Prototype Abstraction Is Reduced by Concurrent Judgments of Learning

Anthony Cruz (acruz27@uwo.ca)[1] & John Paul Minda[1]
[1]Department of Psychology, University of Western Ontario

Abstract

Judgments of learning (JOLs) often modify memory but their impact on category learning remains poorly understood. In this study, we examined how concurrent JOLs affect prototype abstraction, a category learning task believed to rely on implicit memory. Participants studied low- and high-level distortions of a prototypical *Blarg* object either passively, while generating random numbers, or while making JOLs. They were then tested on their ability to identify novel *Blargs*, including the previously unpresented prototypes. Compared to the random number generators, participants who made JOLs while studying were less likely to endorse the prototypical *Blarg*. The JOL group was also less likely than the passive group to endorse foil items. Computational modeling suggested that higher learning rates, which could indicate effortful, explicit processing, reduced both prototype and foil endorsement. These findings suggest that JOLs disrupt prototype abstraction by pushing participants away from implicit learning, promoting explicit encoding strategies.

Keywords: Metacognition; Concepts and categories; Learning

Introduction

Self-assessment is essential to learning, guiding a learner's choice of how to study, what to study, and how much to study (Metcalfe, 2009). One method of self-assessment is the judgment of learning (JOL), wherein a learner predicts their likelihood of remembering material while study is ongoing. Across several word memory studies, JOLs have been associated with changes to memory retention, a phenomenon called JOL reactivity (Double et al., 2018; Ingendahl et al., 2025; Le, 2025). These changes to memory retention are most often positive. Many hypotheses suppose that JOLs change participants' study or encoding strategies (Ingendahl & Undorf, 2026; Janes et al., 2018; Lee & Ha, 2019; Mitchum et al., 2016). Though much work has been done in other areas of learning, JOL reactivity is not well-explored in category learning, the process of learning to organize stimuli into meaningful groups (Ashby & Maddox, 2011; Kéri, 2003). Successful category learning is essential to medical diagnostics (Xu et al., 2016), distinguishing poisonous and non-poisonous foods (Abel, 2023), and other important cognitive processes, making it important to understand how, if at all, JOL reactivity manifests during category learning.

Study strategies are a central theme in the study of category learning. A large body of research posits that there exist two systems of category learning: an explicit system that relies on verbalizable rules and an implicit system that relies on procedural learning (Minda et al., 2024). The explicit system supports categorization strategies that rely on specific features of a stimulus whereas the implicit system supports strategies that rely on the integration of information from multiple stimulus features. According to the COmpetition between Verbal and Implicit Systems (COVIS) model, participants begin category learning by testing explicit categorization strategies, gradually switching to implicit strategies if explicit strategies are ineffective (Ashby & Valentin, 2017). In the absence of significant implicit learning, learners may persist with using explicit categorization strategies even if they are suboptimal (Ashby et al., 2002). Dual-tasks make this persistence on suboptimal explicit strategies more likely (Zeithamova & Maddox, 2006). JOLs are sufficiently demanding that they impose dual-task effects on memory (Mitchum et al., 2016). The dual-task demands associated with JOLs may cause participants to persist with suboptimal explicit categorization strategies during learning.

In a study by Double et al. (2025), participants studied categories that could be effectively learned with either a feature-based strategy or a relational rule. It was found that JOLs impaired participants' ability to discover the relational rule. This could be seen as consistent with a persistence on explicit categorization strategies, which tend to focus on specific features of a stimulus rather than the integration of features. Double et al. argued in support of the conservatism hypothesis, a strategy-oriented account of JOL reactivity which supposes that JOLs increase participants' response thresholds, requiring them to be more confident in their category judgment before responding (Li et al., 2024). This strategy shift likely pushes participants to prioritize immediate success over long-term mastery, making participants respond more conservatively (Double, Goldwater, et al., 2025; Double, Tran, et al., 2025). Existing results in category learning are consistent with the idea JOLs affect a learner's categorization strategies.

In three experiments, Lee and Ha (2019) had participants study one set of categories, then make JOLs, then study another set of categories before taking an abstraction test on all studied categories. In the first experiment, participants made item-level JOLs; they were shown each stimulus-label pair and asked to rate their likelihood of correctly identifying new members of that category in a later test. In the second study, participants made category-level JOLs; they were shown each category label on its own and asked to predict

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

their likelihood of correctly identifying new category members in a later test. In the final experiment, participants made global-level JOLs, where they provided high-level ratings about the monitoring, evaluating, and planning of their learning. Category- and global-level JOLs potentiated the abstraction of to-be-learned categories while item-level JOLs had no effect on such categories. None of the JOLs significantly affected abstraction of the previously studied categories. Another study similarly found no effect of item-level JOLs on abstraction (Ha & Lee, 2024). They suggested that the item-level JOLs encouraged learners to reflect on individual item features instead of the similarities between items in a shared category, which would be highlighted by the broader-level JOLs and would be more essential to successful abstraction. In other words, the authors supposed that appropriately-leveled JOLs guided participants toward more successful encoding strategies in the future. This finding is in line with literature on word memory that highlights how JOLs may impact a participant's use of item-specific vs. relational processing during study (Chang & Brainerd, 2024; Senkova & Otani, 2021). Item-level JOLs may affect category learning by directing participants away from the shared features of category members, instead focusing on features that distinguish them from other category members.

