Modeling category learning using a dual-system approach: A simulation of Shepard, Hovland and Jenkins (1961) by COVIS

C. E. R. Edmunds (ceredmunds@gmail.com)
School of Psychology, Plymouth University, UK

Andy J. Wills (andy.wills@plymouth.ac.uk)
School of Psychology, Plymouth University, UK

Abstract

This paper examines the ability of a dual-system, formal model of categorization COVIS (Ashby, Paul & Maddox, 2011) to predict the learning performance of participants on the six category structures described in Shepard, Hovland and Jenkin's (1961) seminal study. COVIS assumes that category learning is mediated by two dissociable neural systems that compete to control responding. The verbal system explicitly tests verbalizable rules, whereas the implicit system gradually associates each stimulus with the appropriate response. Although COVIS is highly influential, there are no published evaluations of the formal model against classic category learning data (COVIS is most typically applied heuristically to the design of new experiments). In the current paper, we begin to address this gap in the literature. Specifically, we demonstrate that COVIS is able to accommodate the ordinal pattern found by Shepard et al., provided that adjustments consistent with the model's theoretical framework are made.

Keywords: category learning; computational modelling; dual-system; implicit; explicit;

The field of category learning is currently inundated with formal models competing to explain categorization behavior (Wills & Pothos, 2012). These theoretical models vary on features such as how stimuli are initially represented, how category membership is determined and how they explain observed phenomena. One approach, the COVIS model of category learning (COmpetition between Verbal and Implicit Systems; Ashby, Alfonso-Reese, Turken, U, & Waldron, 1998; Ashby, Paul, & Maddox, 2011), hypothesizes that category learning is mediated by two competing systems: one explicit, one implicit. The explicit, verbal system learns using working memory to test hypotheses about category membership. In contrast, the procedural, implicit system learns by gradually associating areas of stimulus space with a response. At the beginning of learning, COVIS predicts that responding is controlled by the verbal system. People only switch to the implicit system if the verbal system cannot learn the category structure.

One of the strengths of COVIS is that it is capable of making predictions that affect three strands of the category learning literature: behavioral studies, cognitive neuroscience and cognitive modeling (Lewandowsky, Palmeri, & Waldmann, 2012). This is because the model contains, as well as the broad theoretical conceptualization above, both neurological and computational implementations. However, research with the COVIS model has not taken full advantage of these strengths. Instead, researchers have mainly focused on the behavioral predictions of COVIS. Crucially, these predictions are stated heuristically without any supporting formal modeling (Ashby & Maddox, 2005, 2010). The few papers that have presented formal modeling of COVIS have focused on its ability to account for specific novel behavioral results, such as impaired learning in Parkinson's patients, rather than a range of already-established seminal experiments in the category learning literature (Hélie, Paul, & Ashby, 2012a, 2012b). This neglect of standard results stands in contrast to the approach taken by proponents of other well-known models of category learning (e.g. Love, Medin, & Gureckis, 2004; Nosofsky, Gluck, Palmeri, McKinley, & Glauthier, 1994).

We note that at least some of COVIS's predictions do not seem to hold under under closer empirical scrutiny (e.g. Dunn, Newell, & Kalish, 2012; Edmunds, Milton, & Wills, 2015; Newell, Moore, Wills, & Milton, 2013; Stanton & Nosofsky, 2007, 2013). However, a different–arguably more fundamental–question is whether the COVIS model is capable of accommodating results already well-known in the category learning literature. As a first step to beginning to answer this question, we describe an investigation of COVIS's ability to account for the ordinal performance of participants learning the six category structures described in (Shepard, Hovland, & Jenkins, 1961). It is important to establish whether COVIS can accommodate these results for two reasons. First, not only have Shepard et al.'s results been found to be robust, they are also known to be problematic for some category learning models (Nosofsky et al., 1994). Second, Shepard et al.'s results have become a standard dataset against which most other formal models have been compared (e.g. Love et al., 2004; Nosofsky et al., 1994). Therefore, testing COVIS against this data set would enable detailed comparisons between COVIS and other leading models of category learning in respect of these particular data.

COVIS

There are slight variations in the model description between papers. The version described here was reported in Hélie et al. (2012a, 2012b).

