Proceedings of the Annual Meeting of the Cognitive Science Society

Gaze Insights into Partially-Encoded Representations of Objects and Categories

Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

Vance, Jared
Darby, Kevin P
Weichart, Emily R.

2025

Gaze Insights into Partially-Encoded Representations of Objects and Categories

Jared Vance (jared.vance@usu.edu)
Department of Psychology, Utah State University
2810 Old Main Hill, Logan, UT 84341 USA

Kevin P. Darby (darbyk@fau.edu)
Department of Psychology, Florida Atlantic University
777 Glades Road, Boca Raton, FL 33431 USA

Emily R. Weichart (emily.weichart@usu.edu)
Department of Psychology, Utah State University
2810 Old Main Hill, Logan, UT 84341 USA

Abstract

Studies of category learning have revealed individual differences in decision-making, such that the same stimulus may be categorized differently across individuals. Modeling accounts have explained these differences in terms of how attention weights are distributed across stimulus dimensions that distinguish between category responses. These weights are typically assumed to reflect an individual's beliefs about which dimensions are most relevant to their goals. The current work investigates the possibility that instead of being purely strategic, attention weights are constrained by what was encoded into memory during learning. Participants (N=120, age 18-25) completed a category learning task while gaze was recorded as an exogenous measure of attention. Model-based analyses using gaze to predict behavior revealed that accounting for partially-encoded representations was necessary for predicting individual differences in feature memory and categorization.

Keywords: concepts and categories; learning; memory; computational modeling; eye tracking

Introduction

Recognizing items in the environment as members of particular cognitive categories and selecting appropriate actions in response are foundational to all forms of goal-directed decision-making. Decision-making, however, is inherently variable. Given the same stimulus and even the same goal, individuals differ in how they interpret and attend to perceptual features in the service of a decision strategy (Medin & Shafer, 1978; Medin & Smith, 1981). A person attending to the botanical properties of produce (e.g., whether it contains seeds) might categorize items like tomatoes, cucumbers, and squash as *fruits*, whereas someone attending to culinary uses (e.g., whether an item is used in sweet vs. savory dishes) might categorize these same items as *vegetables*. More broadly, different individuals may categorize the same t-shirt as tacky or stylish, the same facial expression as friendly or condescending, and the same defendant as a criminal or a victim of circumstance. Understanding how and why people prioritize different dimensions of information is key to explaining a wide range of human preferences, judgments, and decisions.

A common theoretical assumption is that people categorize new stimuli by determining their similarity to stored representations in memory (Logan, 1988). The concept of attentional weighting explains how decision variability emerges, whereby humans and other animals prioritize dimensions that are relevant to their goals or which have proven useful for predicting rewards (Lawrence, 1949; Nosofsky, 1986; Pearce, 1987; Livingston & Andrews, 1995). In his influential Generalized Context Model (GCM; Figure 1A-B), for example, Nosofsky (1986) described attention shifts in terms of psychologically "stretching" stimulus dimensions that are relevant for categorization and "shrinking" those that are irrelevant. In many subsequent computational accounts, attention has similarly been implemented as decision weight parameters that modulate how dimensions influence decision processes (Busemeyer, 1985; Kruschke, 1992; Busemeyer & Townsend, 1993; Love, Medin, & Gureckis, 2004; Love et al., 2004; Krajbich & Rangel, 2011).

Several studies have sought neurophysiological support for the concept of attentional weighting. For example, studies using eye-tracking as a measure of attention during category learning have shown that gaze preferences dynamically shift toward task-relevant dimensions as learning progresses (Blair, Watson, & Meier, 2009; Blair, Watson, Walshe, & Maj, 2009; Rehder & Hoffman, 2005a) and correlate with attention weights derived from GCM (Rehder & Hoffman, 2005b). Using fMRI, Braunlich and Love (2019) found that individual differences in model-estimated attention weights aligned with activation in visual processing areas, such that more attended dimensions were more readily decoded from BOLD signals using multivariate pattern analysis. Together, these findings position attention as a mechanism for optimizing behavior by selectively weighting dimensions at the time of decision. This focus, however, overlooks a key possibility: in the process of acquiring these optimized attentional strategies, learners may inadvertently impose limitations on what is stored in memory.

