model to determine the extent of the differences in RT between feedback types and both category and rewards structures. With categorical feedback participants as the reference group ($\bar{RT} = 0.643s, SD = .39s$), there was no evidence of a difference in response time between categorical and catrwd feedback participants ($\bar{RT} = 0.756s, SD = .47s; b = 0.08, CI = [-0.03, 0.20]$). However, RTs did differ, on average, when comparing categorical feedback data to reward feedback data ($\bar{RT} = 0.951s, SD = .39s; b = 0.266, CI = [0.15, 0.38]$). This result suggests that giving only categorical feedback led to the quickest response times, while giving only reward feedback led to the longest response times. We also explored how RT values changed over time. Over the course
of the four 100-trial blocks, there is evidence that the overall RT values for each feedback type decreased over time as the task was learned (b = -.09, CI = [-0.15, -0.03]).

3.2. Theoretical Analysis

Next, we examined how well each of the three models described above (category-learner, reward-learner, and catrwd-learner) accounted for the data by fitting each model to each participant’s data individually. We then conducted post-hoc simulations to examine how well each model could reproduce the pattern of effects found in each condition. These post-hoc simulations involved simulating the experiment with each model with the best-fitting parameters from each participant. The goal of these post-hoc simulations was to compare how well each model’s simulated, or predicted, behavior aligns with the observed behavior from our participants.

3.2.1. Comparison of Model Fits.

Each model was fit to participants’ data on an individual basis by maximizing the log-likelihood of the model’s prediction for the optimal response on any given trial. We then calculated the Bayesian Information Criterion (BIC), a goodness-of-fit measure (Schwarz, 1978) to compare the fits of different models. Statistics such as BIC penalize models with more free parameters. Smaller BIC values indicate a better fit of the model to the data. Importantly, the category-learner and the reward-learner are considered to be simpler models nested in the full catrwd-learner model. Our two nested models are functionally identical to the full model, but do not include the ‘q’ parameter. This means that while the log-likelihood of the full model cannot be worse than the nested models, the BIC values for full model may be greater, due to the parameter penalty, and thus may indicate a poorer fit.

The BIC values, along with the average best-fitting parameter estimates for each condition are shown in Table 1 below. In general, the catrwd-learner and category-learner models were fairly consistent in terms of BIC. Due to the nested nature of the models, the similar fits of the category-learner and catrwd-learner models suggest that adding reward information to the model did not provide much improvement in fit. This can also be seen by examining the best-fitting ‘q’ parameter values, which weigh categorical versus reward information. These values are less than .5, on average, for all groups, and close to 0 for every group but the RwdCJDI feedback condition. Recall that when q=0 the model relies exclusively on categorical feedback and the catrwd and category-learner models are identical.

Overall, the best fitting model for each condition was the category-learner. In every instance, aside from the RwdCJDI condition, the catrwd-learner showed the second-best fits. Interestingly, in this RwdCJDI condition where the reward-learner was the second-best fit, the catrwd-learner showed a “q” parameter value greater than 0.5, which suggest for this condition alone, reward information was given more weight than categorical information. For the categorical feedback participants, there was no reward information given, so the model outputs should be close to identical as the reward-learner uses reward values of 1 and 0 when fitting. Discrepancies in fits are attributed to the reward scaling function (Eq. 4). Until the full range of rewards are known, 1 and 0 in this case, the deviations between both the category- and reward-learner models may differ.
Table 1