The verbal system

The verbal system generates categorization responses to stimuli by selecting explicit, verbalizable rules from the set $\mathbf{R} = \{R_1, R_2, \ldots, R_m\}$, where each R_i represents a different rule. In previous implementations of COVIS, the set $\mathbf{R}$ has exclusively included all relevant one-dimensional rules, although it is also hypothesized to include some conjunction or disjunction rules. Each rule R_i is described mathematically in terms of a discriminant function $h_V(\underline{x})$, whose form depends on the type of rule being implemented. For a stimulus $\underline{x}$ that varies on r dimensions, denoted $\underline{x} = (x_1, x_2, \ldots, x_r)$, a one-dimensional rule based on dimension i calculates the value of the discriminant function as follows

$$h_V(\underline{x}) = x_i - C_i \tag{1}$$

where C_i is a constant that defines the decision boundary. The value of the discriminant function then determines the response given on that trial as follows

$$\begin{aligned} &\text{If on trial } n \text{ if } h_V(\underline{x}) < \varepsilon_V \text{ respond A} \\ &\quad \text{whereas if } h_V(\underline{x}) > \varepsilon_V \text{ respond B} \end{aligned} \tag{2}$$

where ε_V is a normally distributed random variable with a mean of 0 and standard deviation σ_V.

The rule used on a particular trial depends on the success of the rule used on the previous trial. If the categorization response was correct on the previous trial, the same rule will always be used on the next trial. However, if the response is incorrect the next rule is randomly selected from the set $\mathbf{R}$, weighted by each rule's current weight, $Y_i(n)$. The weight of each rule is dependent on the participant's past experience of the rule, the reward history of the rule and the participant's tendency to perseverate with incorrect rules. $Y_i(n)$ is calculated from the salience of the rule on the current trial, denoted by $Z_i(n)$. Initial saliences are pre-defined and in typical applications of COVIS one-dimensional rules are assigned equal, relatively high, saliences.

The salience of rule R_i used on trial n is adjusted after every trial depending on whether it resulted in a correct response or not. If the response was correct on trial n then the salience of the rule is adjusted by

$$Z_i(n+1) = Z_i(n) + \Delta_C \tag{3}$$

where Δ_C is a positive constant that represents the perceived reward associated with the correct answer. However, if rule R_i resulted in an incorrect response on trial n then the salience of that rule decreases by the rule

$$Z_i(n+1) = Z_i(n) - \Delta_E \tag{4}$$

where Δ_E is a positive constant that represents the perceived cost of an error on any trial. All the remaining rules in $\mathbf{R}$ keep their saliences from the previous trial.

The salience of each rule is then transformed to produce the weight, $Y_k(n)$, according to the following

1. For the rule R_i used on trial n

$$Y_i(n) = Z_i(n) + \gamma \tag{5}$$

   where γ is a constant that represents a participant's tendency to perseverate with a rule in light of receiving disconfirming feedback; as γ increases the participant is less likely to switch rules.

2. For a rule chosen randomly from $\mathbf{R}$, R_j, its weight is adjusted by

$$Y_j(n) = Z_j(n) + \mathbf{X} \tag{6}$$

   where $\mathbf{X}$ is a randomly distributed variable that has a Poisson distribution with mean λ. $\mathbf{X}$ represents the participant's tendency to select novel rules on each trial; the larger λ is, the more likely they are to switch rules. COVIS assumes that rule selection is mediated by the frontal cortex and that λ is related to dopamine levels in the cortex.

3. For the remaining rules

$$Y_k(n) = Z_k(n) \tag{7}$$

Finally, the rule to be used on the next trial is selected with probability

$$P_{n+1}(R_k) = \frac{Y_k^a(n)}{\sum_{s=1}^{m} Y_s^a(n)} \tag{8}$$

where a determines the decision stochasticity. If $a = 1$ then the rule to be used on the next trial is chosen probabilistically, with rules with higher weights being more likely to be picked. As a becomes larger than one, the rule with the largest weight becomes more and more likely to be chosen on each trial, so rule choice behaves more deterministically. Whereas the closer a is to 0, the smaller the differences between probabilities for each rule and so the decisions are more noisy and much less deterministic. Hélie et al. (2012a, 2012b) interpret a as a gain parameter and state a increases with cortical dopamine levels.