Weichart, Galdo, Sloutsky, and Turner (2022) explored this idea by developing a process model of category learning in which within-trial feature sampling and between-trial learning unfold over time (see also Braunlich and Love, 2022). Building upon insights from sequential sampling models (Smith & Ratcliff, 2004), the model proposed that learners shift their gaze among stimulus dimensions only until they acquire enough evidence to make a category decision. While

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

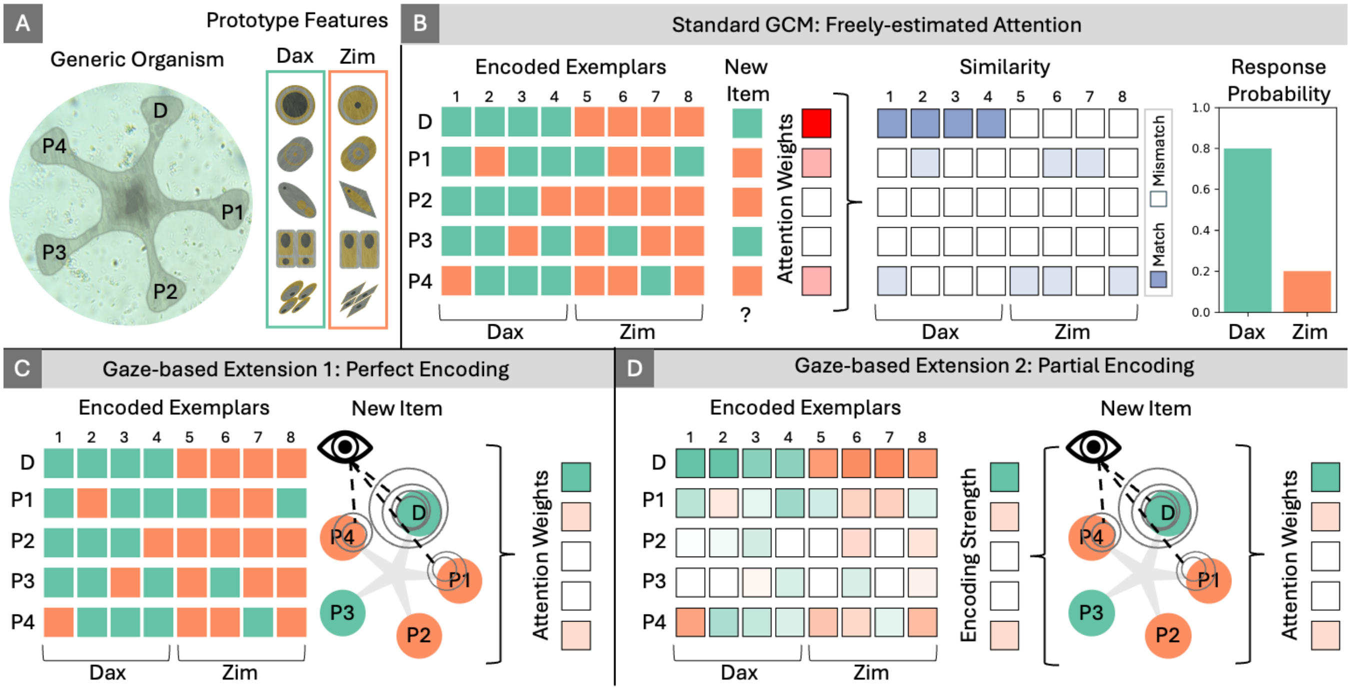

Figure 1: Category structure and theoretical overview. *D=deterministic; P(1-4)=probabilistic; GCM=Generalized Context Model.* **(A)** Dax and Zim prototypes. Organelles appeared in marked locations on the generic organism. **(B)** GCM describes categorization decisions as a feature-level comparison between the new item and labeled exemplars in memory, weighted by attention. Green and orange squares represent features drawn from Dax and Zim prototypes, respectively. Features in the encoded exemplar matrix correspond to the 8 unique items presented during training. **(C)** One interpretation of attention in GCM: decision weights that govern test decisions are imposed upon all features that were presented during training. **(D)** Second interpretation: decision weights are imposed upon the features that were encoded by the individual participant.

computationally efficient, this policy predicts that when sampling terminates before all dimensions are fixated, the resulting representation is only *partially* encoded. In a series of simulation studies, Weichart, Galdo, et al. (2022) showed that accounting for partially encoded representations was necessary to explain individual differences in both eye-tracking behavior and categorization performance. Follow-up analyses using the same framework showed that attentional updating mechanisms recruited a diverse set of brain regions, including those known to be involved in memory encoding (Weichart, Evans, Galdo, Bahg, & Turner, 2022).

