# Characterizing Shifts in Strategy in Active Function Learning

**Rebekah A. Gelpí**
(rebekah.gelpi@mail.utoronto.ca)
Department of Psychology
University of Toronto

**Caiqin Zhou**
(caiqin_zhou@brown.edu)
Dept. of Cognitive, Linguistic, and Psychological Sciences
Brown University

**Christopher G. Lucas**
(clucas2@inf.ed.ac.uk)
School of Informatics
University of Edinburgh

**Daphna Buchsbaum**
(daphna_buchsbaum@brown.edu)
Dept. of Cognitive, Linguistic, and Psychological Sciences
Brown University

**Abstract**

We investigate people's use of strategies for sampling data in an active learning task. In the spirit of resource-rational analysis, we argue that people may often use effective heuristics to guide sampling in lieu of more computationally expensive optimization strategies, but that when they encounter evidence that their heuristics are now ineffective they flexibly shift to new strategies. When the function family changed, participants quickly updated their beliefs about the likely function family on subsequent trials. By clustering participants' sampling behaviour, we show that people can employ varied sampling strategies, shifting strategies more often when encountering unusual function families that are more adversarial to generic sampling strategies. Not all new strategies improved participants' performance on a subsequent prediction task; nonetheless, people's ability to dynamically shift their active learning behaviour may help them understand the abstract features of complex relationships.

**Keywords:** active learning; function learning; sampling

## Introduction

Imagine that you have decided to plant a vegetable garden. Although many guides exist to help you determine what kinds of tools you might need (e.g., a sunny location, soil, fertilizer, a watering can), finding the most successful combination of these ingredients may require you to try out different amounts. Too little fertilizer and your plants may not grow, while too much fertilizer may cause your plants to overgrow and lead to less nutritious vegetables. Trying different amounts of fertilizer and waiting for the vegetables to grow is costly, so choosing the right amount of fertilizer requires adapting your strategy for how much to use as you learn more.

The notion that human learning benefits from being active or self-directed has a long history in the fields of psychology, cognitive science, and education (e.g., Bruner et al., 1966; Gopnik & Wellman, 2012; Hirsh-Pasek et al., 2009; Inhelder & Piaget, 1958; Prince, 2004) and analyses and implementations of computational models of active learning have emphasized the enhanced learning provided by active selection (Cohn et al., 1996; Gureckis & Markant, 2012; Settles, 2012).

At its most basic level, active learning involves selection, or sampling, of the data one would like to learn about. By allowing an individual to test their own hypotheses, active learning provides an opportunity to learn above and beyond characteristics inherent to the data points being sampled (Markant & Gureckis, 2014). Active learning has also been shown to enhance memory for the data (Markant et al., 2016).

These advantages notwithstanding, the opportunity to learn actively can be limited. To the extent that one can sample one's own data, we may still find ourselves constrained by a lack of time, opportunity, or the high amount of effort required to obtain the data; in this situation, we should be judicious about how we use our limited samples. To that end, people may seek out techniques or strategies that optimize their limited cognitive resources to ensure that their opportunities for learning are not wasted.

Nevertheless, optimal active learning strategies also come with costs of their own. Since the data must be actively selected, people must engage in potentially effortful search for the optimal points to sample. Computational models of optimal sampling strategies in active learning environments typically select points or causal interventions that aim to minimize uncertainty, maximize expected information gain, or both (e.g., Bramley et al., 2015; Coenen et al., 2015; Gureckis & Markant, 2012; Jones et al., 2018; Kruschke, 2008; Oaksford & Chater, 1994).