The implicit system

The implicit system of COVIS is based on a procedural learning mechanism that associates a response with each stimulus, whose two or more stimulus dimensions are integrated predecisionally. Broadly, this system consists of a representation of sensory information that leads to a hidden layer representing the striatum, which in turn leads to a decision making process in the prefrontal cortex. The sensory cortex is modeled by COVIS as an ordered array of up to 10,000 units. Each unit responds maximally to one particular stimulus and responds to a lesser extent to stimuli resembling it. The activation of each unit is calculated mathematically by a Gaussian function of the distance between the unit's preferred stimulus and the stimulus currently displayed, $d(K, \text{stimulus})$. So, the activation in sensory cortical unit K on trial n is given by

$$I_K(n) = e^{-\frac{d(K,\text{stimulus})^2}{\alpha}} \tag{9}$$

where α is a positive constant that scales the unit of measurement in stimulus space. The larger α is, the more similar the stimuli are to each other.

The activation of striatal unit $\mathcal{J}$ on trial n is determined by the weighted sum of the activations of all sensory units that project to it, which is formalized as

$$S_{\mathcal{J}}(n) = \sum_K w_{K,\mathcal{J}}(n) I_K(n) + \varepsilon_I \tag{10}$$

where $w_{K,\mathcal{J}}(n)$ is the strength of the synapse between cortical unit K and striatal cell $\mathcal{J}$ on trial n, $I_K(n)$ is the activation of sensory unit K on trial n and ε_I is normally distributed noise (with mean 0 and variance σ_I^2).

Then, the decision rule is

Respond A on trial n if $S_A(n) > S_B(n)$; otherwise respond B

The synapse strengths, $w_{K,\mathcal{J}}(n)$, are adjusted on each trial via reinforcement learning. The initial value, however, must be predetermined. In Ashby et al. (2011), $w_{K,\mathcal{J}}(0)$ was given by

$$w_{K,\mathcal{J}}(0) = 0.001 + 0.0025\ \mathrm{U} \tag{11}$$

where U is a constant sampled randomly from a uniform $[0,1]$ distribution. This means that all initial synaptic strengths will be between 0.001 and 0.0035 and will be randomly assigned. More recent applications of COVIS to experimental data have not defined this parameter.

Then, $w_{K,\mathcal{J}}(n)$ is adjusted on each trial as follows

$$\begin{aligned} & w_{K,\mathcal{J}}(n+1) = w_{K,\mathcal{J}}(n) \\ & + \alpha_w I_K(n)[S_{\mathcal{J}}(n) - \theta_{\mathrm{NMDA}}]^+[D(n) - D_{\mathrm{base}}]^+[1 - w_{K,\mathcal{J}}(n)] \\ & - \beta_w I_K(n)[S_{\mathcal{J}}(n) - \theta_{\mathrm{NMDA}}]^+[D_{\mathrm{base}} - D(n)]^+ w_{K,\mathcal{J}}(n) \\ & - \gamma_w I_K(n)[\theta_{\mathrm{NMDA}} - S_{\mathcal{J}}(n)]^+[S_{\mathcal{J}}(n) - \theta_{\mathrm{AMPA}}]^+ w_{K,\mathcal{J}}(n) \end{aligned} \tag{12}$$

where if $g(n) > 0, [g(n)]^+ = g(n)$, otherwise $[g(n)]^+ = 0$.

As Equation 12 is somewhat complex, further explanation is merited. Broadly, the first line describes the conditions under which synapses would be strengthened whereas lines two and three describe the conditions under which the connections would be weakened.

α_w, β_w, γ_w, θ_{NMDA} and θ_{AMPA} are constants. The first three are learning rates whereas the constants θ_{NMDA} and θ_{AMPA} represent the activation thresholds for post-synaptic NMDA and AMPA glutamate receptors respectively, where $\theta_{\mathrm{NMDA}} > \theta_{\mathrm{AMPA}}$ because NMDA receptors have a higher threshold for activation than AMPA receptors.