Here, we directly tested the hypothesis that attention weights are constrained by partially-encoded representations. Participants learned to sort multidimensional stimuli into novel categories while gaze was recorded. We then leveraged GCM in a *direct input* approach, whereby gaze data replaced the otherwise freely-estimated attention parameters in the model (Turner, Forstmann, Love, Palermi, & van Maanen, 2017; Weichart, Unger, King, Sloutsky, & Turner, 2024). This approach supported a data-driven comparison of two theoretical alternatives regarding what attention weights tell us about individual differences in decision-making: Do they purely reflect the strategic selection of relevant information at the time of the decision (Figure 1C), or do they additionally reveal constraints imposed by what was encoded during learning (Figure 1D)?

Method[1]

Participants

Participants (N=124) were recruited from Utah State University (USU) and compensated with partial course credit. Four were excluded due to eye tracking calibration failure or failure to complete the experiment. Results are reported for the remaining 120 participants (81 females, 36 males, 3 unreported gender; age range 18-25 years, mean=19.8, SD=1.8). All research was approved by the Institutional Review Board at USU.

Stimuli

Stimuli were fictional amoeba-like organisms that were presented as if viewed through a microscope. Each organism consisted of five dimensions that could contain one of two possible features (organelles). Category membership was defined by similarity to two prototypes, “Dax” and “Zim” (Figure 1A). Training exemplars contained four features from one prototype and one from the other, creating a family-resemblance structure. One dimension was perfectly *deterministic* of category membership while the remaining dimensions were *probabilistically* predictive. We refer to these dimensions as “D” and “P(1-4),” respectively. As shown in Figure 1A-B, the D feature from the Dax prototype appeared in

---

[1] Methods are summarized here due to space constraints, but additional details are provided in Weichart and Darby (forthcoming).

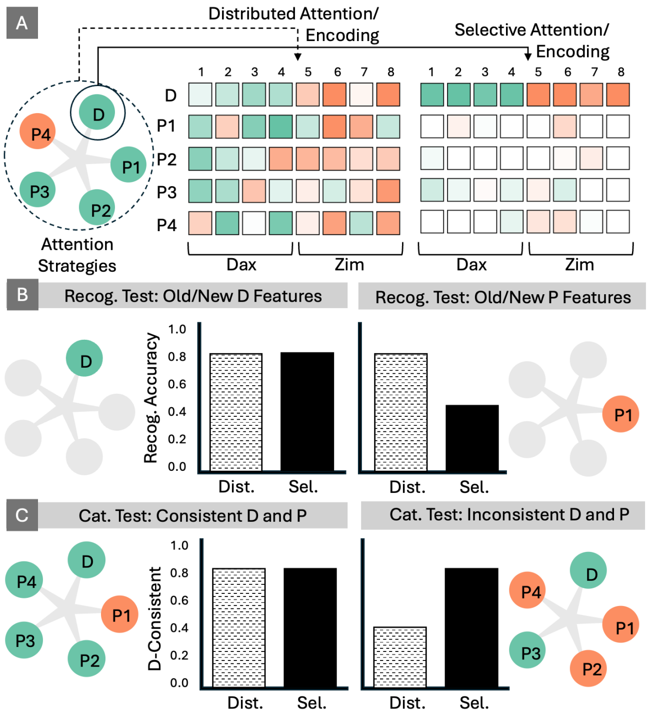

Figure 2: Hypothesized impacts of partial encoding on behavior. *D=deterministic; P(1-4)=probabilistic; Dist=distributed; Sel=selective; Recog=recognition; Cat=categorization.* **(A)** Feature encoding expected to reflect whether participants distributed gaze broadly or narrowly during training. **(B)** Predicted D and P feature recognition performance, given distributed or selective encoding. **(C)** Predicted proportions of responses favoring the D feature when D and P features provide consistent or mixed category evidence.

100% of Dax exemplars, and each P feature from the Dax prototype appeared in 66%. Zim exemplars followed an analogous structure. Dimension designations as D or P and feature designations as Dax or Zim were counterbalanced between participants.

Apparatus

Each participant was seated in an individual testing room in view of an experimenter. A Tobii Pro Spectrum eye tracker recorded binocular eye movements at a sampling rate of 600 Hz. Stimuli were presented on an EIZO FlexScan EV2451 monitor using PsychoPy 1.85.0 (Peirce, 2009). Responses were made using a standard keyboard. Before beginning the experiment, participants provided informed consent, completed an eye tracking calibration and validation routine, and completed 10 practice trials where they learned to map features to categories via trial-and-error. Practice stimuli used different features from the main task stimuli to avoid carryover.