While many JOL reactivity studies in word memory ask participants to make JOLs while stimuli are still being presented (concurrent JOLs; e.g., Mitchum et al., 2016; Myers et al., 2020; Rivers et al., 2023; Soderstrom et al., 2015), existing studies of JOL reactivity in category learning have thus far used non-concurrent JOLs. Many category learning paradigms, especially feedback-based paradigms, are not easily compatible with concurrent JOLs. One task that is compatible with concurrent JOLs is the prototype abstraction task, a classical category learning paradigm in which participants are exposed to distortions of an unpresented prototype and then tested on their ability to identify new category members during a later test (Smith & Minda, 2002). Despite never seeing the prototype during training, participants tend to endorse it more often at test than other stimuli, a phenomenon referred to as prototype enhancement. Prototype abstraction is believed to rely on implicit memory (Squire & Knowlton, 1995). Unrelated dual-tasks are not believed to impact implicit category learning (Waldron & Ashby, 2001). The prototype abstraction task is an excellent way to study JOL reactivity in category learning as it allows the exploration of the relevant strategy-oriented and processing-oriented accounts of JOL reactivity.

The purpose of this experiment was to determine the effect, if any, of concurrent JOLs on prototype abstraction. Although JOL reactivity has been examined in category learning paradigms (Double, Tran, et al., 2025; Ha & Lee, 2024; Lee & Ha, 2019), we believe this to be the first such study using concurrent JOLs, which are directly comparable to those used in existing word memory research (Mitchum et al., 2016; Soderstrom et al., 2015). The dual-task demands of JOLs may cause participants to use suboptimal explicit categorization strategies (Ashby et al., 2002). As prototype abstraction relies on implicit memory, this should result in lower prototype endorsement following JOLs. Alternatively, the dual-task demands of JOLS may leave prototype abstraction intact (Waldron & Ashby, 2001). A task-irrelevant dual-task condition was included in this experiment to dissociate these possibilities. We also explored the conservatism hypothesis, which predicts that JOLs would reduce the likelihood of a stimulus being endorsed as a category member (e.g., induce a "no" response bias). We also explored the item-specific processing hypothesis put forth by Lee & Ha. If this hypothesis holds, then JOLs should emphasize features that make stimuli stand out from each other. We included a mix of novel and studied items at test to determine if JOLs differentially affected endorsement of those items. Our results are supplemented by computational modeling.

Methods

Data were collected online between 18 November 2025 and 27 January 2026. Participants were randomly assigned to one of three study conditions: Passive, random number generation (RNG), or judgment of learning (JOL). They then studied novel objects called *Blargs*. In the RNG and JOL conditions, they completed concurrent tasks while studying the imaginary objects. At test, they were tested on their ability to identify whether novel objects were *Blargs*. Experimental procedures and materials were approved by the University of Western Ontario's Research Ethics Board. The experiment was programmed in jsPsych version 7.1.2 (de Leeuw, 2015).

Participants

Participants were N = 142 ($n_{Passive} = 42$, $n_{RNG} = 54$, $n_{JOL} = 46$) undergraduate students compensated with course credit. Participants all spoke fluent English and had normal or corrected-to-normal vision.

Materials

Stimuli were generated in R version 4.5.1 (R Core Team, 2021). To generate the prototypical *Blarg*, nine *(x,y)* ordered pairs were generated, $X, Y \in U(0,30)$. These ordered pairs were plotted as points in two-dimensional space, and points were connected to each other at random such that the degree of each point was two, generating a closed shape. See example stimuli in Figure 1.

Figure 1: From left to right, the prototypical *Blarg*, a low-level distortion, and a high-level distortion.

To generate distortions, a low or high amount of Gaussian noise was applied to each coordinate. The total amount of noise applied to each low- (LLD) or high-level distortion (HLD) was consistent. This method of prototype generation was validated in a pilot study. The foil items presented at test were low- and high-level distortions of a different prototype (e.g., an unstudied category).

Procedure

Eligible participants signed up for this study on the University of Western Ontario SONA system. The procedure involved a study phase, comprised of two learning blocks, and a test phase wherein participants' category knowledge was assessed. The design is largely modeled after Smith and Minda's (2002) Experiment 2.