Equation 12 requires that we specify the amount of dopamine, $D(n)$, released on every trial in response to feedback. The amount of dopamine released is in turn dependent on the reward prediction error (RPE), which is determined by the following

$$\mathrm{RPE} = \text{Obtained Reward} - \text{Predicted Reward} \tag{13}$$

The reward obtained on each trial is dependent on the feedback given. For example, for applications where all stimuli are rewarded or punished equally, the obtained reward R_n on trial n is defined as $+1$ if correct feedback is given, 0 if no feedback is given or -1 if incorrect feedback is given.

The predicted reward on each trial is calculated using a simplified version of the Rescorla-Wagner model (Rescorla & Wagner, 1972). Assuming that the participant has just responded for the nth time to some stimulus then the reward they should expect to receive is given by COVIS as

$$P_n = P_{n-1} + 0.025(R_{n-1} - P_{n-1}) \tag{14}$$

Then, the dopamine release on each trial, $D(n)$, is calculated from the RPE using the following model

$$D(n) = \begin{cases} D_{\max} & \text{if RPE} > \frac{D_{\max} - D_{\mathrm{base}}}{D_{\mathrm{slope}}} \\ D_{\mathrm{slope}}\mathrm{RPE} + D_{\mathrm{base}} & \text{if } -\frac{D_{\mathrm{base}}}{D_{\mathrm{slope}}} \le \mathrm{RPE} \le \frac{D_{\max} - D_{\mathrm{base}}}{D_{\mathrm{slope}}} \\ 0 & \text{if RPE } \le -\frac{D_{\mathrm{base}}}{D_{\mathrm{slope}}} \end{cases} \tag{15}$$

where $D_{\max}$, D_{s} and D_{b} are constants.

Competition mechanism

COVIS uses a combination of the "confidence" each system has in its response, and the level of "trust" the competition system has for each system, to decide which system will guide responding. The confidence that each system has in its response is related to the degree of activation of each stimulus representation. The confidence in the verbal system equals the absolute value of the discriminant function, i.e. $|h_V(n)|$. If $|h_V(n)|$ is large then the stimulus is a long way from the decision bound and so the verbal system is more confident in its response, whereas if $|h_V(n)|$ is 0 then the stimulus is exactly on the boundary between two categories and the verbal system has no confidence on its categorisation. The confidence in the implicit system, $|h_I(n)|$, is equal to

$$|h_I(n)| = |S_A(n) - S_B(n)| \tag{16}$$

and follows a similar logic. If $|h_I(n)|$ is large then the implicit system favours one response much more than the other. However, if $|h_I(n)|$ is close to zero then both striatal units are activated equally and so the system has little confidence in its decision.

The degree of trust in each system is based on the past successes and failures of the system. The amount of trust in the verbal system is given by

$$\theta_V(n+1) = \theta_V(n) + \Delta_{OC}[1 - \theta_V(n)] \tag{17}$$

if the verbal system suggests a correct response on trial n. However, if the verbal system results in an incorrect response, the amount of trust in the verbal system on the next trial is given by

$$\theta_V(n+1) = \theta_V(n) - \Delta_{OE}\ \theta_V(n) \tag{18}$$

where Δ_{OE} is a parameter. The trust in the implicit system is given by

$$\theta_I(n+1) = 1 - \theta_V(n+1) \tag{19}$$

.

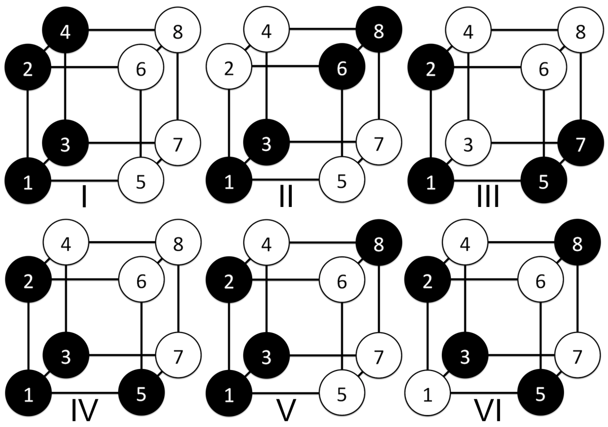

Figure 1: The six category types (I-VI) in Shepard, Hovland & Jenkins (1961). Each number represents a stimulus created from combining the three binary dimensions, which are represented by the sides of the cube. The colour of the circles indicate category membership.