Traditional algorithmic implementations of optimal learning for statistical models in active learning situations often make the assumption that data sampling is myopic (Cohn et al., 1996; Roy & McCallum, 2001); that is, rather than picking the most jointly informative combination of data points to sample, each successive sampled point is treated as if it will be the last sample. For example, Cohn et al. (1996) developed a myopically optimal algorithm for sampling in supervised learning contexts, in order to minimize the number of samples necessary for effective learning—operationalized as minimizing expected future error. Recent work in cognitive science has also argued that human active learning is also best characterized by myopic strategies, rather than engaging in globally optimal search or advance planning (Bramley et al., 2015; Meder et al., 2019). Given the computational intractability of picking a sample by conditioning on all possible yet to be seen data, which would require estimation of all possible values of that data, relying on myopic sampling is understandable. However, it also means that myopic active learning strategies are potentially suboptimal, because they are only resolving uncertainty at the next sampled point, rather than considering in advance how multiple samples might jointly reduce uncertainty.

As a result, simple heuristic strategies for sampling data may sometimes be as good or better at maximizing information gain than myopic active learning policies. For example, low-cost strategies such as sampling a domain evenly (Gelpi et al., 2021) may be broadly applicable and effective strategies to reduce uncertainty at a rate comparable to more expensive myopically optimal policies. Nevertheless, generic heuristics can also be a poor match for some problems. For example, Box and Draper (1987) describe an example of testing the effect of two variables by collecting data at evenly-spaced values. However, for some kinds of underlying relationships, e.g., $y = \frac{1}{x}$, these heuristics are not an efficient way to gather information.

One way to address the tension between the prohibitive cost of fully computing even myopically optimal active learning strategies and the potential pitfalls of heuristic policies is to dynamically shift between different strategies as one encounters situations where one's current strategy is unsuccessful. For example, both children and adults often tackle problem-solving tasks by staying with a current strategy as long as it is successful, and switching to a new strategy if it fails, consistent with implementing a Win-Stay-Lose-Shift heuristic (e.g., Bonawitz et al., 2014; Chierchia et al., 2022; Giron et al., 2022; Yi et al., 2009). In addition, a body of work in children's mathematical problem-solving has argued that children develop and test multiple problem-solving strategies at once as they are learning (Shrager & Siegler, 1998; Siegler & Shipley, 1995).

People may therefore develop meta-reasoning for when different candidate strategies are appropriate for solving a given task, and shift between strategies given the task demands. Although heuristic strategies are not globally optimal, people may effectively deploy them in situations where they are most effective, maximizing the performance of the heuristics while reducing the cognitive cost of solving the task (Gigerenzer & Gaissmaier, 2011; Lieder & Griffiths, 2017). These findings fit within a larger body of work arguing that we engage in resource-rational computations to trade off the expected benefits of a particular strategy against the cognitive costs of engaging in it (Lieder & Griffiths, 2020).

Here, we use the general framework of function learning to evaluate shifts in people's active learning strategies. Our task has been formalized using the general framework of Gaussian processes (Lucas et al., 2015; Schulz et al., 2018), which capture important elements of learning and inference in functional domains. Wu et al. (2018) argue that generic Gaussian process priors shape sampling behaviour and generalization using a task with an explore-exploit dilemma.

However, the structure of more unconventional function families may—in addition to being *a priori* less likely—be much more challenging to represent in terms of a Gaussian process. For example, standard Gaussian processes cannot extrapolate functions with non-stationary kernels, such as sawtooth and step functions (see e.g., Wilson et al., 2015). In these cases, learning about the structure of unusual function families may involve inferences about data above and beyond what is expressible within a Gaussian process framework.

If people can use inductively learned properties of the family of functional relationships they are learning about in a given context, they may be able to adapt their sampling strategies to effectively learn about the exact parameterization of the function they are learning. This could involve changing from a general-purpose heuristic, such as sampling the domain equidistantly, to using a more cognitively costly strategy such as myopically maximizing information gain, perhaps because the cost of the information is outweighed by one's curiosity or the perceived value of the information (e.g., Dubey & Griffiths, 2020), or changing between multiple different heuristics when a heuristic is observed to be sub-optimal (Lieder & Griffiths, 2017).