<table>
<thead>
<tr>
<th>Feedback Type</th>
<th>Structure</th>
<th>Model</th>
<th>a</th>
<th>s</th>
<th>d</th>
<th>s2</th>
<th>q</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Category-Learner</td>
<td>0.3676</td>
<td>0.0879</td>
<td>0.3949</td>
<td>1.9780</td>
<td>NA</td>
<td>426.2529</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reward-Learner</td>
<td>0.4048</td>
<td>0.1220</td>
<td>0.3621</td>
<td>1.9040</td>
<td>NA</td>
<td>427.8166</td>
</tr>
<tr>
<td>CatRwd</td>
<td>Categorical</td>
<td>Category-Learner</td>
<td>0.4435</td>
<td>0.0962</td>
<td>0.3326</td>
<td>2.3212</td>
<td>NA</td>
<td>376.2340</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reward-Learner</td>
<td>0.4356</td>
<td>0.1116</td>
<td>0.3434</td>
<td>2.3618</td>
<td>NA</td>
<td>377.1694</td>
</tr>
<tr>
<td></td>
<td>CJD</td>
<td>Category-Learner</td>
<td>0.4479</td>
<td>0.0898</td>
<td>0.3817</td>
<td>2.1326</td>
<td>NA</td>
<td>402.3978</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reward-Learner</td>
<td>0.4318</td>
<td>0.1007</td>
<td>0.2904</td>
<td>2.6180</td>
<td>0.4346</td>
<td>403.3981</td>
</tr>
<tr>
<td></td>
<td>CJES</td>
<td>Category-Learner</td>
<td>0.4447</td>
<td>0.0964</td>
<td>0.3500</td>
<td>2.3634</td>
<td>NA</td>
<td>407.3898</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reward-Learner</td>
<td>0.4729</td>
<td>0.1035</td>
<td>0.3969</td>
<td>2.1078</td>
<td>0.1271</td>
<td>416.5260</td>
</tr>
<tr>
<td></td>
<td>IIDI</td>
<td>Category-Learner</td>
<td>0.4143</td>
<td>0.0416</td>
<td>0.2481</td>
<td>3.0687</td>
<td>NA</td>
<td>376.9999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reward-Learner</td>
<td>0.4137</td>
<td>0.0795</td>
<td>0.2628</td>
<td>3.0572</td>
<td>0.1182</td>
<td>386.0445</td>
</tr>
<tr>
<td></td>
<td>IIES</td>
<td>Category-Learner</td>
<td>0.4301</td>
<td>0.0781</td>
<td>0.2323</td>
<td>3.1498</td>
<td>NA</td>
<td>369.3368</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reward-Learner</td>
<td>0.4375</td>
<td>0.0651</td>
<td>0.2149</td>
<td>3.2076</td>
<td>0.0741</td>
<td>373.0080</td>
</tr>
<tr>
<td></td>
<td>CJD</td>
<td>Category-Learner</td>
<td>0.5159</td>
<td>0.1159</td>
<td>0.6553</td>
<td>1.4119</td>
<td>NA</td>
<td>500.5736</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reward-Learner</td>
<td>0.5526</td>
<td>0.1529</td>
<td>0.5138</td>
<td>1.9299</td>
<td>0.6234</td>
<td>504.4965</td>
</tr>
<tr>
<td></td>
<td>CJES</td>
<td>Category-Learner</td>
<td>0.4507</td>
<td>0.1212</td>
<td>0.5168</td>
<td>1.6430</td>
<td>NA</td>
<td>462.0999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reward-Learner</td>
<td>0.4657</td>
<td>0.1491</td>
<td>0.5375</td>
<td>1.6463</td>
<td>0.0849</td>
<td>467.0120</td>
</tr>
<tr>
<td></td>
<td>IIDI</td>
<td>Category-Learner</td>
<td>0.5614</td>
<td>0.0949</td>
<td>0.4880</td>
<td>2.0662</td>
<td>NA</td>
<td>450.2910</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reward-Learner</td>
<td>0.5560</td>
<td>0.1099</td>
<td>0.4699</td>
<td>2.2264</td>
<td>0.1550</td>
<td>457.0588</td>
</tr>
<tr>
<td></td>
<td>IIES</td>
<td>Category-Learner</td>
<td>0.5044</td>
<td>0.0908</td>
<td>0.5047</td>
<td>2.1726</td>
<td>NA</td>
<td>421.9575</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reward-Learner</td>
<td>0.5219</td>
<td>0.0950</td>
<td>0.5141</td>
<td>2.1439</td>
<td>0.1564</td>
<td>428.6039</td>
</tr>
</tbody>
</table>
Note: Average best-fitting parameter estimates and BIC values for each condition. Smaller BIC values indicate a better fit of the model to the data. The best fitting model within each feedback type and condition is shaded and the BIC value is bolded. CJDI refers to a conjunctive rule with a difficult reward structure; CJES refers to a conjunctive rule with an easy reward structure; IIDI refers to an information-integration rule with a difficult reward structure, and IIES refers to an information integration rule with an easy reward structure.

To determine if differences in the model-predicted feedback weighting had an impact on learning performance, we assessed the association between the best fitting ‘q’ parameter for each participant and the total proportion of correct choices made. As shown in Figure 7 below, model-predicted focus on reward information (as q approaches 1) led to generally poorer performance in each catrwd and reward feedback condition. Interestingly, however, the two CJDI structure conditions did show a larger average ‘q’ parameter value (CatRwd-CJDI = .435; Rwd-CJDI = .623) as compared to the remaining conditions (~0.10). While this suggests a greater weighting of reward information in these two conditions, the information had a differential impact on the proportion of best responses. In the CatRwd-CJDI condition, the greater weighting of reward information did not seem to impact the overall proportion of correct responses (BF_{10}<1). However, the same cannot be said for the Rwd-CJDI condition as the greater weighting of reward information was associated with poorer categorization performance (r = -0.310, [-.491, -.091], BF_{10} = 6.002). This suggests that when stimuli are harder to categorize, reward information may be relied on more if it is available, but it is possible that this reliance may lead to poorer overall learning.