Study Phase. The study phase comprised two learning blocks of 40 trials each, resulting in 80 study trials altogether. During one study trial, participants saw a fixation cross for 500-800 milliseconds and were then presented with a high-level or low-level *Blarg* distortion with the text "This is a Blarg." written above it. Each *Blarg* distortion was presented for exactly 8 seconds. During each learning block, 20 high- and 20 low-level distortions were shown in a random order. The same stimuli were shown in each learning block. Participants were able to take a self-paced break between each block. After the second learning block, participants played Tetris during the five-minute retention interval.

The experimental manipulation took place in the study phase. In the Passive learning condition, participants passively viewed each stimulus as previously described. In the Random Number Generation (RNG) condition, participants were prompted halfway through each stimulus presentation to type a random number between 0 and 100. In the Judgment of Learning (JOL) condition, participants were prompted halfway through each stimulus presentation to type their likelihood of remembering this *Blarg* later from 0 (*Sure I won't remember*) to 100 (*Sure I will remember*). The RNG condition serves as a control condition for the potential dual-task demands in the JOL condition.

Test Phase. During one test trial, participants were asked "Is this a Blarg?" and were given 5 seconds to choose *Yes* [1] or *No* [0] using their keyboard. The test phase had 84 trials: 4 prototypical *Blarg* trials, 20 low-level *Blarg* distortions (LLD), 20 high-level *Blarg* distortions (HLD), and 40 foil items. Among LLD and HLD items, half were previously studied and half were novel.

Computational Modeling

The task was simulated using MINERVA 2, which assumes that learners maintain a memory trace for each experience (Hintzman, 1984). Each memory trace is represented as a feature vector. Each simulation has a learning rate, *L*, which determines the likelihood of each stimulus feature being encoded. If $L < 1$, two identical experiences might result in distinct memory traces. In this model, a probe activates all memory traces in parallel, with activation being a function of the probe's similarity to a given memory trace. The model returns an echo which includes content, the information that was activated in memory, and intensity, the total strength of memory trace activation. In this study, we focus on intensity. This model allows us to explore how different learning rates may impact prototype abstraction performance.

In this study, features were defined as normalized (x, y) coordinates. We computed the distance between a probe P and a given memory trace t using the L_2 distance metric, described in equation (1). I is the set of points in the given dot pattern, and x and y are normalized coordinates of each point. E.g., t_{x_1} is the x-value of point 1 in the present memory trace.

$$d(P,t) = \frac{\sum_I \sqrt{\sum_{z\in\{x,y\}}\left(P_{z_i} - t_{z_i}\right)^2}}{|I|} \tag{1}$$

The activation (A) of a memory trace t is a monotonically increasing function of its similarity to the probe P, as described in equation (2). The chosen similarity function S is similar to that used in other categorization models (Nosofsky, 1986; Smith & Minda, 1998).

$$A(P,t) = S(P,t)^3 = \left(e^{-d(P,t)}\right)^3 \tag{2}$$

In this model, the intensity of the echo returned by the probe is the sum of the activation of all memory traces. One thousand simulations were run for each learning rate $L \in \{.05, .10, .15, \ldots, .90, .95\}$. Higher echo intensity corresponds to higher probability of endorsing a probe as a *Blarg*.

Results

Behavioural Results

A 2 x 2 (Distortion Level x Learning Block) within-participants Analysis of Variance (ANOVA) was conducted within the JOL condition with mean JOL (M = 33.09, SD = 24.43) as the dependent variable. All main effects and interactions were not significant, each p > 0.1. There is no evidence that JOLs varied over the course of training or with the distortion level of the active stimulus.

A 3 x 4 (Learning Condition x Item Type) Mixed ANOVA was conducted with the probability of an object being classified as a *Blarg* as the dependent variable. The main effect of Item Type was significant, $F(3,420) = 350$, $p = 5.3 \times 10^{-113}$, $\eta_p^2 = 0.71$. Consistent with the typicality effect typically seen in prototype abstraction tasks, participants were most likely to endorse the prototype (M = 0.87, SD = 0.21), followed by low-level distortions (LLD; M = 0.8, SD = 0.18), then high-level distortions (HLD; M = 0.66, SD = 0.2). Foils were endorsed least often (M = 0.22, SD = 0.23). Each pairwise comparison among test item types was significant, each $p < 9.65 \times 10^{-5}$. The main effect of Learning Condition ($F(2,140) = 2.5$, $p = 0.084$, $\eta_p^2 = 0.035$) and the interaction effect ($F(6,420) = 1.3$, $p = 0.25$, $\eta_p^2 = 0.019$) were not significant. Figure 2 shows the *Blarg* classification probabilities across item types and learning conditions.

