Humans Know More Than Exemplar Models Do

Joshua Glass (jglass4@binghamton.edu) & Kenneth Kurtz (kkurtz@binghamton.edu)
Department of Psychology, Binghamton University
Binghamton, NY 13905 USA

Abstract

Exemplar models are the canonical account of human category learning. We investigate the sufficiency of this framework across two experiments demonstrating that human learners spontaneously extract information that exemplar models fail to capture. In Experiment 1, we show that participants classify novel test items based on abstract domain-level regularities rather than summed similarity to stored exemplars. In Experiment 2, we test the design principle of selective attention by providing a perfect unidimensional cue while making a secondary dimension partially diagnostic. Contrary to the attention weight optimization of exemplar theory, participants included the redundant 'extra' features in their category representations as evidenced by single feature classification and generalization performance. These findings constitute an important challenge to core components of exemplar theory and suggest a generative process where learners build rich statistical models of the environment.

Keywords: Categorization; Concepts; Exemplar Models; GCM

Introduction

For decades, the theoretical landscape of category learning has been centered around the broadly excellent fits of exemplar models to human data. Originally formalized by Medin and Schaffer (1978) and extended by Nosofsky (1984, 1986), the exemplar account posits that humans classify novel stimuli by computing attention-mediated similarity to specific, previously stored instances of a category.

Exemplar models have consistently provided superior quantitative fits to benchmark datasets compared to prototype or rule-based alternatives (e.g., Kruschke, 1992, 1993; Nosofsky, Gluck, Palmeri, McKinley, & Glauthier, 1994). The Context Theory (Medin & Schaffer, 1978) accurately predicted the difficulty of the "5-4" category structure; a finding that challenged earlier prototype views. Further, the framework has demonstrated considerable flexibility—offering compelling accounts of complex cognitive phenomena such as shift learning difficulty (Kruschke, 1996), the inverse base-rate effect (Kruschke, 2001; 2003), and typicality effects (Nosofsky, 1988).

Despite this explanatory power, recent research (Conaway & Kurtz, 2017; Kurtz &Wetzel, 2021)has begun to identify divergences between human behavior and exemplar model predictions. We focus here particularly on sensitivity to within-category statistical structure. A strictly exemplar-based view assumes that classification is driven by summed similarity to stored items. However, Conaway (2016) demonstrated that categories possessing strong within-category feature correlations are learned significantly more easily than matched controls. Critically, because the between-category and within-category similarities were equated across conditions, exemplar models predicted identical performance for both, failing to account for the human sensitivity to the internal correlational structure.

This divergence extends beyond initial learning and into generalization. Conaway and Kurtz (2017) presented subjects with novel items that maintained learned feature correlations but were not necessarily proximal to specific training exemplars. Learners tended to classify these items based on whether they conformed to the extracted correlations rather than their similarity to stored instances. These findings suggest that learners are not merely storing instances, but they are sensitive to abstract regularities in ways the standard exemplar view does not capture.

In the current work, we present a set of experiments that pose similar, critical challenges to the exemplar account. In Experiment 1, we demonstrate that humans are sensitive to abstract, domain-level regularities in the category structure—global statistical features that exemplar models are effectively blind to. In Experiment 2, we show that humans learn significantly more about category items than is necessary to solve the classification task. This contradicts the nature of attentionally-mediated exemplar storage, which implies learners only encode what is required to distinguish categories.

Experiment 1

Methods

Participants Participants were undergraduate students recruited to complete the experiment in partial fulfillment of a course requirement. To ensure that analyses focused on learners who had adequately acquired the base category structures, an accuracy criterion was applied: only participants who achieved greater than 80% accuracy by the final block of training were retained for analysis. After applying this exclusion criterion, the final sample consisted of 79 participants distributed across three conditions: Full Circle, 26 participants (from an initial pool of 98); Curtailed Circle, 29 participants (from an initial pool of 61); Top-Middle-Bottom, 24 participants (from an initial pool of 50).