Then, COVIS emits the response suggested by the verbal system if

$$\theta_V(n)\ |h_V(n)| > \theta_I(n)\ |h_I(n)| \tag{20}$$

else it emits the response suggested by the implicit system.

Simulation

The experiment

The category types in Shepard et al. (1961) represent the six qualitatively different ways that eight stimuli, generated from varying three binary dimensions, can be sorted into two, equally sized, categories. These six category structures used are detailed in Figure 1. These category types have been found to systematically vary in difficulty (Shepard et al., 1961; Nosofsky et al., 1994; Smith, Minda, & Washburn, 2004).

Category Type I is a one-dimensional rule and is learned the most easily. Category Type II is an XOR rule and is the next easiest (although see Kurtz, Levering, Stanton, Romero, & Morris, 2013, for a discussion on the relative differences between Types II and IV). Category Types III, IV and V require participants to attend to all three stimulus dimensions and are the next most difficult. Finally, the Type VI category structure has no simplifying rule, is the hardest to learn, and it is generally assumed that the participants learn the stimulus-category assignments separately for each stimulus.

Summary

COVIS is able to predict the pattern of category learning difficulty found in Shepard et al. (1961). However, in order to capture this pattern of data, several adaptations had to be made to the model.

Verbal system adaptations Two changes were made to the verbal system. Although these modifications to the formal model are unique to the current paper, they seem broadly consistent with previously published heuristic descriptions of the COVIS verbal system.

As mentioned above, previous simulations of experimental results using COVIS have limited possible rule selection to only the relevant one-dimensional rules (Ashby et al., 2011; Hélie et al., 2012a, 2012b). However, when using only the 6 available one-dimensional rules (3 dimensions x 2 category mappings) COVIS was unable to capture the ordinal results of Shepard et al. (1961). Specifically, under these conditions COVIS misplaces the difficulty of the Type II problem, placing it as difficult as the Type VI problem.

To rectify this, the model was extended to also include conjunction and disjunction of conjunction rule types, with varying initial saliences. The one-dimensional rules were the most salient, followed by the conjunctions and finally by the disjunction of conjunction rules. We do not include disjunctions as they are behaviorally identical to conjunction models.

The inclusion of rules with different initial saliences implied a knock-on change to part of the rule selection mechanism. It seemed inconsistent with the general operation of the model that rules with different initial saliences would be equally likely to receive the boost produced by Equation 6. Therefore, in this implementation, random selection was weighted by current rule salience.

Implicit system adaptations Two modifications were made to the implicit system. The first was that this simulation included the radial basis function (as stated in Equation 9). This has always been stated as part of the implicit system of COVIS and we state our use explicitly here only because none of the previously simulations have actually used it! (Ashby et al., 2011; Hélie et al., 2012a, 2012b). These previous studies assumed that the stimuli were not confusable and therefore that each sensory unit would respond only to its preferred stimulus. A similar argument could be made here. However, as this simulation aimed to test whether COVIS was capable of learning categories, it seemed important to include the generalizing function as generalization is key to the definition of what it is to have learned a category.

Our second adaptation was to provide separate feedback was given to the implicit system, i.e. feedback as to whether or not the implicit system was correct, not whether the overall response, as determined by the competition system, was correct or not. Providing separate feedback has been included in previous implementations of COVIS, with little explanation as to why (Hélie et al., 2012a, 2012b).

