Procedural and Declarative Category Learning Simultaneously Contribute to Downstream Processes

Priya B. Kalra (pkalra7@uwo.ca), Laura J. Batterink (lbatter@uwo.ca), J. Paul Minda (jpminda@uwo.ca)
Western Institute for Neuroscience, Perth Drive London, ON N6G 0B9 Canada

Abstract

Studies on interactions between procedural and declarative learning have focused on largely on competition during encoding, consolidation, or use (retrieval). Less attention has been paid to interactions between the representations created by each system. In a behavioral study, we demonstrated that information from both declarative and procedural learning can contribute to response selection. Participants were instructed to use a completely diagnostic, verbalizable, shape-based rule to categorize exemplars and received feedback after each trial. However, the categories also differed probabilistically in their color distributions. Participants used both color (learned procedurally) and shape (learned declaratively) to categorize exemplars, making faster responses when both sources indicated the same category judgement, and slower when they conflicted. Debriefing confirmed that most participants were unaware of the color distributions (aware participants were analyzed separately). This result suggests that both the color (procedural) and shape (declarative) information contributed to response selection.

Keywords: memory, category learning, skill learning, knowledge representation, instruction

Introduction

Many mechanisms have been proposed for the process of category learning, and many models suppose a combination of at least two mechanisms (Pothos & Willis, 2011). One such pair of mechanisms is informed by the cognitive neuroscience of memory, procedural/implicit and declarative/explicit. Considerable evidence supports the existence and dissociability of these two forms of category learning, and their relations to procedural and declarative memory systems in the brain (Ashby et al., 1998, 2003; Ashby & Ell, 2001; Ashby & Maddox, 2011; Ashby & O'Brien, 2005; Filoteo et al., 1998, 2005; Knowlton et al., 1994, 1996; Maddox & Ashby, 2004; Nomura et al., 2007; Price et al., 2009; Squire & Knowlton, 1995).

Evidence also suggests that both procedural and declarative/explicit category learning can take place simultaneously, over the same stimulus set (Crossley & Ashby, 2015; Foerde et al., 2006). However, some authors suggest that although learning can take place simultaneously and without interference, only one system contributes to downstream processes such as response selection and decision-making (e.g., Crossley & Ashby, 2015).

Formally, COVIS (competition between verbal and implicit systems) provides a model of these two systems and how they could interact. COVIS integrates the traditions of research in memory and cognitive neuroscience (Ashby & Ell, 2001; Ashby & O'Brien, 2005) and can explain many of the observed phenomena. In the COVIS model, a procedural module and an explicit[1] module each internally reach a category decision, and submit their respective decisions to a gating mechanism, along with a confidence rating. The gating mechanism has a trust or bias parameter θ assigned to each module (based on its success rate during training), and combines this with the confidence rating (h) from each module to choose the category decision with the highest product of confidence and trust[2].

Confidence from the explicit module is based on distance from the category boundary (greater distance→more confident) and confidence in the implicit module is based on the absolute value of the difference between the probability of assigning each category ($|S_A(n) - S_B(n)|$). This gating mechanism can account for observed effects such as trial-by-trial switching between modules (Turner et al., 2017), dominance of one system in response selection despite learning by both systems (Crossley & Ashby, 2015), and gradual shifts over the course of training from dominance of one system to another (Poldrack et al., 2005). Notably, however, the gating mechanism as currently specified does not consider the contents (category decision) of each module, nor whether these are congruent with each other (both submit same category) or not. To the extent that the difficulty or degree of competition in a trial could be modeled by COVIS, it would take into account only the trust in and confidence of each module, not the category decision or the congruence of the two category decisions.

However, outside of the laboratory, we observe scenarios in which information from both systems appears to be used simultaneously: for example, when making diagnoses, medical experts seem to use a combination of both conditional reasoning based on declarative knowledge as well as probabilistic reasoning based on experience (Norman & Brooks, 1997). Similarly, professional musicians performing from memory seem to simultaneously draw on both a non-verbalizable representation of motor sequences as well as a verbalizable understanding of the structure, form, and meaning of a musical piece (Chaffin, Logan, & Begosh, 2009, cited in Reber, 2013).

[1] For simplicity we will use declarative/explicit and procedural/implicit interchangeably.