**Figure 7:** Scatter plots detailing the trends between the best fitting 'q' parameter from the catrwd-learner model and the proportion of correct responses for the catrwd and reward feedback groups. CJDI refers to a conjunctive rule with a difficult reward structure; CJES refers to a conjunctive rule with an easy reward structure; IIDI refers to an information-integration rule with a difficult reward structure, and IIES refers to an information integration rule with an easy reward structure. For the values of ‘q’, more weight is placed on categorical feedback as ‘q’ approaches 0.

**3.2.2. Post-Hoc Simulations.**

Using the process detailed by Ahn, Busemeyer, Wagenmakers, & Stout (2008), we ran 100 simulations using the best-fitting parameters for each participant as the input for each of the respective three models. The data for each of the 100 simulations for each participant were aggregated by trial to produce a single averaged simulated dataset for each model/condition combination. Thus, for each condition and each model we generated the average predicted
proportion of best responses on each trial, across all 100 simulations, using each participant’s best-fitting parameters. The participant data was aggregated in the same manner which yielded the observed proportion of correct choices made on each trial, across all participants in each condition. We then used both sets of data to compute the mean square deviation values for each combination of datasets using the formula in Equation 7 below:

$$MSD = \frac{1}{n} \sum_{t=1}^{n}(D_{exp,t} - D_{sim,t})^2$$

where \( n \) is the number of trials each participant observed, and \( t \) represents the individual trial number. For each experiment, since there were only two alternatives, we calculated the mean square difference (MSD) of the average proportion of correct responses (\( \hat{D} \)) between the experimental data and simulations. As performed in Ahn et al. (2008), we used the percentage values, instead of proportion values, when computing the MSD. We report the MSD values in their root form, as it is a more understandable metric of model performance, for each Feedback type, Model, and Structure in Table 1 below. The mean deviation values (MD) indicate the percentage of time the simulated data deviated from the observed data on each trial.

Based on Table 2 above, each of the models showed less than a 15% deviation in total, and about 10% deviation on average (10.224), between the post hoc simulated data and the observed behavioral data. This means that on average, over the course of the 400-trial task, a simulated participant using the same best fitting parameters as a human, would show incongruent behavior on ~40 trials. Within each feedback and structure condition, the MD values for each model are relatively similar, showing a difference in MD values of 0.2444 on average, 0.0893 median, between models with the categorical feedback conditions showing the highest deviations on average. This would equate to approximately ±1 deviation between models over the course of 400 trials. Overall, each of the models were fairly consistent within each condition in reproducing the experimental data from participants’ best fitting parameters. However, the reward-learner did show the smallest deviations on average in reproducing the behavioral data. Below, in Figure 8, we show the post hoc simulation learning curves for each condition based on the best fitting values.

**Table 2**

<table>
<thead>
<tr>
<th>Feedback</th>
<th>Structure</th>
<th>Category-Learner</th>
<th>CatRwd-Learner</th>
<th>Reward-Learner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Categorical</td>
<td>CJ</td>
<td>14.1796</td>
<td>NA</td>
<td>15.3241</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>12.7686</td>
<td>NA</td>
<td>14.2685</td>
</tr>
<tr>
<td>CatRwd</td>
<td>CIDI</td>
<td>7.7376</td>
<td>7.7170</td>
<td>7.7393</td>
</tr>
<tr>
<td></td>
<td>CIES</td>
<td>9.1892</td>
<td>9.1183</td>
<td>9.1751</td>
</tr>
<tr>
<td></td>
<td>IID</td>
<td>8.2149</td>
<td>8.1521</td>
<td>7.9289</td>
</tr>
<tr>
<td></td>
<td>IIES</td>
<td>8.4790</td>
<td>8.5085</td>
<td>8.2112</td>
</tr>
<tr>
<td>Reward</td>
<td>CIDI</td>
<td>9.1578</td>
<td>9.1181</td>
<td>9.1389</td>
</tr>
<tr>
<td></td>
<td>CIES</td>
<td>12.3855</td>
<td>12.4063</td>
<td>12.4232</td>
</tr>
<tr>
<td></td>
<td>IID</td>
<td>9.2743</td>
<td>9.2787</td>
<td>9.1626</td>
</tr>
<tr>
<td></td>
<td>IIES</td>
<td>10.1036</td>
<td>10.0732</td>
<td>9.9518</td>
</tr>
</tbody>
</table>
Note: Table of mean deviation (MD) values. Lower values indicate that there were fewer deviations between both datasets on average. Shaded cells indicate which model produced the lowest deviation for a given feedback type and condition. CJDI refers to a conjunctive rule with a difficult reward structure; CJES refers to a conjunctive rule with an easy reward structure; IIDI refers to an information-integration rule with a difficult reward structure, and IIES refers to an information integration rule with an easy reward structure.