Simulation

1000 simulations were run for each category type using the model and adaptations described above. As in previous reports of the empirical data Nosofsky et al. (1994), blocks 1 and 2 contained 8 trials whilst the remainder contained 16 trials. The stimuli were presented to the verbal system as

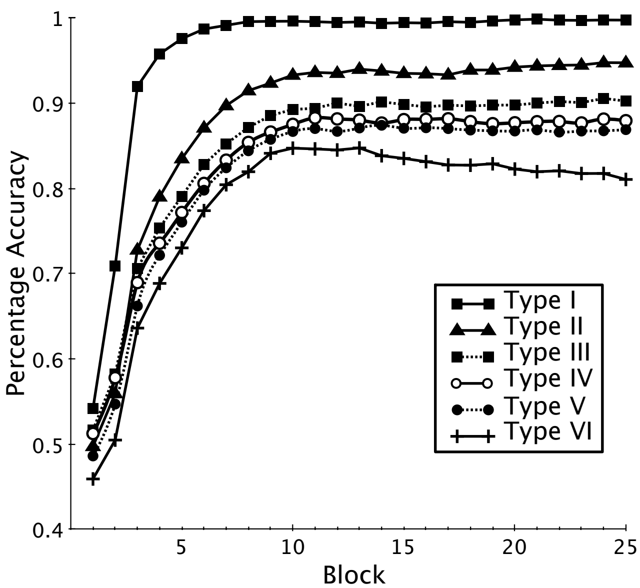

Figure 2: COVIS's predictions of learning the six structures in Shepard, Hovland and Jenkins (1961).

vectors that varied on three binary dimensions as per previous simulations of COVIS (Ashby et al., 2011; Hélie et al., 2012a, 2012b). However, unlike previous simulations and as discussed above, stimuli were presented to the implicit system as vectors that varied on 8 dimensions, corresponding to the eight possible stimuli, with values determined according to Equation 9. This was so that the verbal system received a feature based representation of the stimulus and the implicit system received an object-based representation which included some generalisation across stimuli. The parameters used are displayed in Table 1. No optimisation procedures were used, rather parameters estimates were taken from previous COVIS simulations and adjusted until the predictions shown in Figure 2 were generated.

The results indicate that COVIS can capture the ordinal pattern of difficulty of the six category structures found in Shepard et al. (1961).

Table 1: Parameter values for all learning task simulations.

Parameter	Value	Parameter	Value
Verbal System		**Implicit System**	
σ_V	0	α	0.14
δ_C	0.1	σ_I	0.0125
δ_E	0.2	α_w	0.65
γ	1	β_w	0.1
λ	5	γ_w	0.02
a	10	θ_{NMDA}	0.0022
		θ_{AMPA}	0.01
Competition System		D_{base}	0.2
Δ_{OC}	0.2	D_{slope}	0.8
Δ_{OE}	0.9	D_{max}	1

Discussion

This simulation demonstrates that COVIS is capable of capturing the ordinal pattern of difficulty found by participants in learning the six distinct category structures described by Shepard, Hovland and Jenkins (1961). However, it is interesting to note that although COVIS captured the ordinal pattern, there were still some quantitative discrepancies between this simulation and behavioral replications of the original study. For example, Nosofsky et al. (1994) demonstrated that well before the end of the experiment category types II to V had reached maximal accuracy, whereas COVIS predicts that accuracy does not exceed approximately 85 and 95%. One might argue that this is due to slower learning on the part of COVIS relative to the participants. However, as learning appears to have reached a plateau, this seems unlikely. Another possibility is that there are other parameters values that could allow COVIS to more precisely capture learning of these categories. This seems more likely due to the lack of formal optimization in the current demonstration; an exhaustive search of the parameter space could conceivably discover improved parameter estimates that would enable COVIS to predict the quantitative aspects of Nosofsky et al. (1994) more closely. However, we would also argue that capturing the ordinal aspects of a data set are more important than the precise level of quantitative fit, at least in the early stages of model testing and comparison (Wills & Pothos, 2012).

It should also be emphasised that this and the simulations conducted by Ashby and colleagues (Ashby et al., 2011; Hélie et al., 2012a, 2012b) are first steps in determining whether the formal aspects of COVIS provide the best description of the mechanisms of category learning. Reasons to be cautious come from both simulations and behavioral work conducted to test the COVIS theoretical framework. First, it is possible that the predictions from this simulation and the others mentioned above are parameter dependent. Although COVIS has been shown to be sufficient to capture the ordinal predictions of several experiments, this only determines the model's behaviour at one point in parameter space for each experiment (Pitt, Kim, Navarro, & Myung, 2006). A harder, and more important, question is whether there is a set of parameters that would permit COVIS to capture the ordinal patterns of several experiments simultaneously (as recommended by Wills and Pothos, 2012). One way future work could address this issue would be to employing analysis techniques that look at the model's behaviour across all parameter values, such as parameter space partitioning (Pitt et al., 2006) or landscaping (Navarro, Pitt, & Myung, 2004).