Planned pairwise comparisons were conducted comparing prototype and foil classification probabilities among learning conditions. The RNG group (M = 0.9, SD = 0.17) was marginally more likely than the JOL group (M = 0.82, SD = 0.25) to endorse the prototype, p = 0.027, d = 0.43. The Passive group (M = 0.29, SD = 0.27) was more likely than the JOL group (M = 0.16, SD = 0.18) to endorse foil items, p = 0.01, d = 0.51.

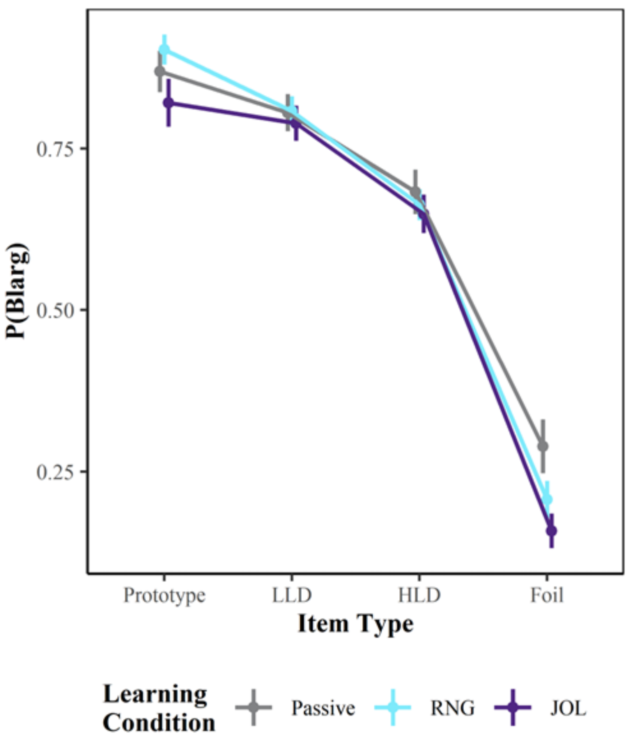

Figure 2: Mean (±SE) probability of identifying stimulus as a *Blarg* across test item types and learning conditions.

A generalized linear mixed-effects model with a logit link function was conducted to examine the effect of learning condition on prototype endorsement likelihood. The RNG condition was the reference for dummy coding. Relative to the RNG condition, the JOL condition was associated with a lower likelihood of correctly endorsing the prototypes, $b = -0.99$, $SE = 0.43$, $z = -2.29$, $p = 0.022$. There was no significant difference in prototype endorsement between the RNG and Passive conditions, $b = -0.46$, $SE = 0.47$, $z = -1.0$, $p = 0.32$. Random intercepts were used for participants to account for between-participant variability (SD = 1.29).

A 3 x 2 x 2 (Learning Condition x Distortion Level x Novelty) mixed ANOVA was conducted with the probability of a prototype distortion being correctly identified as a *Blarg* as the dependent variable. The "Novelty" factor refers to whether the test item is one that was studied during the study phase; half of the LLD and HLD items classified at test were studied during the study phase. Foil and Prototype items could not have been previously studied, so they were dropped from this analysis. Much like the previous analysis, the main effect of distortion level was significant, $F(1,140) = 99$, $p = 4.5 \times 10^{-18}$, $\eta_p^2 = 0.41$. Participants were better at identifying low-distortion items (M = 0.8, SD = 0.20) than high-distortion items (M = 0.66, SD = 0.22). All other main effects and interactions were not significant, each p > 0.2. Participants do not have a significant advantage when classifying previously studied *Blargs* compared to novel *Blargs*.

A one-way ANOVA predicting response bias as a function of learning condition was significant, $F(2,140) = 3.8$, $p = 0.024$, $\eta_p^2 = 0.052$. Pairwise comparisons revealed that the JOL group (M = 0.46, SD = 0.1) was less likely to endorse an item than the passive group (M = 0.53, SD = 0.17), p = 0.01. Other pairwise comparisons were not significant, p > 0.1. See Figure 3.

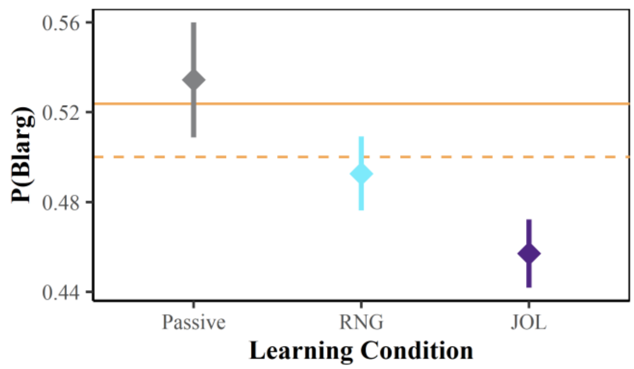

Figure 3: Response bias (Mean ± SE) as a function of learning condition. Solid horizontal line: .52, actual proportion of *Blargs* in test set. Dashed line: .5.