[2] Formally, this can be expressed as:
(Eq 1) $|h_E(n)|\ \theta_E(n) <> |h_P(n)|\ \theta_P(n)$

where $\theta_e(n)$ = trust in explicit input on the current trial, $h_e(n)$ = confidence of explicit input on the currrent trial, $\theta_p(n)$ = trust in procedural input on current trial, and $h_p(n)$ = confidence of procedural input on the current trial

To our knowledge, no study has been specifically designed to test whether information from both systems simultaneously contributes to a response within a single trial. However, some studies have hinted at this possibility. For example, Brooks and Allen (1991) observed that even though participants were given a perfectly predictive classification rule, their responses on a transfer task were nevertheless affected by similarity to previously seen exemplars along irrelevant dimensions—almost as if the similarity information were "contaminating" the rule-based classification. This study was originally designed to contrast rule-based and exemplar-based learning, not declarative and procedural learning, so it is possible that that the generalization based on exemplars could have been mediated by either procedural or declarative mechanisms.

Similarly, Schoenlein and Schloss (2022) trained participants to classify stimuli based on a completely diagnostic shape difference, but using a cool-biased color distribution for one category and a warm-biased color difference for the other category. Participants were able to use the shape rule effectively, and they did not demonstrate explicit knowledge of color differences between categories. However, after training, participants were asked to rate how associated different colors were with each category (using the category names) on a continuous scale labeled from "not at all" to "very much." The participants rated cool colors as more associated with the cool-biased category and warm colors as more associated with the warm-biased category, even generalizing to warm and cool colors that had not been included in the training set. Again, these findings demonstrate that unconscious learning of probabilistic information can occur simultaneously with explicit rule use. However, since the color-category associations were assessed outside of the categorization task, this study could not test whether this probabilistic color information directly influenced category response selection itself.

More closely, Batterink and colleagues (Berger & Batterink, under review; Batterink et al., 2014) found that participants who were taught an explicit rule governing the use of novel articles showed an effect of the training distribution along a covert, second dimension (noun animacy), as reflected by reaction times, even when they were not aware of the article-animacy contingency. This finding shows that information acquired without awareness can affect intentional response decisions that are based on a deterministic rule, delaying response times when the two forms of information conflict. These results suggest that both information that the participant is aware of, as well as information that the participant is not aware of, are interacting at the point of response selection. In the current study, we attempt to extend this paradigm to category learning.

The goal of the current study is to observe whether procedural and declarative knowledge contribute simultaneously to categorization response selection behavior. It can be difficult to disentangle the contributions of multiple learning systems in a single task because the systems usually converge on common responses. Here, we have designed a stimulus set in which information learned by each system can support different responses, and we have created trials in the test phase that are designed to maximize such divergence.

For the declarative system, we provide—explicitly, verbally, and with examples—a deterministic, verbalizable category rule based on feature combinations. For the procedural system, the distribution of colors differs probabilistically between the two categories. If an association between color distribution and category membership is learned, it must be learned gradually, with immediate feedback. We also include measures of whether the color-category knowledge is available to awareness or not. If it is learned gradually, based on feedback, and without awareness, we may infer that this learning is mediated by the striatal procedural system. The structure and task demands of the probabilistic color-category association are very much like the probabilistic classification task, which has been reliably demonstrated to use the striatal procedural system (Knowlton, Mangels, & Squire, 1996).

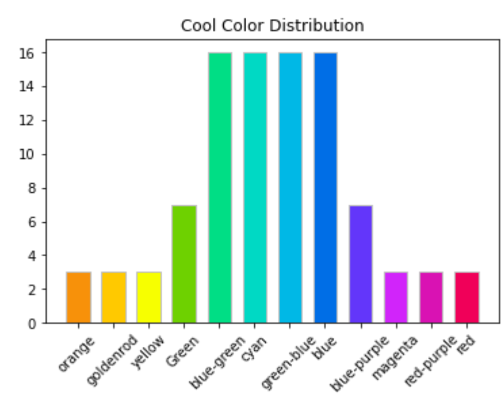

Figure 1a: Cool-biased color distribution.

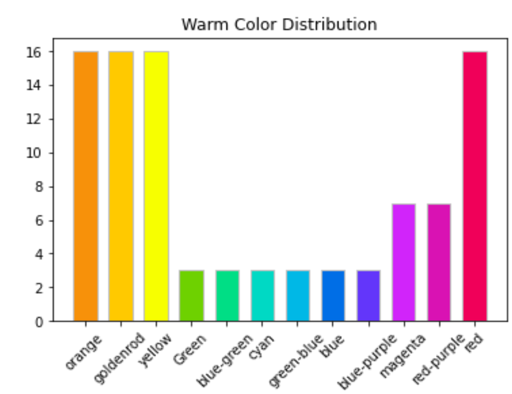

Figure 1b: Warm-biased color distribution.