Procedure

Training Participants were instructed to use trial-and-error to categorize each organism as a Dax or Zim by looking at its organelles. Participants examined each stimulus for as long as they wanted while the eye-tracker recorded dimension-level dwell times. When they were ready to make a category judgment, they pressed the spacebar to proceed. The stimulus was removed, eye-tracking was suspended, and text was presented to indicate the two species options and their respective response mappings (e.g. DAX F; ZIM J). Feedback appeared for 1200ms after the response: “correct” (green text) or “incorrect” (red text). The category structure (Figure 1A-B) resulted in 4 unique stimuli representing the Dax and Zim species, respectively. Each stimulus was presented 15 times during training in randomized order, resulting in 120 trials. Figure 2A illustrates our hypotheses regarding how one’s attention strategy during training would affect the encoded category representations. Distributing attention broadly across dimensions should result in broad feature encoding, whereas selectively attending to the most informative (D) dimension should result in more focused representations and reduced encoding of unattended features.

Recognition Test Following training, we examined how well participants could distinguish between previously-seen and novel features. Stimuli were the 10 organelles that comprised the Dax and Zim prototypes, and 10 visually-similar lures (2 corresponding to each dimension). Organelles were presented individually against the organism backdrop shown in Figure 1A. Participants responded *”OLD”* if they recognized the organelle from training, and otherwise responded *”NEW.”* Each organelle was presented 5 times, resulting in 100 trials. Gaze measures were not recorded, and correct/incorrect feedback was not provided. Figure 2B shows predicted feature recognition when participants used either a distributed or selective attention strategy. While both strategies were expected to support accurate discrimination between old and new D features, we predicted that participants who selectively attended to D would exhibit poorer recognition of P features compared to those who distributed attention more broadly.

Categorization Test Participants categorized old and new organisms while dimension-level dwell times were recorded, and feedback was not provided. Organisms contained a majority of P features that were either *consistent* (34 items) or *inconsistent* (30 items) with the species associated with the D feature. Figure 2C illustrates the predicted response patterns for the two attention strategies of interest. When D and P features provide consistent category evidence (as in training), participants are expected to categorize items according to the D feature, regardless of whether they adopted a distributed or selective attention strategy. However, when presented with novel items containing mixed evidence from D and P features, we expected divergent patterns: distributed attention should lead to responses reflecting the majority of features, while selective attention should produce responses aligned with the D feature alone.

Computational Modeling

We used the GCM framework in two ways. First, we fit the standard model to each participant's Categorization Test behavior and freely-estimated parameters representing attention (Figure 1B). This step allowed us to identify participants whose responses were consistent with broadly distributed or D-selective attention strategies. We next further examined what these freely-estimated attention parameters represent: decision weights spontaneously allocated across the dimensions of the test stimulus (Figure 1C), or decision weights constrained by a partially-encoded category representation (Figure 1D). To do this, we developed and compared variants of GCM in which the typically freely-estimated attention parameters were replaced by gaze data. We will first provide the relevant computations of GCM as described by Nosofsky (1986) for reference, then we will provide further details of our two-stage GCM analysis.

The Generalized Context Model (GCM) The exemplar presented on trial n of Training is denoted by vector $x^{(n)} = [x_1\ x_2\ \dots\ x_J]$ where each element corresponds to the feature value in dimension j. All N exemplars are stored in matrix $X = [x^{(1)}\ \dots\ x^{(N)}]^\top$ and associated feedback is stored in vector $F = [f^{(1)}\ \dots\ f^{(N)}]$. On each test trial i, the observer is presented with probe $e^{(i)} = [e_1\ e_2\ \dots\ e_J]$. To categorize $e^{(i)}$, the observer computes the perceived similarity between the probe and each exemplar $x^{(n)}$:

$$s_j\left(e^{(i)}, x^{(n)}\right) = exp\left(-\delta d_j\left(e^{(i)}, x^{(n)}\right)\alpha_j\right). \qquad (1)$$

Here, free parameters δ and α represent similarity kernel specificity and dimension-level attention, respectively. Distance $d_j\left(e^{(i)}, x^{(n)}\right)$ is simply a match/mismatch comparison between corresponding features:

$$d_j\left(e^{(i)}, x^{(n)}\right) = \begin{cases} 1 & \text{if } e_j{}^{(i)} \neq x_j{}^{(n)} \\ 0 & \text{otherwise.} \end{cases} \qquad (2)$$

The observer combines feature similarity across dimensions to compute the activation of each exemplar, given the probe. In categorization, the probability of assigning the probe to a given category is determined by the proportion of total exemplar activation associated with that category label. In recognition, the probability of responding "old" is based on the total activation across all exemplars relative to a threshold.