Computational Results

In each simulation, echo intensity was highest for the prototype, followed by the low-level distortions, then the high-level distortions, then the foils. Thus, the prototype enhancement effect was present despite there being no explicit prototype representation in the MINERVA 2 model.

In the simulations, echo intensity for the prototype, low-level distortions, and high-level distortions peaks at L = 0.25. At higher values of L, echo intensity gradually decreases. Echo intensity for the foil items peaks earlier at L = 0.20, seeing a more noticeable decrease at higher values of L. In summary, both prototype and foil activation decrease at higher learning rates. See Figure 4 for a visualization of these model results.

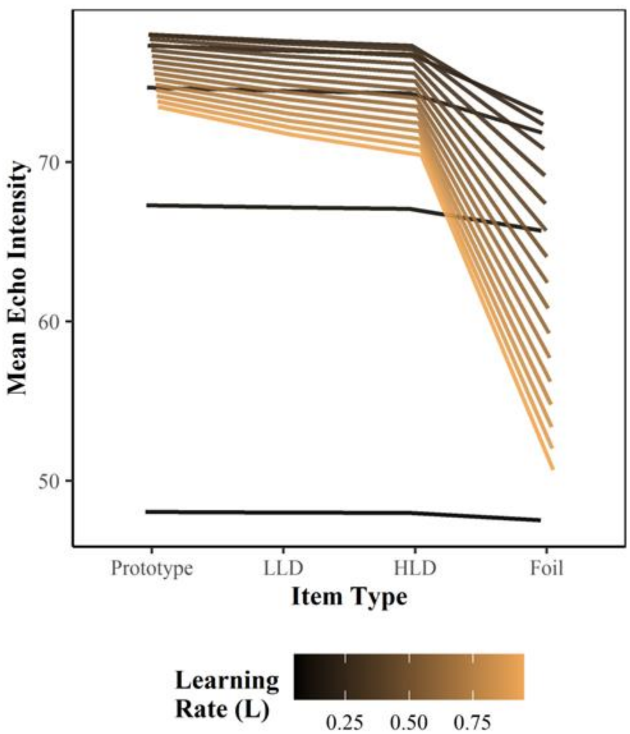

Figure 4: Mean echo intensities for each probe item type in the simulations at each learning rate, $0.05 \leq L \leq 0.95$.

Discussion

The JOL group showed lower prototype endorsement than the RNG group. At a glance, this result may seem counterintuitive; the group performing a learning-relevant dual-task achieved lower prototype abstraction than the group performing a learning-irrelevant dual-task. However, prototype abstraction relies on implicit memory (Squire & Knowlton, 1995; Zannino et al., 2012), and dual-tasks can cause participants to use explicit categorization strategies even when an implicit strategy is optimal (Zeithamova & Maddox, 2006). In particular, JOLs may cause participants to more explicitly engage with the learning task, which is counterproductive in this task where implicit engagement is essential. Our results are consistent with an account of JOL reactivity in category learning wherein JOLs lead to persistent use of explicit categorization strategies. Compared to the Passive group, the JOL group had a stronger bias against endorsing stimuli as category members. This is consistent with the conservatism hypothesis endorsed by Double at al. (2025), wherein JOLs cause participants to require more evidence to endorse an item as a category member. Broadly, our results build upon prior research which suggests that JOLs can change participants' study or encoding strategies (Chang & Brainerd, 2024; Ingendahl & Undorf, 2026; Janes et al., 2018; Lee & Ha, 2019; Mitchum et al., 2016).

If JOLs induced item-specific processing, as suggesting by Lee & Ha (2024; 2019), then JOLs should be associated with an advantage at classifying previously studied items, as they should stand out from the novel items at test. At test, the learning groups did not differ in their abilities to classify novel items compared to previously studied items. This is inconsistent with the idea that JOLs support item-specific processing in category learning. However, our stimuli are low-level shape stimuli whereas Lee & Ha trained participants on complex painting stimuli. The complexity of the stimulus set may explain differences in learning strategy, broadly, as well as how JOLs impact them. It should be noted that there are also mixed results related to item-specific vs. relational processing in word memory (Chang & Brainerd, 2024; Senkova & Otani, 2021).