Stimuli and Design The stimuli consisted of squares varying along two dimensions: shading and size (Figures 1-5 represent the locations of stimuli plotted in a coordinate plane). An independent scaling study ensured that both dimensions were equally salient and separable, making them appropriate for the city-block distance metric. Participants were assigned to one of three experimental conditions which

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

were each defined by a unique category structure designed to dissociate exemplar-based similarity from domain-level regularities.

Full Circle Condition: In this condition, the categories were arranged as two concentric circles in the two-dimensional feature space (see Figure 1; roughly inspired by the structure used by Qiu & Minda, 2023). The inner circle constituted the "Beta" category, while the outer circle constituted the "Alpha" category. This structure creates a distinct abstract regularity: Alpha items are always extreme on at least one dimension, whereas Beta items are moderate on both. We introduced novel test items at coordinates including (0.1, 0.3) and (0.9, 0.7). While these points are equidistant from the nearest training exemplars of both categories, standard exemplar models (which compute summed similarity across all items) predict a slight bias toward the Beta category. However, if participants learn the domain-level regularity (extreme vs. moderate), they should classify these items as Alphas (for full distance calculations and model simulations see the supplemental GitHub repository[1]).

Curtailed Circle Condition: This structure was a reduced version of the Full Circle design (see Figure 2). The extreme versus moderate global regularity was removed, and a different probabilistic regularity was introduced: the Alpha category was predominantly *high* on the shading dimension (and sometimes *medium*), while the Beta category was predominantly *medium* on the shading dimension (and sometimes *low*). We introduced a test item at (0.9, 0.7) to adjudicate between strategies. Under standard exemplar models, the summed similarity across all training items predicts that classification should be near chance (50%). However, because this item conforms to the extracted 'Alpha is high-shading' regularity, the domain-level account predicts robust classification into the Alpha category. Thus, we can test for the presence of domain-level learning by evaluating whether Alpha responding on this item significantly exceeds the chance baseline.

Top-Middle-Bottom Condition: This structure was designed to implement a similar type of abstract regularity as the Curtailed Circle (high vs. low dimension bias – though it is important to note that there is no clear rule-like boundary; see Figure 3). Like the other conditions, this condition utilized test items where similarity and regularity made opposing predictions. The test item at (0.1, 0.7) has greater summed similarity to the Beta category but conforms to the Alpha regularity (high feature value). Conversely, the item at (0.9, 0.3) has greater summed similarity to the Alpha category but conforms to the Beta regularity (low feature value).

Procedure Participants were instructed that their task was to learn to classify items into "Alpha" or "Beta" categories, and that they would later be tested on unfamiliar items. The experiment began with a training phase consisting of 15 blocks. During each block, training items were presented in a randomized order and participants received immediate corrective feedback after each classification decision.

Following training, participants completed a single test block. This block included the original training items intermixed with the novel critical and reference test items. No feedback was provided during this phase. To obtain a more stable estimate of the participants' true category representations and mitigate random response errors, each unique test item was presented five times; the order of presentation was fully randomized.

Results

Full Circle Results The primary question in this condition was whether participants would generalize the 'extreme versus moderate' domain-level regularity to the critical test items.

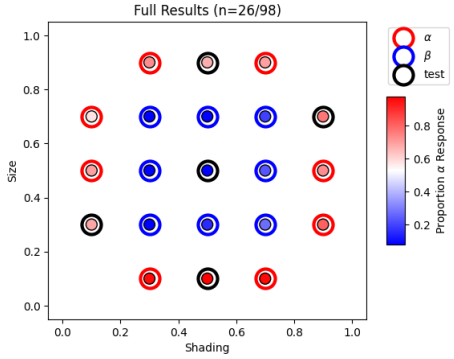

Figure 1: Stimulus structure and generalization performance- Full Circle condition. Colored outlines denote true labels. Interior color- mean 'Alpha' responses at test.

To test whether participants generalized based on the domain-level regularity rather than exemplar similarity, we compared Alpha responding on the test items against a conservative chance baseline (0.50). Participants demonstrated a high proportion of Alpha responding (M = 0.73, SE = 0.04). A one-sample t-test confirmed that this performance was significantly above chance ($t(25) = 5.3$, $p < 0.01$, $d = 1.03$)..