During the test phase, if we see a difference between trials in which procedural and declarative learning point to the same response (congruent) compared to those trials in which they point to different responses (incongruent), we will conclude that both sources of information contribute to response selection within a given trial. If we do not observe such a difference, we will conclude that information from both systems is not used simultaneously in response selection, though this does not rule out the possibility that parallel encoding may have taken place.

Methods

Participants

A total of 249 undergraduates (ages 17 to 25, mean age 18.4 years; 44% male, 54% female, 2% other/decline to state/non-binary/genderqueer) participated via an online platform for course credit. After all exclusions (see below), the final sample size for analysis was n=190.

Materials

Alien Stimuli

"Alien" images were created using custom code and Python's PIL package (https://pypi.org/project/pillow/) to combine geometric figures (ovals, rectangles, etc.) Each alien consisted of a large oval with some configuration of the following features: number of eyes (1-4), mouth type (round or angular), nose type (round or angular), ears (present or absent). All thirty-two possible combinations of these features were generated and used in the experiment (see Figures 2 & 3 for example images).

Categorization Features and Explicit Rule. Stimuli were divided into two categories based on an exclusive or (XOR) rule over the eye and mouth dimensions. Stimuli in Category A had either a round mouth and odd number of eyes OR a square mouth and even number of eyes. Conversely, stimuli in Category B had either a square mouth and odd number of eyes OR a round mouth and even number of eyes. Thus, each category included 16 possible configurations of features (including both the diagnostic and nondiagnostic feature dimensions [nose and ears]). The XOR rule was chosen for two reasons. First, an XOR rule is difficult or impossible to learn from feedback alone, so any use of the XOR rule could be assumed to be via the declarative system. Second, the complexity of the XOR rule requires considerable working memory and attention resource allocation, so it was unlikely that participants could attend to and become aware of the biased color distributions.

Color selection and distribution. Each configuration of features (category token) was then generated in a variety of colors. Colors were divided into warm and cool colors based on their hue angle. Even hue-angle spacing was used to choose hue angles of 0, 30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 330 (where 0 corresponds to red and 180 to cyan). Warm colors were defined as those within 90 degrees (+/-) of 0; cool colors were defined as those greater than 90 degrees (+/-) from 0. For each hue angle position, saturation (chroma) was adjusted to minimize differences in saturation across hue angles and between warm and cool color groups. Given the inherent asymmetry of the visual color space, it was not possible to equalize luminance (brightness) between warm and cool colors; cool colors were systematically lower in luminance than warm colors. See Figures 1a & 1b. The colormath package in Python (Taylor, 2018) was used to select colors in CIELCh$_{uv}$ colorspace (CIE, 1986) and to calculate the distances between colors in that colorspace (delta e).

Training phase color distributions.

Unbeknownst to participants, stimuli in the training condition followed biased color distributions. Two color distributions were created: one warm-biased and one cool-biased, and each was assigned to a category in counterbalanced fashion (i.e., for half of participants, Category A followed the warm-biased distribution and Category B the cool-biased distribution, and vice versa for the other half). Each training color distribution included a total of 88 tokens distributed across all 12 hue values. Each distribution contained 64 congruent (e.g., warm colors in the warm distribution) and 24 incongruent (e.g., cool colors in the warm distribution) items. Within each category, each color was also distributed evenly across subcategory and number of eyes. The distribution of non-diagnostic features (nose type and ears/no ears) was also matched between the two categories.

Procedure

Overview

Participants completed the task on their personal computers (option to complete the task on tablet or phone was disabled) via the Pavlovia online testing platform (www.pavlovia.org). Both training and test phase tasks were created using PsychoPy (Peirce et al., 2019). After giving informed consent by selecting an online box in Qualtrics, participants viewed an explanation of the explicit rule on a screen with visual examples. Participants then progressed through the training phase, followed by the test phase. Finally, participants answered demographic and debriefing questions and completed a brief test of color-blindness using Ishihara plates. All instructions were provided in text form on the computer screen.

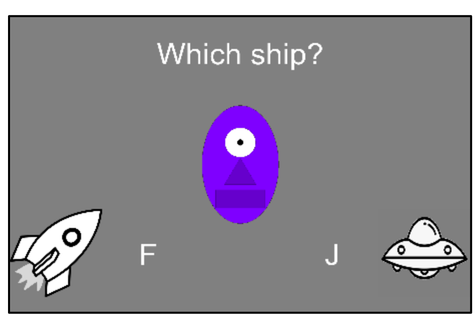

Figure 2: Screenshot of training phase trial.