Individual Differences We used a switchboard analysis of GCM variants with freely-estimated attention to delineate participants according to Categorization Test behavior. We fit 3 models to data from each participant, instantiating the candidate attention strategies via Equation 1 as follows:

1) **Selective attention**: The value of attention allocated to the D dimension (α_D) was freely-estimated with constraints $\frac{1}{5} < \alpha_D < 1$. Attention to each P dimension was set to $\alpha_{P_n} = \frac{1-\alpha_D}{4}$. δ was freely estimated, resulting in 2 total free parameters.

2) **Distributed attention**: Attention values across all dimensions were constrained to be equal, such that $\alpha_D = \alpha_{P_n} = \frac{1}{5}$. δ was freely estimated, resulting in 1 free parameter.

3) **Random responding**: Values of α and δ were fixed to 0, resulting in predictions of chance performance across trials. This null model thus had 0 free parameters.

After fitting the models, a comparison of Akaike Information Criterion (AIC; Akaike, 1974) was used to select the most representative strategy for each participant. Because this fit statistic accounts for parsimony, freeing δ and/or α_D was favored only if it significantly improved model fit compared to the null (random responding) model.

Gaze-based Modeling We conducted a second GCM comparison in which freely-estimated attention parameters were eliminated, and gaze was a direct input to α instead. This model-based cognitive neuroscience technique is ideal for examining whether a measured signal (in this case, dimension-level dwell time) plays a mechanistic role in the cognitive process under study (Turner et al., 2017). We developed a suite of models in which 1) *partially encoded* memory traces constrained decision processes at Test, or 2) Training stimuli were *perfectly encoded*, such that attention weights directly reflect gaze in the presence of the Test stimulus. To this end, we specified:

$$\alpha_j = \eta_j^{(n)} \zeta_j^{(i)} \qquad (3)$$

where $\eta_j^{(n)}$ represents encoding of the feature presented in dimension j on Training trial n, and $\zeta_j^{(i)}$ represents the decision weight allocated to dimension j on Test trial i.

We selected the functions in Table 1 to transform feature-level dwell times to a 0-1 range, freely estimating shape parameters θ and ω as needed. The input $dwell_j^{(t)}$ refers to the total time spent looking at the feature in dimension j on Trial t, and output is denoted $v_j^{(t)}$. This transformed value was used directly as $\eta_j^{(n)}$ when applied to training fixations, and as $\zeta_j^{(i)}$ when applied to test fixations.

Function	Equation
X: Perfect encoding	$v_j^{(t)} = 1$
A: Binary step	$v_j^{(t)} = \begin{cases} 0 & \text{if } dwell_j^{(t)} \leq \theta, \\ 1 & \text{otherwise.} \end{cases}$
B: Piecewise linear	$v_j^{(t)} = \min\left(\frac{dwell_j^{(t)}}{\theta}, 1\right)$
C: Softmax	$v_j^{(t)} = \frac{\exp\left(\theta dwell_j^{(t)}\right)}{\sum_k \exp\left(\theta dwell_k^{(t)}\right)}$
D: Logistic	$v_j^{(t)} = \frac{1}{1+\exp\left(-\theta\left(dwell_j^{(t)} - \omega\right)\right)}$

Table 1: Encoding functions for dimension-level $v_j^{(t)}$.

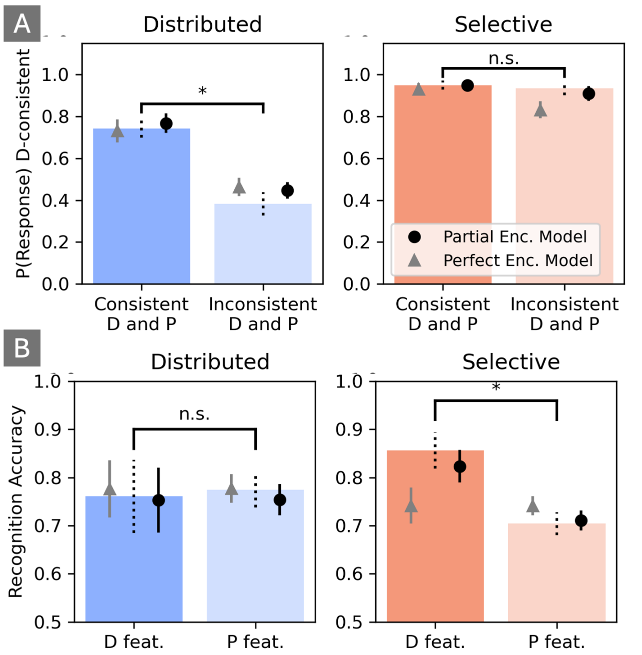

Figure 3: Observed and gaze-predicted behavior at test. *D=deterministic; P=probabilistic; Enc=encoding; n.s.=not significant; *=Bayesian HDI does not contain 0.* Bars show observed means. Points show predictions from best-fitting partial and perfect encoding models. **(A)** Recognition accuracy for D and P features. Facets are strategy groups identified using GCM with freely-estimated attention parameters. **(C)** Proportion of D-consistent responses when D and P features were drawn from the same (consistent) or opposite (inconsistent) category prototypes.