Curtailed Circle Results A similar pattern of results emerged in the Curtailed Circle condition. As shown in Figure 2, participants displayed dominant Alpha generalization to the critical test item ($\boldsymbol{M} = 0.86$, $\boldsymbol{SE} = 0.05$), classifying it as belonging to the Alpha category despite it being equidistant from the training exemplars of both categories.

[1] https://github.com/Josh-Glass/Humans-Know-More-Than-Exemplar-Models

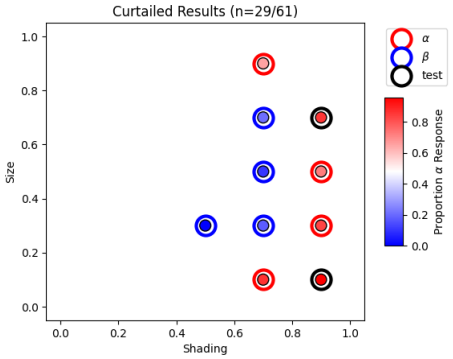

Figure 2: Stimulus structure and generalization performance- Curtailed Circle condition.

A one-sample t-test confirmed that Alpha responding to the test item was significantly above the chance baseline predicted by exemplar similarity ($M=0.86$, $SE=0.05$; $t(28)=7.5$, $p<0.01$, $d=1.40$). This suggests that participants successfully extracted the 'high-shading' regularity characterizing the Alpha category and applied it to the novel critical item, rather than relying solely on exemplar similarity.

Top-middle-bottom Results In this condition, the critical test items placed the domain-level regularity in direct conflict with exemplar similarity. To adjudicate between these strategies, we conducted a profile analysis to categorize individual participants based on their generalization patterns.

Domain-Level Regularity Profile: 7 participants fit the profile predicted by the domain-level regularity account (see Figure 3). These participants classified the item at (0.1, 0.7) as an Alpha (conforming to the high feature value regularity) and the item at (0.9, 0.3) as a Beta (conforming to the low feature value regularity), despite similarity favoring opposite classification. Exemplar Similarity Profile: 10 participants fit the profile predicted by exemplar models (see Figure 4). These participants classified the item at (0.1, 0.7) as a Beta and the item at (0.9, 0.3) as an Alpha; thereby aligning with the summed similarity of the stimuli. Concordant Profile: The remaining 7 participants displayed a concordant profile showing performance close to chance for both items (see Figure 5), indicating an ambiguous resolution to the conflict between similarity and regularity.

These data indicate that while some learners default to similarity-based generalization, a substantial subset of participants spontaneously override similarity to follow abstract domain-level regularities. While this profile analysis provides clear qualitative evidence of individual differences (with a distinct subset of learners actively defying exemplar similarity to follow the domain-level regularity) we acknowledge that the resulting sub-groups are small. Future work with larger samples is required to draw definitive conclusions about the true population distribution of these strategies.

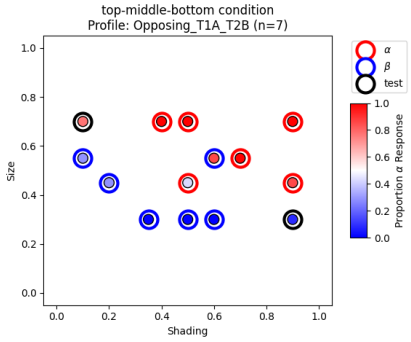

Figure 3: Stimulus structure and generalization performance- "Domain-Level Regularity" profile of Top-Middle-Bottom condition.

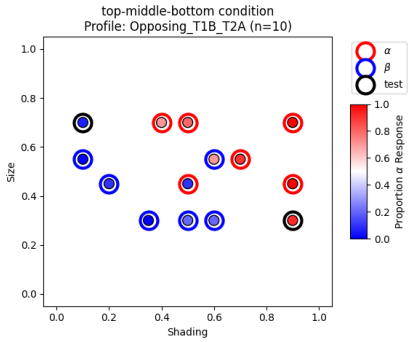

Figure 4: Stimulus structure and generalization performance- "Exemplar Similarity" profile of Top-Middle-Bottom condition.