Conclusion

In conclusion, this simulation demonstrates that COVIS is able to capture the order of difficulty of the six category types found by Shepard et al. (1961). Although encouraging, further research is needed to establish whether COVIS can accommodate the broad range of well-established phenomena already present in the category learning literature. Establishing this seems likely to be of similar importance to testing novel predictions of the model.

References

Ashby, F. G., Alfonso-Reese, L. A., Turken, U, & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*(3), 442–481.

Ashby, F. G., & Maddox, W. T. (2005, February). Human Category Learning. *Annual Review of Psychology*, *56*(1), 149–178.

Ashby, F. G., & Maddox, W. T. (2010, December). Human category learning 2.0. *Annals of the New York Academy of Sciences*, *1224*(1), 147–161.

Ashby, F. G., Paul, E. J., & Maddox, W. T. (2011, February). COVIS. In *Formal approaches in categorisation* (pp. 1–13). New York: Cambridge University Press.

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

Edmunds, C. E. R., Milton, F., & Wills, A. J. (2015). Feedback can be superior to observational training for both rule-based and information-integration category structures . *The Quarterly Journal of Experimental Psychology*, *68*(2), 1203–1222.

Hélie, S., Paul, E. J., & Ashby, F. G. (2012a, July). A neurocomputational account of cognitive deficits in Parkinson's disease. *Neuropsychologia*, *50*(9), 2290–2302.

Hélie, S., Paul, E. J., & Ashby, F. G. (2012b). Simulating the effects of dopamine imbalance on cognition: From positive affect to Parkinson's disease. *Neural Networks*.

Kurtz, K. J., Levering, K. R., Stanton, R. D., Romero, J., & Morris, S. N. (2013). Human learning of elemental category structures: Revising the classic result of Shepard, Hovland, and Jenkins (1961). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 552–572.

Lewandowsky, S., Palmeri, T. J., & Waldmann, M. R. (2012). Introduction to the Special Section on theory and data in categorization: Integrating computational, behavioral, and cognitive neuroscience approaches. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(4), 803–806.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A Network Model of Category Learning. *Psychological Review*, *111*(2), 309–332.

Navarro, D. J., Pitt, M. A., & Myung, I. J. (2004, August). Assessing the distinguishability of models and the informativeness of data. *Cognitive Psychology*, *49*(1), 47–84.

Newell, B. R., Moore, C. P., Wills, A. J., & Milton, F. (2013). Reinstating the frontal lobes? Having more time to think improves implicit perceptual categorization: A comment on Filoteo, Lauritzen, and Maddox (2010). *Psychological Science*, *24*(3), 386–389.

Nosofsky, R. M., Gluck, M. A., Palmeri, T. J., McKinley, S. C., & Glauthier, P. (1994, October). Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961). *Memory & Cognition*, *22*(3), 1–20.

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

Rescorla, R. A., & Wagner, A. R. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, 64–99.

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied*, *75*(13), 1–42.

Smith, J. D., Minda, J. P., & Washburn, D. A. (2004). Category Learning in Rhesus Monkeys: A Study of the Shepard, Hovland, and Jenkins (1961) Tasks. *Journal of Experimental Psychology: General*, *133*(3), 398–414.

Stanton, R. D., & Nosofsky, R. M. (2007, January). Feedback interference and dissociations of classification: Evidence against the multiple-learning-systems hypothesis. *Memory & Cognition*, *35*(7), 1747–1758.

Stanton, R. D., & Nosofsky, R. M. (2013). Category number impacts rule-based and information-integration category learning: A reassessment of evidence for dissociable category-learning systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(4), 1174–1191.

Wills, A. J., & Pothos, E. M. (2012). On the adequacy of current empirical evaluations of formal models of categorization. *Psychological Bulletin*, *138*(1), 102–125.