Training phase

Participants were given a cover story in which they were asked to match aliens to their appropriate vehicles—aliens of one category used rockets, while the other category used saucers (no verbal labels were given for vehicle type). Participants were instructed on the explicit rule through slides that explained the rule (e.g. "Group 1 aliens have square mouths and an odd number of eyes") and provided examples. The training task then began. A short practice block of 12 trials preceded the four training blocks. In all other ways practice trials were identical to training trials.

On each trial, participants were shown an alien and two images of spaceships—one rocket, one saucer-shaped. The rocket was always on the left side of the screen and the saucer on the right. Participants categorized each alien by choosing rocket or saucer (left/right) with a key press (f/j). Participants received feedback in the form of the words "correct" or "incorrect" displayed on the screen. Feedback was based solely on the explicitly instructed XOR mouth/eyes rule. Since color correctly predicted category membership on 64/88 training trials, the maximum accuracy that could be reached using only color was less than or equal to 64/88 = .73. Category-ship combination was counterbalanced across participants, so for half the participants Category A aliens used the rocket and Category B aliens used the saucer, and vice versa for the other half. To encourage accurate performance, incorrect responses were followed by a 3 second delay with countdown. The task screen also included a "power bar" showing cumulative accuracy; in the pre-task instructions, participants were told that cumulative accuracy above 70% was necessary "to win the game."

Stimuli were presented in 4 blocks of 44 trials; each block was roughly even in terms of category, color, subcategory, and non-diagnostic features. Frequency of individual features (e.g. number of eyes) was balanced across categories to deter the formation of individual feature-color associations[3]. Stimulus order was pseudorandom such that no more than 3 trials from the same category appeared consecutively and consecutive same-color or same-eye-number trials were similarly limited. Trial order within a block was fixed, but block order was random across participants.

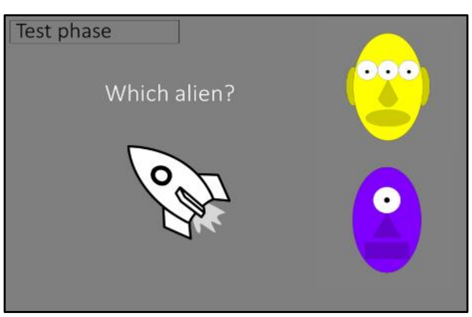

Figure 3: Screenshot of test phase trial.

Test phase

On each trial, participants were shown two aliens, one from Category A and one from Category B, and one ship, and were instructed to choose which alien corresponded to the presented ship. To avoid confusion with the training left/right configuration of the ships, the aliens were stacked vertically and response keys were u/n (upper or lower).

Participants were presented with a total of two blocks of 38 trials each (one saucer block, one rocket block). In half the trials, the colors of both aliens were congruent with the training color distribution (e.g., A-warm/B-cool); in the other half of trials, both aliens were presented in incongruent colors. All stimuli presented were novel (i.e., not previously seen in the training task).

Pairs in the congruent and incongruent conditions were balanced for total shared features, shared diagnostic features, shared non-diagnostic features, binned warmness difference, and mean distance between colors in colorspace (delta e).

Post-tests and survey questions

Colorblindness items. Participants were asked to type the numbers visible to them in a set of 5 Ishihara plates selected to probe for deficiencies in color vision. Before any exclusions, about 1% of participants responded to the Ishihara plates in ways consistent with some form of colorblindness. However, performance for these participants was comparable to that of other participants and the main RT difference between congruent and incongruent trials at test was similar in colorblind and non-colorblind participants. It is likely the case that they were able to perceive the differences in color distributions based on brightness differences even though they may not have been able to perceive all hue differences. For this reason, we did not exclude participants on the basis of their Ishihara plate responses or self-reports of colorblindness.

Strategy use and color awareness questions. Participants were also asked to complete the following open-ended questions:

1. Describe the rule you used to classify the aliens (in your own words, to the best of your ability).
2. Other than the rule you were instructed to use, did you use any strategy or rule of thumb to decide which aliens went with which ships? (if yes, please describe briefly if you can)
3. Did you notice anything about the colors of the aliens? If yes, please describe below
4. Describe what (if anything) you noticed about the colors of the aliens.
5. Did you use the colors to help the aliens find their ships? (yes/no)

[3] If an association between color and a particular combination of diagnostic features were learned (e.g., aliens with square mouths and odd number of eyes are typically green/blue) that would be in essence learning of the rule-color association.