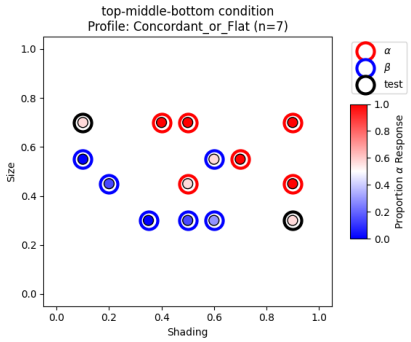

Figure 5: Stimulus structure and generalization performance- "Concordant" profile of Top-Middle-Bottom condition.

Discussion

The results of Experiment 1 provide compelling evidence that human learners are sensitive to domain-level regularities that standard exemplar models fail to capture. In both the Full Circle and Curtailed Circle conditions, the classification of critical test items violated the mechanics of standard exemplar models. Because standard models compute

summed similarity across all stored instances, the dense inner structure of the Beta category slightly greater gravitational pull on these test items. Exemplar theory thus dictates that generalization should hover near chance, or even lean toward Beta (in the full circle condition). Instead, participants systematically classified these items into the Alpha category at rates significantly above chance. This provides strong affirmative evidence that fitting the abstract role of a category member (e.g., being "extreme" or "high-valued") completely overrode local, exemplar-based similarity computations.

.

A potential counter-argument to the findings in the Full Circle condition involves the inherent difference in variability between the concentric categories; the outer "Alpha" ring naturally spans a larger area than the inner "Beta" ring. Proponents of exemplar theory might argue that the preference for the outer category is explained by a response bias parameter—a mechanism often invoked when categories differ in variability. Cohen, Nosofsky, and Zaki (2001) attempt to offer a principled justification for this parameter based on classification competition. They argue that a bias is necessary when a stimulus from a high-variability category has few neighbors of its own but faces strong competition from a dense cluster of low-variability items.

However, this rationale does not apply to the concentric circle structure used here. In this design, all exemplars of the high-variability category (the outer circle) are subject to the *same* level of competition from the low-variability category (the inner circle). Because the competition is uniform across the structure, there is no local density imbalance that necessitates a corrective bias. Consequently, attributing these results to a response bias is not an explanation, but merely a re-description of the data—simply stating that subjects chose the outer category more often. In contrast, our framework offers a generative explanation: subjects chose the outer category because they learned the domain-level regularity that Alpha items are extreme-valued relative to the moderate Beta items.

The Top-Middle-Bottom condition further sharpened this distinction by placing similarity and regularity in direct conflict. While some participants did align with exemplar model predictions, a distinct subset of learners (the "Domain-Level Regularity Profile") actively classified items against the predictions of exemplar similarity. This demonstrates that domain-level regularities are not merely used as tie-breakers when similarity is ambiguous. For a substantial proportion of learners, these abstractions appear to override local similarity computations entirely.

Collectively, these findings challenge the sufficiency of the exemplar account. The standard view posits that category knowledge is built bottom-up through the accumulation of specific instances. However, our data suggest that learners extract top-down principles regarding the distribution of features in the domain.

If humans are indeed extracting this richer generative information, it raises a further question about the breadth of their learning. Exemplar models are often characterized as "lazy" or economical—storing the data itself rather than estimating a function or model of the data and learning to ignore unneeded information. If human learning is driven by sensitivity to broader within-category regularities, then learners may encode information about category members that is not strictly necessary for the immediate classification task. We investigate this possibility in Experiment 2.

Experiment 2

Exemplar models optimize selective attention weights either as a hyperparameter (GCM) or based on an error-driven optimization process (e.g., ALCOVE; Kruschke, 1992). When faced with a multidimensional categorization task, these models shift attention broadly toward diagnostic dimensions and away from non-diagnostic ones. If a single dimension provides a perfect solution (100% accuracy), an exemplar system should attend exclusively to that dimension, effectively "ignoring" any additional, partially diagnostic information on other dimensions (e.g., such as in the classic SHJ Type 1 benchmark; Nosofsky, Gluck, Palmeri, McKinley, & Glauthier, 1994).