In JOL experiments, participants tend to rate easier study items as more memorable (e.g., Mueller et al., 2013; Soderstrom et al., 2015). In this experiment, low-level distortions were classified more accurately at test than high-level distortions. Despite this behavioural evidence that the low-level distortions were easier to classify than the high-level distortions, JOLs did not vary over the course of study or with distortion level. Participants metacognitive predictions did not reflect the difficulty of the studied items. As participants had enough category knowledge to successfully perform the categorization task but could not make accurate memory predictions, it is possible that a prototype was not represented explicitly in memory. One possibility is that the prototype is represented in implicit memory, unable to be accessed during an explicit JOL task (Squire & Knowlton, 1995; Zannino et al., 2012). An alternative interpretation is that the prototype is simply not stored in memory. Instance models, like MINERVA 2 (Hintzman, 1984), show prototype enhancement effects even when prototypes are not explicitly represented in the models (Hintzman, 1984, 1986; Palmeri & Nosofsky, 2001).

In the MINERVA 2 simulations, if mean echo intensity is assumed to have a monotonic relationship with *Blarg* classification probability, then the highest *Blarg* endorsement rates would be seen at a low learning rate, L = 0.25. A higher or lower learning rate would decrease *Blarg* endorsement likelihood. The lower learning rates have shallower typicality gradients, so a higher learning rate seems a more appropriate comparison for our behavioural data. This comparison also makes theoretical sense. If JOLs cause participants to persist with explicit categorization strategies, then participants would be paying more attention to specific stimulus features, perhaps remembering them at a higher rate. Some accounts of category learning consider forgetting to be essential to abstraction (Vlach, 2014; Vlach et al., 2014). When exemplars are encoded as memory traces, these accounts suggest that forgetting the features that differentiate items within a category improves abstraction, as within-category differences are de-emphasized. Perhaps JOLs increase participants' memory for specific stimulus features, lowering their ability to abstract to novel instances of a category. Future work will explore whether JOLs cause participants to attend to stimulus features that differentiate items within a category.

It is worth noting that many well-supported accounts of JOL reactivity rely on the idea that JOLs direct participants to attend to semantic information, a helpful cue which naturally aids memory (Chang & Brainerd, 2024, 2023; Maxwell & Huff, 2023; Myers et al., 2020; Rivers et al., 2021; Soderstrom et al., 2015). These accounts are not immediately relevant to category learning as, by definition, novel categories are not associated with semantic information. These accounts might be worth considering in a paradigm wherein participants are trained extensively on novel items (e.g., Gauthier et al., 1998). These accounts are not entirely inconsistent with strategy-oriented accounts.

With only 4 instances of the prototype shown at test, our measure of prototype abstraction is quite coarse, resulting in floor and ceiling effects. Adding more instances of the prototype would not remedy this issue, as participants continue learning through the duration of a categorization test (Palmeri & Flanery, 1999; Palmeri & Mack, 2015) and extensive repetitions of a target in a memory test results in substantial decreases in error (Fischler & Juola, 1971). Future work should examine the abstraction of multiple prototypes at once to get a more granular view of group differences, should they exist.

Category learning is qualitatively distinct from the memory tasks in which JOL reactivity is typically studied. Although we have made attempts, it should be acknowledged that direct connections to the broader JOL reactivity literature are not easily made. Our interpretations should not be taken as outright endorsements or rejections of existing hypotheses, as the mechanisms proposed in these hypotheses may be different in distinct learning tasks. Still, we believe that this experiment builds upon existing theories related to JOL reactivity in a way that warrants further investigation. Our results demonstrate, counterintuitively, that actively reflecting on one's learning can be detrimental to later abstraction. Given the importance of abstraction in learning, JOL reactivity in category learning warrants more extensive investigation.

Conclusion

Our results are consistent with accounts of JOL reactivity in category learning which suggest that JOLs affect participants' categorization strategies, either by promoting explicit strategy use or by promoting more conservative strategies. These results provide preliminary evidence that concurrent JOLs, often helpful in word memory retention, can negatively impact abstraction, a fundamental learning process. Future work should more deeply explore the mechanisms by which JOLs impact abstraction and identify situations wherein it does and does not.

References

Abel, R. (2023). Some fungi are not edible more than once: The impact of motivation to avoid confusion on learners' study sequence choices. *Journal of Applied Research in Memory and Cognition*. https://doi.org/10.1037/mac0000107

Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy of Sciences*, *1224*, 147–161. https://doi.org/10.1111/j.1749-6632.2010.05874.x

Ashby, F. G., Maddox, W. T., & Bohil, C. J. (2002). Observational versus feedback training in rule-based and information-integration category learning. *Memory & Cognition*, *30*(5), 666–677. https://doi.org/10.3758/bf03196423