It should be noted that the post hoc simulations were conducted in a different manner than our a priori simulations presented above. The a priori simulations are considered to be data-uninformed (Wagenmakers et al., 2004), as they take values uniformly sampled across the entire parameter space. In the post hoc simulations below, these are considered to be data-informed, as the parameter values that were best fit by each participant are used for the simulations. This means that the parameters used for the post-hoc simulations are more constrained in range than parameters used in our a priori simulations, thus resulting in different simulated behavior. For the post-hoc simulations the model predicted behavior is more similar because these models are, for better or worse, fairly flexible, and can account for a wide range of patterns of behavior with different parameter values. However, the general trends in learning shown by each model below seem to reflect the behavioral data shown above in Figure 5B, particularly the better predicted performance within the reward feedback group for participants in the ES conditions, compared to the DI conditions.

Figure 8: Learning curves for the post hoc predictions of each model, feedback type, and condition. The best fitting line is plotted for each data set.

3.2.3. Prediction Error Analysis

The a priori simulations of our task showed distinct differences in the magnitude of the average, summed prediction errors between feedback types and conditions. We postulated that the nature of the feedback received on each trial would lead to magnitude differences in prediction errors. We also stated that the feedback types with larger prediction error magnitudes would also lead to better learning of the task overall. To verify our claims, we extracted the prediction error values from the post hoc simulations we conducted in Section 3.2.2. In Figure 9A below, we show the average, summed predictions errors for each feedback type and category/reward structure. Similar to the a priori simulations, categorical feedback produced the largest summed prediction errors based on our models, whereas reward feedback produced the smallest.
Reward prediction error research suggests that, people, on average, theoretically exhibit smaller prediction errors as learning progresses (e.g., Schultz et al., 2017; Tian et al., 2016; Watabe-Uchida et al., 2017). The data in Figure 9B below provide further evidence of these trends. Similar to the behavior observed in general learning where people become proficient with time, the expected value of each category was sufficiently predicted, and correct categorizations led to smaller and smaller prediction errors as the task progressed. However, based on Figure 9B below, the rate at which the magnitude of predictions errors declined differed based on feedback type. For both categorical and catrwd feedback, the decline is linear with slight differences between conditions. In contrast, for reward feedback, the prediction errors asymptote early in the task. This result for reward feedback suggests that expected values may be learned quickly but do not promote any significant changes throughout the remainder of the task.

Figure 9: A.) Plot of summed prediction errors (absolute and raw values) by category/reward structure and feedback type. Larger values denote larger predictions over the course of the task. B.) Average absolute values of prediction errors across trials by category/reward structure and feedback type.

Additionally, we conducted an exploratory analysis to examine the average prediction errors for each stimulus in each feedback/condition. In Figure 10 below, we recreated Figure 3 with each condition and colored each individual stimulus, created as a combination of line length and orientation, based on the average prediction error across all participants. Importantly, the order of the stimuli was randomized for each participant, so any influence of when an individual stimulus was observed should be averaged out. The largest disparity in prediction errors can be seen in the DI reward structures; the structures where the largest rewards are given for the stimuli
closest to the category bounds. For these conditions, the smallest prediction errors seem to be closest in distance to the prototype of each category, and the largest prediction errors are associated with the stimuli at the bounds of each category distribution.

While the smaller prediction errors for the most typical stimuli could be an explanation for the poorer performance in these reward conditions, it also prompts a question about the role of the larger prediction errors, and associated rewards, at the bound. When the rewards are larger for these boundary stimuli, it could prompt a false sense of learning. The participant may receive a large reward signifying that they are correct, but on subsequent trials, similar stimuli may not give any reward. Thus, the prediction error would likely remain near .5, similar to what we observe in Figure 10 below, and decision-making behavior may be similar to chance. This could have contributed to the poorer performance we observed in the reward-DI conditions.