By contrast, if human learners are characteristically "rich" or generative-style encoders, learning the statistical structure of the environment rather than merely solving the immediate classification problem, then they should encode information about features that are redundant or only partially diagnostic. Experiment 2 tests this by presenting category structures where a simple unidimensional rule suffices to achieve perfect accuracy, but secondary dimensions contain 'extra' information. If subjects learn this extra information, it would violate the design principle of selective attention that is central to how exemplar models extend stimulus generalization theory to account for human category learning. It is important to note that while some previous work has shown that learners distribute attention more broadly when learning mode altered by manipulating task demands (e.g. Minda & Ross, 2004), the current work makes a more dramatic claim by employing a standard artificial classification training procedure without trying to actively push learners to be more generative.

Methods

Subjects As in Experiment 1, participants were undergraduate students. The same exclusion criterion was applied in retaining only those who achieved greater than 80% accuracy by the end of training. After applying this exclusion criterion, the final sample consisted of 141 participants distributed across two conditions: the Learn Extra Condition, 83 participants met the criterion (from an initial pool of 87); And the Uni vs. Proxi Condition, 58 participants met the criterion (from an initial pool of 61).

Stimuli and Design The stimuli were identical in type to Experiment 1 (squares varying in size and shading).

Participants were assigned to one of two conditions designed to pit a sufficient unidimensional solution against proximity based generalization on a weakly-diagnostic or non-diagnostic dimension.

Learn Extra Condition: The category structure was constructed such that one dimension (e.g., shading) provided a perfect unidimensional solution (see Figure 6). The second dimension (e.g., size) was only partially diagnostic. To determine if participants encoded the 'extra' information on the partially diagnostic dimension, we introduced critical test items located across the unidimensional boundary. Importantly, in the full two-dimensional space, these items were highly similar to the *opposite* category via the partially diagnostic dimension. Standard exemplar models, relying on selective attention, predict that learners will focus solely on the perfectly diagnostic dimension. Consequently, these test items should be classified according to the unidimensional boundary. However, if participants learn the extra information about the partially diagnostic dimension, their classification of these items should be influenced by similarity on that dimension which pulls them away from the unidimensional solution. As a secondary probe of knowledge, we included a single-feature classification task in which participants were asked at test to classify items based on only one dimension cue.

Uni vs. Proxi Condition: This condition followed a similar design principle (see Figures 7 and 8). A clear unidimensional rule (Uni) was available, but critical test items were positioned such that their proximity (Proxi) to the training items in 2D space conflicted with the strict linear boundary. This condition served to replicate the "learning extra" effect in a structure where the distribution of exemplars was more sparse.

Procedure The training procedure mirrored Experiment 1 (15 blocks with feedback). However, for the Learn Extra condition, the testing phase was expanded. In addition to the standard Full-Feature Classification test (classifying the 2D squares), participants completed a Single-Feature Classification test. To control for order effects, the sequence of these test phases (Full-Feature vs. Single-Feature) was counterbalanced across participants.

Results

Learn-Extra Results The primary hypothesis for this condition was that human learners would encode information about a partially diagnostic dimension even when a perfect unidimensional solution existed—a direct violation of selective attention optimization predicted by exemplar models.

The full-feature classification results support this hypothesis. As illustrated in Figure 6, generalization performance on the critical test items was dominantly driven by the partially diagnostic feature (Size) rather than aligning with the strict unidimensional solution (Shading).

To confirm that participants genuinely encoded the partially diagnostic dimension, we analyzed accuracy during the Single Feature test phase. Crucially, performance in this phase was robustly above chance for both single feature first (M = 0.74, SE=0.014, t(42)=16.672, p<0.01, d=2.50) and single feature second groups (M = 0.75, SE=0.014, t(39)=18.366, p<0.01, d=2.90). This confirms that participants did not merely downgrade ignore the partially diagnostic dimension as predicted by selective attention mechanisms; rather, they acquired a stable representation of its structure despite it being redundant for the classification task.

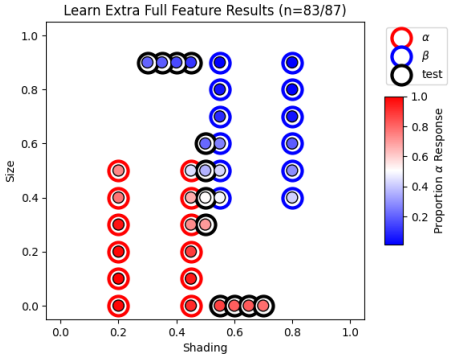

Figure 6: Stimulus structure and generalization performance- "Learn Extra" condition.