Ashby, F. G., & Valentin, V. V. (2017). Chapter 7 - Multiple systems of perceptual category learning: Theory and cognitive tests. In H. Cohen & C. Lefebvre (Eds.), *Handbook of Categorization in Cognitive Science* (second, pp. 157–188). Elsevier. https://doi.org/10.1016/B978-0-08-101107-2.00007-5

Chang, M., & Brainerd, C. (2024). Judgments of Learning Reactivity on Item-Specific and Relational Processing. *Journal of Intelligence*, *12*(1). https://doi.org/10.3390/jintelligence12010004

Chang, M., & Brainerd, C. J. (2023). Changed-goal or cue-strengthening? Examining the reactivity of judgments of learning with the dual-retrieval model. *Metacognition and Learning*, *18*(1), 183–217. https://doi.org/10.1007/s11409-022-09321-y

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Double, K. S., Birney, D. P., & Walker, S. A. (2018). A meta-analysis and systematic review of reactivity to judgements of learning. *Memory*, *26*(6), 741–750. https://doi.org/10.1080/09658211.2017.1404111

Double, K. S., Goldwater, M. B., & Birney, D. P. (2025). Reactivity to confidence ratings: evidence of impaired rule-learning. *Metacognition and Learning*, *20*(1), 8. https://doi.org/10.1007/s11409-025-09413-5

Double, K. S., Tran, D., & Goldwater, M. B. (2025). Judgments of learning impair rule-based discovery. *Memory & Cognition*. https://doi.org/10.3758/s13421-025-01737-6

Fischler, I., & Juola, J. F. (1971). Effects of repeated tests on recognition time for information in long-term memory. *Journal of Experimental Psychology*, *91*(1), 54–58. https://doi.org/10.1037/h0031838

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training 'greeble' experts: a framework for studying expert object recognition processes. *Vision Research*, *38*(15–16), 2401–2428. https://doi.org/10.1016/s0042-6989(97)00442-2

Ha, H., & Lee, H. S. (2024). Evaluating the judgment of learning: Its limited impact and the power of retrieval on inductive learning. *Educational Psychology Review*, *36*(2), 1–28. https://doi.org/10.1007/s10648-024-09899-6