![Figure 10: Plot of the absolute values of the average prediction errors for each stimulus shown in each category structure and feedback group. In the above plot, the data shows the prediction errors for matching feedback types and models.](image)

**3.2.4. Exploration of Feedback Entropy**

A concern could be raised about the possibility that the difference between categorical and reward feedback, from our design, leads to a difference in the amount of information received, or the uncertainty surrounding the outcomes of each choice. A potential alternative explanation to our reward-prediction error account of why categorical feedback leads to better performance, compared to continuous reward feedback, is that there is less uncertainty in processing categorical feedback compared to continuous feedback. Simplifying the definition of Shannon entropy, the more certain outcomes are, the less information is received from the outcome (Shannon, 1948). A major motivation and goal in learning is the reduction of uncertainty, and entropy provides for a formal means of quantifying the amount of uncertainty associated with a stimulus or outcome (Aron et al., 2004; Crupi et al., 2018; Hasson, 2017). Prior research has also observed that large contrasts in potential rewards were associated with greater dopaminergic activity in the brain, suggesting an association between uncertainty and reward prediction errors (Fiorillo et al., 2003).

In terms of our task, we have categorical information comprised of only two possible outcomes. In contrast the reward-based feedback has ~51 possible outcomes across all 400
stimuli, possibly resulting in less certainty about what potential outcomes a participant might expect on any given trial despite learning the general features of each category. However, this difference in outcome possibilities introduces a contrast with the categorical feedback being more certain over time, but giving larger prediction errors, as compared to numerical reward feedback having less certainty, but smaller prediction errors. To compute the weighted average of the information gained across all 400 trials, we used Shannon’s entropy formula (Equation 8):

$$H(X) = H(p_1, ..., p_n) = - \sum_{i=1}^{n} p_i \log_2(p_i)$$

where $H(X)$ is the information for or uncertainty surrounding an event, in our case the information for the entirety of the task, in bits. In the formula, $p_i$ refers to overall proportion of outcomes observed over the course of all 400 trials. Additionally, we corrected the proportion of outcomes observed by the mean probability that the outcome should have been observed. Simply, this changes Equation 8 above to account for the overall information obtained from the outcomes given the probability of outcomes actually being observed. Equation 9 below reflects these changes:

$$H(X) = H(p_1, ..., p_n | p(x_1), ..., p(x_n)) = - \sum_{i=1}^{n} p_i \log_2(p_i)p(x_i)$$

where $p(x_i)$ is the probability that a unique outcome $x_i$ for each respective $p_i$, had for occurring for a given participant. As an example, in the DI version of this task, an outcome of 100 points would have a classification probability of 0.5 (based on the GCM-computed values discussed in Section 2.2.2.). Within each condition, we scaled the entropy values to the range of 0 to 1, thus controlling for any overall differences in the amount of information received between conditions.

Using the Bayesian mixed effects regression models, we first compared the amount of feedback entropy for participants who received categorical versus reward feedback; in this analysis we did not include participants who received catrdw feedback, as the modeling results suggest that most participants did not weigh continuous reward feedback very strongly, and it is unclear whether subjects attended to both types of feedback information. Accuracy and feedback accuracy were averaged for each of the four 100-trial blocks in the task. The first Bayesian multilevel model predicted feedback entropy, from the interaction between feedback type and block (centered), with a random intercept for block. The interaction between feedback type and block was not significant, ($b = -0.0000001, CI = [-0.0000020, 0.0000012]$), and there was also no effect of block, ($b = 0.0000003, CI = [-0.0000002, 0.0000002]$). However, there was a significant effect of feedback type, ($b=0.153, [0.148, 0.167]$). The coefficient value represents the average difference in entropy units, measured in bits, at the mean level of block. Collapsed across blocks the mean information value for the category conditions was 0.695 bits, compared to 0.849 bits in the reward condition. A Bayesian t-test indicates extreme support for an effect of group ($BF_{10}>100$). Entropy or decisional uncertainty was much greater in in the reward feedback condition, than in the categorical feedback condition.

Next, we ran a Bayesian multilevel model predicting accuracy for each block from the interaction between feedback group and block, with entropy added as a regressor to control for its relationship with accuracy. This model also included random effects for block. Even with entropy added as a nuisance variable in this model, there was still significantly lower accuracy for participants in the reward feedback condition, ($b = -0.050, CI = [-0.074, -0.025]$). As a comparison we ran the same model, but with feedback entropy not included. The effect of group was similar across the two models, ($b = -0.044, CI = [-0.068, -0.023]$). Thus, although feedback
entropy was considerably higher with continuous reward feedback, including entropy as a regressor in our model did not reduce the effects of feedback type on accuracy.