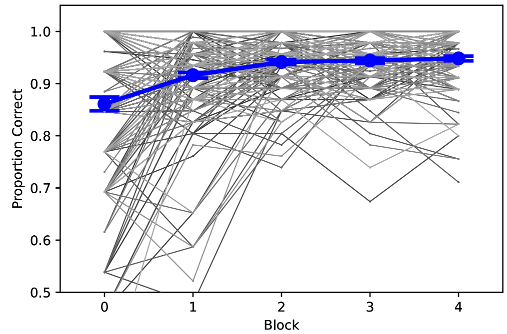

Figure 4a: Training phase accuracy by block

As we were primarily interested in effects of implicit sensitivity to the color distribution on categorization performance, responses to any of these questions that suggested a use of color to classify stimuli, or awareness of the biased color distributions resulted in the participant's data being excluded from further analysis.

Participant exclusions

Low performance. Thirteen (13) participants performed below 70% accuracy on the training phase and were excluded from further analysis on the assumption that they had not engaged with the task in earnest. In addition, 17 participants were excluded from test phase analysis for test phase scores lower than 70% accuracy. Test phase performance was used to exclude participant because data were collected online and low performance was used as a marker for lack of engagement with the task (confirmed by unusually low reaction times coupled with low accuracy). The distribution of test phase scores was bimodal, such that most participants performed very well (>80% accuracy) but a significant cluster performed in the chance range (40-60% accuracy).

Survey responses. In addition, participants were excluded for the following reasons based on their responses to open-ended questions. Participants whose responses to survey questions indicated any type of color-category association (n=9) and/or who answered "Yes" to "did you use color to categorize the aliens" (n=11) were excluded from further analysis (total n=17). After these exclusions, the final sample size for accuracy analysis was 190**.**

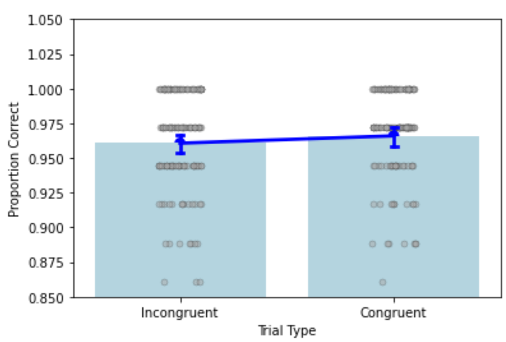

Figure 5a: Test phase accuracy by condition

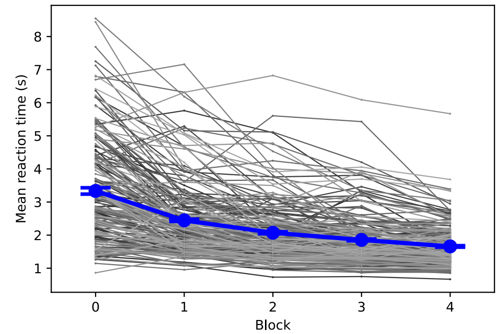

Figure 4b: Training phase reaction time by block.

Reaction time data cleaning and analysis. Before reaction time by participant and condition was analyzed, reaction time data were cleaned as follows: 1) only correct trials were included 2) each participant's mean and SD for RT were calculated; trials that were above or below 3SDs of that participant's mean were excluded. 3) participants with exceptionally long (mean >6 seconds) or exceptionally variable (SD > 6 seconds) were excluded; these cutoffs were determined by visual inspection of the histograms for participant means and standard deviations of reaction time, respectively[4]. After these exclusions, 173 participants were included for reaction time analysis.

Stimuli, code for creating stimuli, analysis scripts, and de-identified data (including full survey responses) can be found at:

https://osf.io/cb3zu/?view_only=e7eab3114113414285431fa6d53b4f26

Results

Task performance

Training phase

Mean accuracy was computed for each participant and each block and analyzed using a one-way ANOVA with blocks (1-4) as a within-subjects factor**.** Mean reaction time was computed for each participant and was analyzed using a one-way ANOVA with blocks (1-4) as a within-subjects factor. Mean accuracy (M=.93 SD= .05) was generally high and mean reaction time (M = 2.07s. SD =1.46s) on the training task was within expected limits. In addition, performance improved over training blocks, as seen in

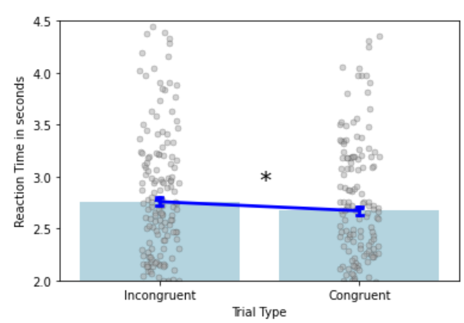

Figure 5b: Test phase reaction time by condition.