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods,*

Instruments, & Computers, *16*(2), 96–101. https://doi.org/10.3758/BF03202365
Hintzman, D. L. (1986). “Schema abstraction” in a multiple-trace memory model. *Psychological Review*, *93*(4), 411–428. https://doi.org/10.1037/0033-295x.93.4.411
Ingendahl, F., Halamish, V., & Undorf, M. (2025). Do immediate judgments of learning alter memory performance? A meta-analytical review. *Psychological Bulletin*, *151*(7), 892–929. https://doi.org/10.1037/bul0000487
Ingendahl, F., & Undorf, M. (2026). Instructed learning strategy use eliminates negative reactivity of immediate judgments of learning. *Psychonomic Bulletin & Review*, *33*(1), 47. https://doi.org/10.3758/s13423-025-02844-w
Janes, J. L., Rivers, M. L., & Dunlosky, J. (2018). The influence of making judgments of learning on memory performance: Positive, negative, or both? *Psychonomic Bulletin & Review*, *25*(6), 2356–2364. https://doi.org/10.3758/s13423-018-1463-4
Kéri, S. (2003). The cognitive neuroscience of category learning. *Brain Research. Brain Research Reviews*, *43*(1), 85–109. https://doi.org/10.1016/s0165-0173(03)00204-2
Le, N. D. (2025). Is metamemory judgements reactivity a concern or a topic worth investigating? An overview of literature and future directions. *Roczniki Psychologiczne*, *27*(Online First), 299. https://doi.org/10.18290/rpsych2024.0017
Lee, H. S., & Ha, H. (2019). Metacognitive judgments of prior material facilitate the learning of new material: The forward effect of metacognitive judgments in inductive learning. *Journal of Educational Psychology*, *111*(7), 1189–1201. https://doi.org/10.1037/edu0000339
Li, B., Hu, X., Shanks, D. R., Su, N., Zhao, W., Meng, L., Lei, W., Luo, L., & Yang, C. (2024). Confidence ratings increase response thresholds in decision making. *Psychonomic Bulletin & Review*, *31*(3), 1093–1102. https://doi.org/10.3758/s13423-023-02380-5
Maxwell, N. P., & Huff, M. J. (2023). Is discriminability a requirement for reactivity? Comparing the effects of mixed vs. pure list presentations on judgment of learning reactivity. *Memory & Cognition*, *51*(5), 1198–1213. https://doi.org/10.3758/s13421-022-01381-4
Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science*, *18*(3), 159–163. https://doi.org/10.1111/j.1467-8721.2009.01628.x
Minda, J. P., Roark, C. L., Kalra, P., & Cruz, A. (2024). Single and multiple systems in categorization and category learning. *Nature Reviews Psychology*, 1–16. https://doi.org/10.1038/s44159-024-00336-7
Mitchum, A. L., Kelley, C. M., & Fox, M. C. (2016). When asking the question changes the ultimate answer: Metamemory judgments change memory. *Journal of Experimental Psychology. General*, *145*(2), 200–219. https://doi.org/10.1037/a0039923
Mueller, M. L., Tauber, S. K., & Dunlosky, J. (2013). Contributions of beliefs and processing fluency to the effect of relatedness on judgments of learning. *Psychonomic Bulletin & Review*, *20*(2), 378–384. https://doi.org/10.3758/s13423-012-0343-6
Myers, S. J., Rhodes, M. G., & Hausman, H. E. (2020). Judgments of learning (JOLs) selectively improve memory depending on the type of test. *Memory & Cognition*, *48*(5), 745–758. https://doi.org/10.3758/s13421-020-01025-5
Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology. General*, *115*(1), 39–61. https://doi.org/10.1037//0096-3445.115.1.39
Palmeri, T. J., & Flanery, M. A. (1999). Learning about categories in the absence of training: Profound amnesia and the relationship between perceptual categorization and recognition memory. *Psychological Science*, *10*(6), 526–530. https://doi.org/10.1111/1467-9280.00200
Palmeri, T. J., & Mack, M. L. (2015). How experimental trial context affects perceptual categorization. *Frontiers in Psychology*, *6*, 180. https://doi.org/10.3389/fpsyg.2015.00180
Palmeri, T. J., & Nosofsky, R. M. (2001). Central tendencies, extreme points, and prototype enhancement effects in ill-defined perceptual categorization. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *54*(1), 197–235. https://doi.org/10.1080/02724980042000084
R Core Team. (2021). *R: A Language and Environment for Statistical Computing* (4.1.1). R Foundation for Statistical Computing. https://www.R-project.org/
Rivers, M. L., Janes, J. L., & Dunlosky, J. (2021). Investigating memory reactivity with a within-participant manipulation of judgments of learning: support for the cue-strengthening hypothesis. *Memory*, *29*(10), 1342–1353. https://doi.org/10.1080/09658211.2021.1985143
Rivers, M. L., Janes, J. L., Dunlosky, J., Witherby, A. E., & Tauber, S. K. (2023). Exploring the Role of Attentional Reorienting in the Reactive Effects of Judgments of Learning on Memory Performance. *Journal of Intelligence*, *11*(8). https://doi.org/10.3390/jintelligence11080164
Senkova, O., & Otani, H. (2021). Making judgments of learning enhances memory by inducing item-specific processing. *Memory & Cognition*, *49*(5), 955–967. https://doi.org/10.3758/s13421-020-01133-2
Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *24*(6), 1411–1436. https://doi.org/10.1037/0278-7393.24.6.1411
Smith, J. D., & Minda, J. P. (2002). Distinguishing prototype-based and exemplar-based processes in dot-pattern category learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *28*(4), 800–811. https://doi.org/10.1037/0278-7393.28.4.800
Soderstrom, N. C., Clark, C. T., Halamish, V., & Bjork, E. L. (2015). Judgments of learning as memory modifiers. *Journal of Experimental Psychology. Learning, Memory,*

and Cognition, *41*(2), 553–558. https://doi.org/10.1037/a0038388

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

Vlach, H. A. (2014). The spacing effect in children's generalization of knowledge: Allowing children time to forget promotes their ability to learn. *Child Development Perspectives*, *8*(3), 163–168. https://doi.org/10.1111/cdep.12079

Vlach, H. A., Sandhofer, C. M., & Bjork, R. A. (2014). Equal spacing and expanding schedules in children's categorization and generalization. *Journal of Experimental Child Psychology*, *123*, 129–137. https://doi.org/10.1016/j.jecp.2014.01.004

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. https://doi.org/10.3758/bf03196154

Xu, B., Rourke, L., Robinson, J. K., & Tanaka, J. W. (2016). Training melanoma detection in photographs Using the perceptual expertise training approach. *Applied Cognitive Psychology*, *30*(5), 750–756. https://doi.org/10.1002/acp.3250

Zannino, G. D., Perri, R., Zabberoni, S., Caltagirone, C., Marra, C., & Carlesimo, G. A. (2012). Performing prototype distortion tasks requires no contribution from the explicit memory systems: evidence from amnesic MCI patients in a new experimental paradigm. *Neuropsychologia*, *50*(12), 2907–2915. https://doi.org/10.1016/j.neuropsychologia.2012.08.024

Zeithamova, D., & Maddox, W. T. (2006). Dual-task interference in perceptual category learning. *Memory & Cognition*, *34*(2), 387–398. https://doi.org/10.3758/bf03193416