A potential reason as to why including entropy as a predictor did not reduce the effect of feedback type on accuracy, is that in the continuous reward conditions entropy was positively related to accuracy, \((b = 1.085, CI = [1.036, 1.137])\). This is due to the way the reward structure was constructed; less accurate participants received more “0” rewards, while more accurate participants received more variable rewards from 50-100, leading to higher feedback entropy. In the categorical feedback condition, in contrast, there was a negative relationship between feedback entropy and performance, \((b = -0.246, CI = [-0.288, -0.205])\). This is due to less accurate participants receiving more of an equal combination of 0s and 1s for incorrect and correct through the task, whereas more accurate participants received more 1s, leading to less feedback entropy. Overall, this analysis suggests that while there was greater feedback entropy from continuous than categorical reward feedback the task was designed such that there were also different relationships between accuracy and feedback entropy within each group. Future work could explore the alternative explanation for our findings, based on feedback entropy, with a design that better controlled for relationships between entropy and accuracy due to aspects of the task like the reward structure. For example, if we had given losses between -50 and -100 for incorrect categorizations, instead of 0, this probably would have reduced the positive relationship between entropy and accuracy in the continuous reward feedback condition.

4. Discussion

Real-world feedback often takes different forms, from feedback that is discrete and categorical to feedback that lies on a more continuous scale. In the current study, we simulated and experimentally tested how continuous and categorical feedback affect human category learning. Our results suggest that when learning to categorize novel line stimuli, categorical feedback produces better learning on average as compared to continuous reward feedback alone. The rate at which the continuous reward feedback participants learned to make optimal responses was slower, and overall performance was poorer as compared to the other two feedback conditions. This seems to imply that the definitive information gained from categorical feedback better facilitates learning as compared to the nuanced, and possibly more ambiguous information gained from continuous reward feedback. These findings corroborate our model-based hypotheses and provide evidence for how different forms of feedback could potentially impact real world learning.

When given categorical and continuous rewards simultaneously, our results suggest that people tend to disregard the continuous information in favor of the categorical feedback. Participants given both categorical and continuous reward feedback learned at the same rate as participants given categorical feedback alone. Our model-fitting results indicated that most participants who were given both types of reward feedback tended to weight categorical feedback much more than continuous feedback. However, participants in the CJD1 conditions gave slightly more weight to continuous reward feedback based on the value of the ‘q’ parameter for these conditions. This finding is likely due to these conditions giving the largest rewards for the stimuli nearest the boundary of category membership. Since these stimuli were likely more difficult to categorize, these participants may have relied more on the numerical reward, rather than categorical information.
4.1. Differences Between Feedback Types

Based on fits and simulations of a connectionist category-learning model (Kruschke, 1992), we have provided an algorithmic model-based account of the observed differences between participants who received categorical versus continuous numerical reward feedback. Under this account reward prediction errors were largest when categorical feedback was given, and the larger error signals prompted better learning. While speculative, there is extensive evidence from the neuroscience literature that prediction errors are correlated with dopaminergic spikes and activity in the ventral striatum (Glimcher, 2011). Cellular level models of category-learning assume that connections between cells in the ventral striatum and areas of the visual and premotor cortex are strengthened by dopamine reward signals (Wickens & Kotter, 1995; Worthy, Markman, & Maddox, 2013). We believe it is consistent with these cellular-level models to propose that larger prediction errors lead to better long-term potentiation of connections between striatal and cortical regions, which would enhance learning. Our prediction-error related account is also consistent with classic work following Rescorla and Wagner’s seminal paper (1972; Sutton & Barto, 1981; Pearce & Hall, 1980). An implicit assumption in much of this work is that learning and attention are reduced over time because learning improves and the prediction errors between the model’s predicted values and the observed outcomes become smaller. This assumption is similar to our account whereby manipulating the feedback to be either categorical or continuous affects prediction error magnitude, which subsequently affects learning. Categorical feedback provides the largest learning signals, which promote the best learning.

We also presented an analysis of feedback differences in feedback entropy in the categorical versus the continuous reward conditions. We observed that continuous feedback led to significantly more feedback entropy than categorical feedback, however, adding entropy as a regressor in our model predicting accuracy from feedback type, did not reduce the effects of feedback type on accuracy. However, future work could explore this intriguing, computational level hypothesis that increased feedback entropy leads to poorer learning, in a design that is better prepared to examine such a question. There are also several other potential verbal explanations for our findings. One such explanation, which is related to the issue of differences in feedback entropy, is that categorical feedback may have seemed more certain to participants than reward feedback. Categorical feedback is discrete, and outcomes can be understood with a degree of expectation. Comparatively, continuous numerical feedback often has more variability surrounding the initial outcomes. Receiving a value of 85 and nothing else when attempting to categorize a line stimulus, like the ones used on our task, may not give much information initially. A person may have questions about the range of rewards: whether the value was high or low; or simply be unsure how to apply the knowledge they just learned. This might be due, in part, to the contrast in the variability, or uncertainty surrounding the expectation, of each type of outcome and feedback (e.g., Walker et al., 2019).