[4] Without excluding participants with exceptionally long or variable reaction times, the main effected reported for this experiment is still significant: ($M_{INCONGRUENT}$ =3.28s, $M_{CONGRUENT}$ = 3.07s; M_{RTDIFF} = .143s, t(189)= 4.005, p<.0001).

significant effects of block on accuracy (increasing, $F(4, 760) = 47.81$, $p<.001$) and reaction time (decreasing, $F(4, 756) = 228.43$, $p<.001$) See Figures 4a and 4b[5].

Test Phase

Overall performance. Mean accuracy in the test phase was comparable to training phase accuracy (M =.96, SD = .19). Mean response times were longer than in the training phase (M = 2.90s SD = 2.54s) after reaction time data cleaning (see above).

Test Phase Differences by Condition. We compared accuracy and reaction time in the test phase across the main condition of interest (congruency).

Accuracy analysis. Although accuracy was numerically higher for congruent trials than incongruent trials, this difference did not reach significance ($M_{\text{CONGRUENT}} = .966$, $M_{\text{INCONGRUENT}} = .960$; $t(189) = 1.75$, $p = .08$). See Figure 5a.

Reaction time analysis. Participants' RTs were significantly slower for incongruent than congruent trials: ($M_{\text{INCONGRUENT}}$ =2.862s, $M_{\text{CONGRUENT}}$ = 2.719s; M_{RTDIFF} = .143s, $t(172)= 4.68$, p<.0001). Figure 5b shows the comparison of RTs across condition, within participant.

Strategy use

87% of participants gave some response to the question "Describe the rule you used to classify the aliens (in your own words, to the best of your ability)." Of these, 85% gave a response that referred generally to the eyes and mouth or to the parity of the eyes and the shape of the mouth. An additional 3% referred to using the rule given. We interpret these responses as evidence that shape (XOR rule)-based classification in these participants was implemented by declarative mechanisms.

Discussion

In the current study, we have demonstrated that the contents—not merely the engagement—of both procedural and explicit categorization modules can contribute to categorization behavior and downstream processes such as decision-making and response selection. If both modules suggest the same category (congruent condition), response selection is facilitated (faster reaction time); if the modules suggest conflicting categories, then response selection is hindered (slower reaction time).

Very few participants reported any explicit knowledge of the biased color distributions (<10%); those who did were excluded from the analysis intended to show interaction between explicit and procedural information. For this reason, we feel comfortable interpreting the effects of the color distribution as contributions of the procedural category learning module, despite the fact that the reversed form of the test phase could potentially disadvantage a procedural learning system (Jacoby, 1991; Fincham & Anderson 1994; Vaquero et al, 2020). We interpret the robustness of the effect despite this reversal as evidence of abstract learning beyond stimulus-response association (Reber,1991;Seger, 1994).

In contrast, in their debrief responses almost all participants explicitly referred to the eyes and mouth (diagnostic features), or to "the rule that was given," and sometimes reported verbalizable heuristics based on the XOR rule. Furthermore, performance accuracy in the first block of training trials is well above chance: this pattern suggests that participants are applying an explicit rule that they already know (from the instructions provided) rather than searching for a rule or gradually accumulating information about the shape-category relation.

All stimuli shown in the test phase were novel, i.e. not shown in the training phase, so participants could not use memory for specific exemplars to categorize the test stimuli. It is possible that they were comparing each test stimulus to stored representations of previous stimuli (exemplars); however, the high accuracy levels, particularly early in training, suggest that a similarity-based approach alone cannot explain the results. Furthermore, the nature of the XOR rule deters similarity-based reasoning based on shape features and is difficult to learn inductively.

With regard to the COVIS model, some results support the current specification, but some argue for adjustments to the gating mechanism. Specifically, the high accuracy (defined by responses consistent with the explicit rule) is consistent with COVIS because the trust parameter for the explicit module should be high ($\theta_E = 1$) by the end of training while the trust parameter for the procedural module must be lower since it only inconsistently ($\theta_P < .73$) predicts the correct (reinforced) response.