Uni vs. Proxi Results In this condition, the critical test items pitted the strict unidimensional rule against similarity on the non-diagnostic dimension. A profile analysis was conducted to classify participants based on their generalization strategies. Proximity Dominant (Alpha Response): 27 subjects displayed an "Alpha Dominant" response pattern. For these learners, the high similarity of the test items to the Alpha items on the non-diagnostic dimension overrode the unidimensional boundary, leading them to classify the items contrary to the perfect rule. Unidimensional solution (Beta Response): 31 subjects displayed a "Beta Dominant" response pattern. These learners adhered to the unidimensional solution, classifying the test items based on the diagnostic dimension and ignoring the proximity conflict.

The presence of a substantial "Proximity Dominant" group (n=27) further corroborates the findings of the Learn Extra condition—demonstrating that nearly half of the participants encoded and utilized unnecessary information even when it conflicted with a perfect deterministic rule.

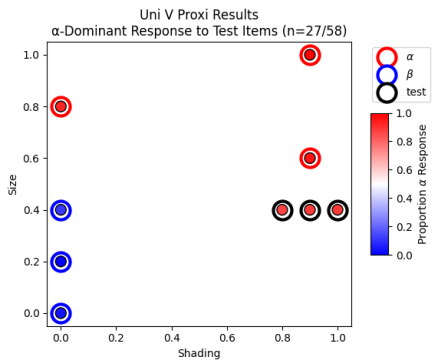

Figure 7: Stimulus structure and generalization performance- “Alpha dominant” profile of “Uni vs Proxi” condition.

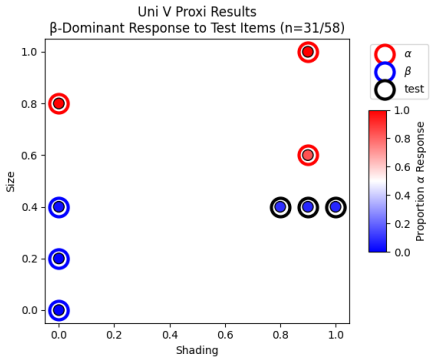

Figure 8: Stimulus structure and generalization performance- “Beta dominant” profile of “Uni vs Proxi” condition.

Discussion

According to exemplar theory, learners should allocate attention to diagnostic dimensions while learning to ignore irrelevant or redundant features. In the Learn Extra condition, participants were provided with a perfect unidimensional solution. Despite this, they spontaneously encoded significant information about the partially diagnostic dimension. This was evident in two distinct ways. First, generalization to the critical test items was heavily influenced by the secondary feature, pulling classification away from the strict unidimensional boundary. Second, performance on the single-feature test phase was robustly above chance. Methodological checks confirmed that this knowledge was not an artifact of the testing sequence.

The Uni vs. Proxi condition reinforced this finding by placing the extra information in direct conflict with a deterministic rule. The fact that nearly half of the participants followed a proximity-based strategy—classifying items based on their overall similarity to the Alpha category along the non-diagnostic dimension rather than the clear unidimensional rule—presents a challenge for exemplar models. Exemplar models could potentially fit this behavior by turning selective attention off, but this would impact ease of learning predictions and there is no evident reason why half the learners would not invoke selective attention.

General Discussion

Exemplar models have loomed large in the field of category learning—due to their ability to account for learning and generalization performance in terms of storage of instances. However, the present findings challenge the completeness of this account. Across two experiments, we demonstrate that human learners diverge from exemplar predictions by extracting abstract regularities and encoding ‘extra’ information.

Standard exemplar models rely on summed similarity and mechanisms of discriminative efficiency. Our results contradict both principles. In Experiment 1, participants consistently classified novel test items based on domain-level regularities, doing so at rates that deviated from the predictions of summed distance/similarity.In Experiment 2, participants proved to be ‘rich’ encoders. Even when a perfect unidimensional rule was available, subjects employed additional information to guide generalization and spontaneously encoded secondary, redundant features, as evidenced by their ability to classify items based on these unnecessary dimensions alone.