Further, in the modeling framework we presented, categorical feedback represented the maximum possible reward, whereas continuous feedback varied. This difference may have implicitly modeled a reduced level of uncertainty in the categorical condition because reward prediction errors were consistently larger than in the continuous condition. In doing so, it may have led to larger updates of the connection weights between the hidden nodes of the model and response output nodes, and better learning of the categories. Thus, it is possible that the less variable or more discrete the feedback is (i.e., coin images, static values, category labels), the better the predicted learning is, compared to feedback consisting of a more variable range of values (i.e., continuous range of rewards, distribution of values, etc.). This improved
performance for discrete feedback over continuous feedback may be an explanation of why past research has shown that there is very little difference in categorization performance when comparing cognitive and monetary feedback (e.g., Daniel & Pollmann, 2010) or when comparing groups given differing magnitudes of discrete rewards (e.g. Bellebaum et al., 2010; Miller & Estes, 1961; Peterson & Seger, 2013).

It is also possible that the process of asking participants in the categorical feedback condition to categorize objects, versus asking participants in the continuous feedback condition to predict which option would give greater rewards, partially led to the discrepancy in performance between both groups. In a blocking paradigm by Bott and Hoffman (2007), participants who were asked to predict outcomes demonstrated significantly worse performance than participants asked to learn categories; this is reflective of our own findings. Further, an additional verbal account of our observed differences could be due to the option labeling used for each option. The simple action of labeling something as a member of a category has been shown to lead to quicker classification and learning as compared to the use of generic ‘option’ labels (Lupyan, 2012; Lupyan et al., 2007). Reasoning about categories may be different than reasoning about which option to choose. For example, reasoning with categories may involve more generalization than making simple multi-alternative choices. Our modeling framework offers one plausible mechanistic account (prediction errors) for the pattern of data we observed, but it is also possible that higher level processes, such as reasoning with or without categories, played a role in the observed differences. Future work should address this issue.

4.2. Impact of Reward on Task Stimulus Difficulty

Prior research has shown that people show increasingly poor performance as observed stimuli become harder to categorize or discriminate between (Daniel et al., 2011; Krebs et al., 2012; Schevernels et al., 2014). While our study did not directly compare performance between easier and harder stimuli within participants, we did compare the effect of rewarding either type of stimuli more than the other between participants. In the ES conditions of our task, the most rewarding stimuli were the most typical of their respective category, whereas the DI conditions gave the most rewards to the stimuli at the bounds of each category. When given only reward feedback, participants in the ES conditions showed far better performance than their counterparts in the DI conditions. This finding suggests that when learning from reward values alone, learning is best facilitated when the most typical stimuli give the greatest amount of reinforcement. An explanation for this difference, based on our model fitting, is that the increased difficulty promotes greater reliance on reward information. Interestingly however, the feedback weighting parameter ‘q’ in the catrwrd-learner model also shows that poorer performance is associated with weighting reward information more heavily in the reward feedback conditions.

4.3. Limitations and Future Directions

Since a wealth of categorization research examines how well people generalize learned knowledge to new stimuli (e.g. Ashby & Maddox, 2005a; Nosofsky et al., 2019; Yamauchi & Markman, 1998), one possible question is how well continuous feedback would prepare people for new stimuli. In a paradigm used in previous research, participants were asked to train on a set of novel stimuli, and then tested on new stimuli that populate the bounds of category membership (Seger, Braunlich, Wehe, & Liu, 2015). If we were to adopt a similar paradigm, there would likely be a difference between participants in the continuous feedback condition who received the largest rewards for either the most typical or atypical category members. However,
it is unclear whether categorical or continuous feedback in training will lead to better generalization. Numerical values representing how correct one is could lead to a more defined idea of category memberships at the bound of categories. However, the overall better learning demonstrated when given categorical feedback may translate to better generalization.

Additionally, in this study, we only used one model of category learning, the coverage map version of ALCOVE. There are multiple models of category learning that have been developed (see Wills & Pothos, 2012), and it is possible that one could be more suitable in using reward information. For example, we could compare our results against another exemplar-based model such as the Generalized Context Model (Nosofsky, 1986). Similarly, we could attempt to use models with different assumptions such as prototype models (Minda & Smith, 2012) or the Competition between Verbal and Implicit Systems model (COVIS; Ashby et al., 1998). Similar to how we modified ALCOVE in this paper, future research could modify the above models to utilize reward information and possibly create novel predictions about the role that continuous, numerical information has in category learning. Further, for a complete picture, it may also be useful to utilize basic reinforcement learning models that are agnostic to categorical information such as the Delta model (Don et al., 2019; Rescorla & Wagner, 1972) or the Decay model (Erev & Barron, 2005; Yechiam & Busemeyer, 2005). These models could be used to predict choice from outcomes alone, and possibly serve as a baseline comparison for models that include categorical information.