However, the difference in reaction times between congruent and incongruent trials is not predicted by COVIS; the gating mechanism in COVIS does not consider the actual category choices of each model, nor whether they are the same, but only the confidence from each model and the trust parameter for each model. In our stimulus set, there are examples of stimuli that could be considered high and low confidence for each module, but these are distributed evenly across congruent and incongruent trials. In other words, in our stimulus set congruency and confidence are uncorrelated, so the observed difference in reaction time based on congruency cannot be explained away by confidence.

In order for COVIS to accommodate the current findings, the gating mechanism would have to be modified to consider not only the confidence and trust in each module, but also whether the results of each module are consistent with each other. This could be accomplished by modeling the gating mechanism as receiving a probability for a given category decision from each module, and combining them to reach a final decision. The ATRIUM model uses a "mixture of experts" gating mechanism like this (Krushke, 2011), but models a different set of category learning modules (not procedural and explicit).

[5] Congruency effects appeared to develop over the course of training such that incongruent trials were (on average) slower and less accurate by block 4 of training, but these differences did not quite reach statistical significance

Substantively, our findings suggest that representations formed by both procedural and declarative learning may simultaneously contribute to categorization and response selection behavior. Although many forms of interaction between procedural and declarative learning have been observed (see e.g. Freedberg, 2020), to our knowledge this finding represents a completely novel form, specifically that the contents, not only the activity or confidence of each system, interact.

The implication of this finding is that in complex, real-world situations, both rule-based or explicit decision making and procedural and similarity-based decision making may interact. For example, in formal instruction students are often given rules or necessary and sufficient criteria for category membership (e.g. mammals have hair and live young etc.), but if they experience only a biased selection from the space of possible category members (e.g. only dogs and cats as mammals), they may have difficulty transferring their knowledge of mammals in general to unfamiliar exemplars, despite the fact that they have an explicit, perfectly diagnostic rule.

Similarly, it is possible that results such as those of Gleitman, Armstrong & Gleitman (1983) could be reinterpreted as an example of behavior influenced by both explicit (deterministic) and implicit (probabilistic) representations for the same category. The current results cannot reveal whether different learning modules or memory systems create discrete or overlapping representations of the stimulus space, but our in-progress fMRI-RSA study is designed to answer this question.

In another example, the current findings provide further evidence that behavior can be based on a combination of both implicit representations based on accumulated experiences and explicit understandings or beliefs, as seen in many studies using the Implicit Association Test (Greenwald & Banaji, 2017).

Thus, the findings of the current study not only deepen our understanding of how memory systems may interact, but also provide a potential mechanism to explain and predict behavior in the real-world.

References

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

Ashby, F. G., & Ell, S. W. (2001). The neurobiology of human category learning. Trends in Cognitive Sciences, 5(5), 204–210. https://doi.org/10.1016/S1364-6613(00)01624-7

Ashby, F. G., & Ennis, J. M. (2006). The role of the basal ganglia in category learning. *Psychology of Learning and Motivation*, *46*, 1-36.

Ashby, F. G., Noble, S., Filoteo, J., Waldron, E. M., & Ell, S. W. (2003). Category learning deficits in Parkinson's disease. *Neuropsychology*, 17(1), 115–124. https://doi.org/10.1037/0894-4105.17.1.115

Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy of Sciences*, *1224*, 147–161.

Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. TRENDS IN COGNITIVE SCIENCES, 9(2), 83–89. https://doi.org/10.1016/j.tics.2004.12.003

Batterink, L. J., Oudiette, D., Reber, P. J., & Paller, K. A. (2014). Sleep facilitates learning a new linguistic rule. *Neuropsychologia*, *65*, 169-179.

Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of experimental psychology: General*, *120*(1), 3.

Norman, G. R., & Brooks, L. R. (1997). The Non-Analytical Basis of Clinical Reasoning. In *Advances in Health Sciences Education* (Vol. 2). Kluwer Academic Publishers.

Chaffin, R., Logan, T. R., & Begosh, K. T. (2009). Performing from memory. In:

S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology*

(pp. 352–363). Oxford: Oxford University Press.

CIE (1986) *Colorimetry*, second edition: CIE publication 15.2. Vienna: Bureau Central

Crossley, M. J., & Ashby, F. G. (2015). Procedural learning during declarative control. *Journal of Experimental Psychology: Learning Memory and Cognition*, *41*(5), 1388–1403. https://doi.org/10.1037/a0038853

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

Anderson, J. R., & Fincham, J. M. (1994). Acquisition of procedural skills from examples. *Journal of experimental psychology: learning, memory, and cognition*, *20*(6), 1322.