A suite of 25 models represented every pairwise combination of functions for transforming gaze during training and test. Models were fit to subject-level behavior from recognition and categorization tests and were evaluated using AIC.

Results

Feature-level dwell times during training predict recognition accuracy. To examine the relationship between gaze and encoding, we used Bayesian mixed-effects logistic regression to predict recognition accuracy from feature-level dwell times during Training. Feature type (D or P) was a within-subjects factor. We identified an extremely strong main effect of dwell time on recognition accuracy (Bayes factor >100; mean (M) = 2.51; 95% highest density interval (HDI) = [1.26, 3.74]).

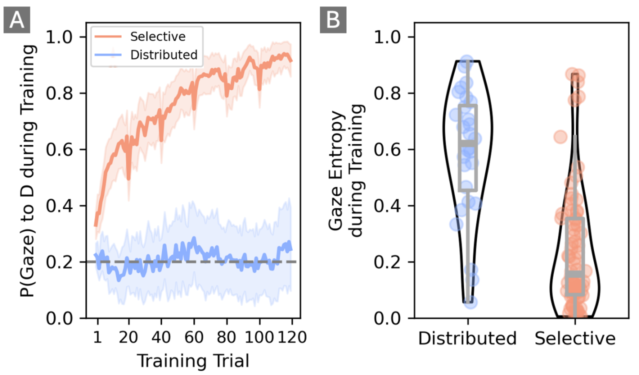

Figure 4: Gaze during training, given behavior at test. *D=deterministic.* **(A)** Group-level proportion of fixations to the deterministic (D) dimension during training. 0.2 is chance, given 5 organelles. **(B)** Group-level mean gaze entropy during training. Values are maximized when gaze is allocated equally to all dimensions.

Selective attention is linked to selective feature memory. The strategy group delineation analysis revealed that 74 participants selectively attended to D and 28 distributed attention across dimensions. The remaining 18 responded randomly, and were excluded from subsequent analyses. Figure 3A shows that these GCM-derived attention weights bore measurable impacts on categorization test behavior that were consistent with our predictions (Figure 2C). The selective attention group responded consistently with D regardless of the P features (Bayesian t-test: $M_{incDP} - M_{conDP} = -0.01,\ 95\%\ HDI = [-0.05, -0.02]$), and the distributed attention group made fewer D-consistent responses when P features provided evidence for the opposite category ($M_{incDP} - M_{conDP} = -0.36,\ 95\%\ HDI = [-0.42, -0.30]$). Although groups were defined by categorization test behavior, we retrospectively examined feature recognition performance as well. If GCM attention weights reflect purely strategic decision biases imposed on a perfectly-encoded representation, then both groups should be able to recognize D and P features. If weights instead reflect differences in the encoded category representations, selective and distributed attention should differ in their ability to recognize P features (Figure 2B). The results shown in Figure 3B favor the latter account: while the distributed attention group showed no difference between feature types ($M_P - M_D = 0.01,\ 95\%\ HDI = [-0.07, 0.09]$), the selective attention group showed worse recognition of P features compared to D ($M_P - M_D = -0.15,\ 95\%\ HDI = [-0.20, -0.11]$).

Categorization biases at test are preceded by gaze differences during training. As a precursor to the gaze-based modeling analyses to follow, we back-sorted gaze acquired during training based on the attention strategies utilized during test. As shown in Figure 4A, the selective attention group tended to focus their gaze upon the D dimension during training, while the distributed group did not look at D more often than expected by chance (20% of dwell time, given 5 dimensions). We quantified the spread of fixation probabilities across the available dimensions using gaze entropy (H) (Shannon, 1948):

$$H = -\frac{1}{m}\sum_{i=1}^{5} p_j \log_2(p_j) \tag{4}$$

where p_i is the proportion of dwell time allocated to each dimension j, and m is the maximum possible entropy given 5-dimensional stimuli: $m = -\sum_{i=1}^{5}(0.2\log_2(0.2)) \approx 2.32$. As shown in Figure 4B, the distributed attention group showed