Further, this work only captures learning over the course of 400 trials which could be considered early-stage learning as compared to other category learning paradigms which have participants learn to a specific criterion or over the course of multiple sessions (e.g. Ashby et al., 2003; Homa & et al, 1973; Richler & Palmeri, 2014). Research has shown that strictly positive feedback, like we utilize here, wanes in effectiveness as categories are learned. While we have shown that reward information is not as effective as categorical feedback in the timeframe of 400 trials, it is possible that reward information may become more effective in finely differentiating between categories close in distance once the general features of a category are learned. There is precedence for this idea in the machine learning literature. Problems such as the Cart-Pole Balancing problem utilize large ‘reward’ values during the initial time steps to learn the gross movements, and then utilize more fine-tuned rewards to identify and maintain the optimal policy (Nagendra et al., 2017). More related to the present work, research by Zamir and colleagues (2017) showed that neural networks were more effective at categorizing images when iterative, categorical feedback was presented in a coarse to fine structure as the images were learned (ex. Object -> Vehicle -> Truck).

Alternatively, it may not be the reward values that are the cause for the differences between feedback types, but rather the structure of the categories. In our task, the categories are enmeshed and overlapping, so the stimuli at the bounds could equally be in one category or the other. Thus, it is possible that while the prediction errors would remain large for these stimuli, it would likely hinder learning, especially in the DI condition where the largest rewards are given for the stimuli at the bounds. In these conditions, the boundary stimuli would be highly rewarded, but largely uncertain, which would explain the poorer performance. It is also possible that if more deterministic categories were used, we would not see as poor of performance, particularly in the DI conditions.

In recent research by Liu and colleagues (2020), stratified and deterministic categories with discrete rewards and punishments resulted in quicker category learning as compared to categorical feedback alone. This finding is largely in contrast to our present findings but
highlights the idea that learning proficiency may be dependent on feedback and the structure of the categories and rewards. As such, future research could expand on our current paradigm to investigate if our contrasting results are due to the lack of punishment, range of rewards, non-deterministic categories, or something else entirely. A simple test would be to discretize our current reward structure into three levels of reward. This change may also lessen the information gap between the categorical and reward feedback groups that we observed in Section 3.2.4. as the number of potential outcomes between feedback groups would be closer in range. Additionally, it would be interesting if this discretization of rewards would also produce the same trend where more information leads to better performance, or if the relationship between information and proportion of optimal responses would be more similar to that of categorical feedback.

Finally, our paradigm assumed that different stimulus values, here line length and orientation, are commensurable, and can be evaluated as real numbers on a common scale, and that category membership can be expressed probabilistically. While the validity of these assumptions is subject to empirical scrutiny, a large body of work suggests that it is valid to assume that people can integrate values across stimulus dimensions, and that stimulus similarity can be compared on different scales (Shepard, 1987; Nosofsky, 1988). Future work should continue to explore these issues at both the algorithmic and computational levels.

4.4. Conclusion

We have shown that giving continuous versus discrete feedback leads to important learning differences in a categorization task. We detailed a possible mechanistic account of these learning differences using connectionist learning models. Both the models and the observed data suggest that, when learning to categorize novel stimuli, giving feedback that includes categorical information will lead to significantly better performance compared to feedback consisting of only reward information. Importantly, when given both types of information simultaneously, categorical information is likely to be more heavily weighted than reward information. We detailed that these learning differences likely stem from differences in the magnitude of predictions errors associated with each form of feedback, and that the perceivably larger amount of uncertainty surrounding the reward information had an effect on how well the categories were learned. Additionally, when given reward information alone, both the modeling and behavioral data showed that the relative difficulty of categorizing the stimuli affected learning. When the most typical stimuli of a category are associated with the largest rewards, we should expect performance similar to that of categorical feedback alone. Likewise, when the least typical stimuli, or those near the bounds of category membership, are associated with the largest rewards, poorer performance is to be expected. Thus, the present behavioral results and theoretical account suggest that feedback can be structured in different ways to promote better learning.
Acknowledgments

This work was supported by NIA grant AG043425 to D.A.W.

Supplementary Materials

Data and analysis code for this article can be found on the Open Science Framework via the following link: https://osf.io/uygx4/
References


Neuroscience, 30(1), 47–55. https://doi.org/10.1523/JNEUROSCI.2205-09.2010
https://doi.org/10.1109/ICACCI.2017.8125811


Schultz, W., Stauffer, W. R., & Lak, A. (2017). The phasic dopamine signal maturing: from reward via behavioural activation to formal economic utility. *Current Opinion in...*
Neurobiology, 43, 139–148. https://doi.org/10.1016/j.conb.2017.03.013