Filoteo, J. V, Maddox, W. T., & Davis, J. (1998). Probabilistic category learning in patients with amnesia, Huntington's disease, or Parkinson's disease: The role of the hippocampus and basal ganglia. *Journal of Cognitive Neuroscience*, *10*(S), 108.

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of The National Academy of Sciences of The United States of*, *103*(31), 11778–11783. https://doi.org/10.1073/pnas.0602659103

Foerde, K., & Shohamy, D. (2011b). The role of the basal ganglia in learning and memory: Insight from Parkinson's disease. *Neurobiology of Learning and Memory*, *96*(4, SI), 624–636. https://doi.org/10.1016/j.nlm.2011.08.006

Freedberg, M., Toader, A. C., Wassermann, E. M., & Voss, J. L. (2020). Competitive and cooperative interactions between medial temporal and striatal learning systems. *Neuropsychologia*, *136*, 107257.

Greenwald, A. G., & Banaji, M. R. (2017). The implicit revolution: Reconceiving the relation between conscious and unconscious. *American Psychologist*, *72*(9), 861.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of memory and language*, *30*(5), 513-541.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science (New York, N.Y.)*, *273*(5280), 1399–1402. http://www.ncbi.nlm.nih.gov/pubmed/8703077

Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic Classification Learning in Amnesia. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *1*, 106–120. https://doi.org/10.1101/lm.1.2.106

Kruschke, J. K. (2011). Models of attentional learning. Formal approaches in categorization, 120, 120-152.

Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. Cognition, 13(3), 263-308.

Maddox, W. T., & Ashby, F. G. (2004). Dissociating explicit and procedural-learning based systems of perceptual category learning. *Behavioural Processes*, *66*(3), 309–332

Maddox, W. T., Ashby, F. G., Ing, A. D., & Pickering, A. D. (2004). Disrupting feedback processing interferes with rule-based but not information-integration category learning. *Memory & cognition*, *32*(4), 582-591.

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

Norman, G. R., & Brooks, L. R. (1997). The Non-Analytical Basis of Clinical Reasoning. In *Advances in Health Sciences Education* (Vol. 2). Kluwer Academic Publishers.

Reber, P. J. (2013). The neural basis of implicit learning and memory: A review of neuropsychological and neuroimaging research. *Neuropsychologia*, *51*(10), 2026–2042.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods.* 10.3758/s13428-018-01193-y

Poldrack, R. A., Sabb, F. W., Foerde, K., Tom, S. M., Asarnow, R. F., Bookheimer, S. Y., & Knowlton, B. J. (2005). The neural correlates of motor skill automaticity. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *25*(22), 5356–5364. https://doi.org/10.1523/JNEUROSCI.3880-04.2005

Pothos, E. M., & Wills, A. J. (Eds.). (2011). *Formal approaches in categorization.* Cambridge University Press.

Price, A., Filoteo, J. V., & Maddox, W. T. (2009). Rule-based category learning in patients with Parkinson's disease. *NEUROPSYCHOLOGIA*, *47*(5), 1213–1226

Salmon, D. P., & Butters, N. (1995). Neurobiology of skill and habit learning. *Current Opinion in Neurobiology*, *5*(2), 184–190.

Schoenlein, M. A., & Schloss, K. B. (2022). Colour-concept association formation for novel concepts. *Visual Cognition*, *30*(7), 457–479.

Squire, L. R., & Knowlton, B. J. (1995). Learning about categories in the absence of memory. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(26), 12470–12474. https://doi.org/10.1073/pnas.92.26.12470

Squire, L. R., Hamann, S., & Knowlton, B. (1994). Dissociable Learning and Memory-Systems of The Brain. *Behavioral And Brain Sciences*, *17*(3), 422–423.

Turner, B. O., Crossley, M. J., & Ashby, F. G. (2017). Hierarchical control of procedural and declarative category-learning systems. *NeuroImage*, *150*, 150–161. https://doi.org/10.1016/J.NEUROIMAGE.2017.02.039

Taylor, G. (2018) Colormath Package for Python. https://github.com/gtaylor/python-colormath

Vaquero, J. M., Lupiáñez, J., & Jiménez, L. (2020). Asymmetrical effects of control on the expression of implicit sequence learning. *Psychological Research*, *84*, 2157-2171.

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

Zeithamova, D., & Maddox, W. T. (2006). Dual-task interference in perceptual category learning. *Memory & cognition*, *34*(2), 387-398.