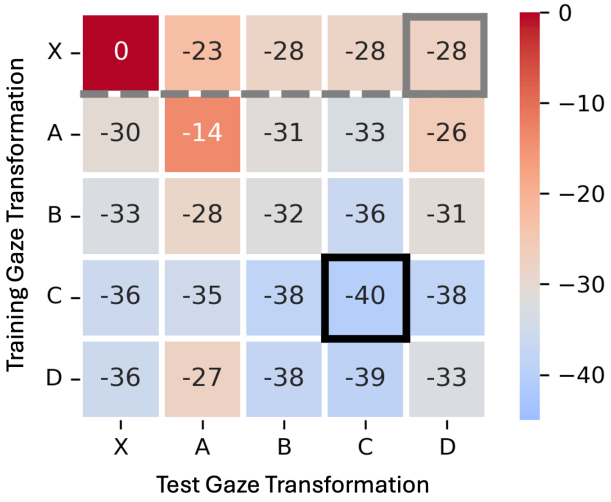

Figure 5: Gaze-based model comparison. Mean AIC differences relative to null model X-X. Lower values indicate better fit. Each cell represents a model with transformations applied to gaze during training (rows) and categorization test (columns). Gray and black boxes indicate the best-fitting perfect- and partial-encoding models, respectively.

higher gaze entropy during training compared to the selective attention group ($M_{dist} - M_{sel} = 0.35$, 95% HDI$=[0.26, 0.45]$). Taken together, Figures 4A and B show that the distributed attention group did not prioritize D because they spread attention across multiple dimensions, not because they erroneously focused on a single P feature.

Test decisions are constrained by partially-encoded representations. We compared 25 models in which dimension-level gaze replaced the typically freely-estimated attention parameters in GCM. Models are labeled in the format "T1-T2," where T1 and T2 refer to transformations applied to gaze data from training and test, respectively, as denoted in Table 1. Comparisons of mean AIC values are shown in Figure 5, where lower values indicate better fit relative to the null model X-X ($\eta_j^{(i)} = \zeta_j^{(i)} = 1$). The best-fitting model overall was partial encoding Model C-C (black box, Figure 5). This model outperformed even the best perfect encoding model (X-D; gray box, Figure 5). Aggregate simulations using each participant's best-fitting parameters from Models C-C and X-D are shown as points in Figure 3. Model C-C, along with most other partial encoding models, successfully reproduced key behavioral differences between strategy groups despite not including any free parameters to account for dimension-level biases. By contrast, Model X-D failed to capture the behavioral profile of the selective attention group.

Discussion

Findings from our model-based analyses demonstrate that gaze not only functions as an exogenous expression of attention, but also imposes constraints on the information that is available for use in future decisions. On the surface, this may not seem very surprising. The idea that selective attention imposes a gate on memory encoding is one of the most well-known principles in Psychology. Any undergraduate who has taken an introductory course can recount the classic "invisible gorilla" study, in which participants failed to remember a salient figure while their attention was directed elsewhere (Simons & Chabris, 1999). Despite intuitive appeal, formal models of category learning have largely overlooked the representational consequences of attention and have instead treated attention weights as a strategic filter applied at the time of decision (but see Weichart, Galdo, et al., 2022). The distinction, however, is meaningful. Rather than reflecting the observer's beliefs about which dimensions best-support the current goal, we suggest that attention weights are subject to flaws incurred through early experience, which may be skewed by salience, distraction, or pursuit of a different goal. This interpretation echoes Simon's (1955) "bounded rationality" perspective that cognitive processes reflect the adaptive use of limited resources, rather than optimal reasoning under conditions of complete knowledge about the environment (see also Anderson, 1990).

Our findings are also consistent with recent work by Weichart et al. (2024), who used an analogous direct input approach to examine category learning in 3- to 4-year-old children. Like the present study, their results favored a partial encoding account: gaze data collected during learning improved behavioral predictions by constraining the availability of features at test. Many studies have documented the profound cognitive changes that occur between early childhood and adulthood, particularly in the acquired capacity for engaging selective attention during learning (Huang-Pollock, Maddox, & Karalunas, 2011; Rabi & Minda, 2014). Converging evidence across samples of young children and adults, despite known differences in attentional control, therefore make a compelling argument that partial encoding may be a general property of human learning.

However, our findings tentatively contrast with results reported by Mack, Love, and Preston (2016) in an fMRI study of category learning. Evidence from a model-based analysis showed that the hippocampus can adaptively reshape encoded representations to reflect changing categorization rules. The current findings instead suggest that when information is not attended and encoded during learning, it cannot be retrospectively reclaimed as needed. While some degree of representational updating is likely possible, we propose that the capacity to reorganize is bounded by the contents of memory itself.