To ensure our conclusions are not artifacts of a specific model parameterization, we evaluated our findings against an exhaustive grid search of the standard Generalized Context Model’s (GCM) entire parameter space (footnote 1). This comprehensive mapping reveals that to account for our data, the exemplar framework must rely on psychologically implausible assumptions.

In Experiment 1, the GCM structurally struggled to represent the base circular categories. While human-like generalization could be forced using certain parameterizations (e.g., sensitivity c>7, combined with response determinism >4), doing so required the model to apply strong selective attention (w ≥0.7). It is psychologically implausible that human learners would aggressively selectively attend to a single dimension of a symmetrical circular structure where both dimensions possess equal variance and diagnostic utility.

Experiment 2 traps the exemplar account in the exact opposite theoretical contradiction. While the GCM can qualitatively recreate the "Learn Extra" results, it does so by assuming an even distribution of attention across dimensions. However, this violates a core design principle of exemplar theory: attention-weight optimization. We intentionally provided a perfect, highly diagnostic unidimensional feature; standard exemplar mechanisms dictate that learners should heavily weight this feature to maximize cognitive economy. Instead, participants abandoned selective attention where it would have been optimal, choosing instead to encode the redundant features. Therefore, to fit human behavior across these tasks, the GCM must assume that learners fiercely apply selective attention where it is structurally useless (Experiment 1), while completely failing to apply it where it is a perfect solution (Experiment 2)

These divergences highlight a theoretical shortcoming. Exemplar theory characterizes learning as a discriminative process—optimizing attention to distinguish Category A from Category B with minimal ambiguity. Consequently, it treats domain-level regularities and redundant features as irrelevant if they do not aid immediate discrimination. In contrast, the present results support a generative view: learners appear to model the underlying nature of the categories. By encoding that a category has extreme values or has multiple correlated features, humans build a robust model that supports category use and inference beyond the immediate task requirements. This aligns with categories serving as building blocks for knowledge representation and inferential reasoning as opposed to strictly task-optimized stimulus generalization.

References

Cohen, A. L., Nosofsky, R. M., & Zaki, S. R. (2001). Category variability, exemplar similarity, and perceptual classification. *Memory & Cognition*, *29*(8), 1165-1175.

Conaway, N. (2016). *Re-evaluating the reference point view of human classification learning*. State University of New York at Binghamton.

Conaway, N., & Kurtz, K. J. (2017). Similar to the category, but not the exemplars: A study of generalization. *Psychonomic bulletin & review*, *24*(4), 1312-1323.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99*(1), 22–44. https://doi.org/10.1037/0033-295X.99.1.22

Kruschke, J. K. (1993). Human category learning: Implications for backpropagation models. *Connection Science*, *5*(1), 3-36.

Kruschke, J. K. (1996). Dimensional relevance shifts in category learning. *Connection Science*, *8*(2), 225-248.

Kruschke, J. K. (2001). The inverse base-rate effect is not explained by eliminative inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1385.

Kruschke, J. K. (2003). Attentional Theory Is a Viable Explanation of the Inverse Base Rate Effect: A Reply to Winman, Wennerholm, and Juslin (2003).

Kurtz, K. J., & Wetzel, M. T. (2021). On the generalization of simple alternating category structures. *Cognitive Science*, *45*(4), e12972.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological review*, *85*(3), 207.

Minda, J. P., & Ross, B. H. (2004). Learning categories by making predictions: An investigation of indirect category learning. *Memory & cognition*, *32*(8), 1355-1368.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, memory, and cognition*, *10*(1), 104.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of experimental psychology: General*, *115*(1), 39.

Nosofsky, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: learning, memory, and cognition*, *14*(1), 54.

Nosofsky, R. M., Gluck, M. A., Palmeri, T. J., McKinley, S. C., & Glauthier, P. (1994). Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961). *Memory & cognition*, *22*(3), 352-369.

Qiu, T., & Minda, J. (2023). Novel Concentric-Circle Technique Interrogates Implicit Category Learning.