Taken together, our results show that the flexibility of category representations depends not only on learners' goals, but also on the limits of what was—and was not—encoded during learning.

Acknowledgments

This work was funded by Research Catalyst grant A62271 awarded to Emily Weichart by the Office of Research at USU. The authors thank Ethan Bardsley and Hailey Johnson for assisting with data collection.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. (doi: 10.1109/TAC.1974.1100705)

Anderson, J. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Blair, M., Watson, M., & Meier, K. (2009). Errors, efficiency, and the interplay between attention and category learning. *Cognition*, *112*(2), 330–336.

Blair, M., Watson, M., Walshe, C., & Maj, F. (2009). Extremely selective attention: Eye-tracking studies of the dynamic allocation of attention to stimulus features in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(5), 1196–1206.

Braunlich, K., & Love, B. (2019). Occipitotemporal representations reflect individual differences in conceptual knowledge. *Journal of Experimental Psychology: General*, *148*(7), 1192–1203.

Braunlich, K., & Love, B. (2022). Bidirectional influences of information sampling and concept learning. *Psychological Review*, *129*(2), 213–234.

Busemeyer, J. (1985). Decision making under uncertainty: A comparison of simple scalability, fixed-sample, and sequential-sampling models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(3), 583–564.

Busemeyer, J., & Townsend, J. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432–549.

Huang-Pollock, C., Maddox, W., & Karalunas, S. (2011). Development of implicit and explicit category learning. *Journal of Experimental Child Psychology*, *109*(3), 321–335.

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, *108*(33), 13852–13857.

Kruschke, J. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, *99*, 22–44.

Lawrence, D. (1949). Acquired distinctiveness of cue. i. transfer between discriminations on the basis of familiarity with the stimulus. *Journal of Experimental Psychology*, *39*(6), 770–784.

Livingston, K., & Andrews, J. (1995). On the interaction of prior knowledge and stimulus structure in category learning. *The Quarterly Journal of Experimental Psychology Section A*, *48*(1), 208–236.

Logan, G. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492–527.

Love, B., Medin, D., & Gureckis, T. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, *111*(2), 309–332.

Mack, M., Love, B., & Preston, A. (2016). Dynamic updating of hippocampal object representations reflects new conceptual knowledge. *Proceedings of the National Academy of Sciences*, *113*(46), 13203–13208.

Medin, D., & Shafer, M. (1978). Context theory of classification learning. *Psychological Review*, *85*(3), 207–238.

Medin, D., & Smith, E. (1981). Strategies and classification learning. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(4), 241–253.

Nosofsky, R. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*(1), 39–57.

Pearce, J. (1987). A model for stimulus generalization in pavlovian conditioning. *Psychological Review*, *94*(1), 61–73.

Peirce, J. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*.

Rabi, R., & Minda, J. (2014). Rule-based category learning in children: The role of age and executive functioning. *PlOs One*, *29*(9), e85316.

Rehder, B., & Hoffman, A. (2005a). Eyetracking and selective attention in category learning. *Cognitive Psychology*, *51*(1), 1-41.

Rehder, B., & Hoffman, A. (2005b). Thirty-something categorization results explained: Selective attention, eyetracking, and models of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(5), 811–829.

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423–1074.

Simon, H. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99–118.

Simons, D., & Chabris, C. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*(9), 1059–1074.

Smith, P., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neuroscience*, *27*(3), 161–168.

Turner, B., Forstmann, B., Love, B., Palermi, T., & van Maanen, L. (2017). Approaches to analysis in model-based cognitive neuroscience. *Journal of Mathematical Psychology*, *76*(3), 282–299.

Weichart, E., & Darby, K. (in press). Discovery-based learning induces elaborate processing: Evidence from an eye-tracking study.

Weichart, E., Evans, D., Galdo, M., Bahg, G., & Turner, B. (2022). Distributed neural systems support flexible attention updating during category learning. *Journal of Cognitive Neuroscience*, *34*(10), 1761–1779.

Weichart, E., Galdo, M., Sloutsky, V., & Turner, B. (2022). As within, so without; as above, so below: Common mechanisms can support between- and within-trial category learning dynamics. *Psychological Review*, *129*(5), 1104–1143.

Weichart, E., Unger, L., King, N., Sloutsky, V., & Turner, B. (2024). "The eyes are the window to the representation": Linking gaze to memory precision and decision weights in object discrimination tasks. *Psychological Review*, *131*(4), 1045–1067.