To test whether and how people can exploit information about the class of functional relationships they are learning about to change their sampling strategies, we introduce an active function learning task in which participants encounter multiple different blocks of functions, some of which are designed to be relatively adversarial to a generic heuristic sampling policy: namely, equidistant sampling. We hypothesize that participants will use information about the function families they encounter to shift their sampling strategies away from generic "one-size-fits-all" heuristics more often when they are faced with function families for which these heuristics are poorly adapted. If participants are able to learn abstract features of the function families they anticipate encountering, this could result in the use of more well-adapted sampling strategies, and subsequently more accurate representations of the functions.

## Experimental Design

### Participants and Design

Ninety-seven adult participants ($M_{\text{age}}$ = 37.36, $SD_{\text{age}}$ = 12.08) were recruited through Prolific and paid £1.75 for their participation. One participant's data was not recorded due to a technical error, leaving 96 participants in the final analysis.

### Procedure

**Familiarization and Exposure** Participants were told they would play the role of a scientist on an alien planet whose goal was to map the distribution of an element in various (linear) regions through collecting measurements. They were also told that the detector for the element consumed a large amount of energy to generate each measurement, so they could only choose eight locations to measure the quantity of the element. Participants were randomly presented with one of two slightly different cover stories. They were either a geologist searching for underground Unobtanium (a rare earth element) along several coastlines, or a nuclear physicist searching for Ethereum (a radioactive element) along several fissures.

Participants practiced collecting measurements and graphing the distribution of the element in a region (i.e., a coastline or fissure) through a warm-up trial. Participants first saw an empty coordinate grid, with its x and y axes representing the

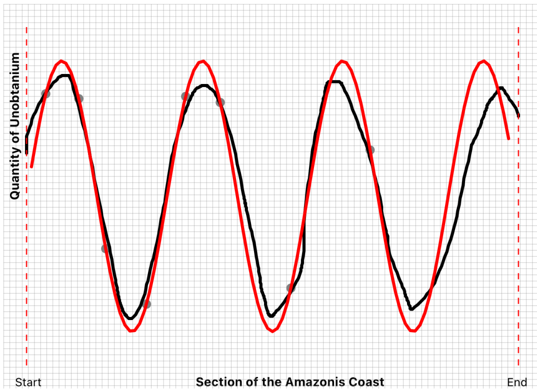


Figure 1: Example task setup for a trial where a participant had to learn one of the cyclic functions. Participants sample (grey dots) areas along the X axis, learning the corresponding Y value (quantity of element at the given location). After drawing 8 samples, participants must draw their estimate of the distribution (black line). After completing the drawing, participants are provided feedback in the form of the ground truth function (red line).

location along a linear region and the element quantity, respectively. To sample a location, participants could click on a spot on the x-axis, and a dot reflecting the quantity of the element would appear and remain on the grid. Participants then chose the next location to sample. After sampling the allocated number of locations (two in the warm-up trial for brevity), participants graphed the distribution of the element by free drawing a curve through the sampled dots, using their mouse. Lastly, participants learned about the true distribution pattern, which was it was superimposed on the participants' drawing in red.

**Experimental Task** Participants completed three task blocks. Each block included four mapping trials that showed different parameterizations of the same function family. As we aimed to create a task in which some trials would be adversarial to generic heuristic sampling policies, such as evenly sampling across the entire domain, two of the blocks involved functions with unusual patterns.

The first was a function we termed a "hurricane" function. This function contained a sharp, narrow local minimum at a particular point near the middle of the sampling region (Equation 1; see also Figure 5, right). The second, a cyclic function based on a sinusoidal curve of varying wavelengths and amplitudes (Equation 2; see also Figure 1), was similarly intended to pose a challenge to evenly-spaced samples. To provide a baseline against these unusual function families, one block consisted of a parameterization of a more common function: quadratic (Equation 3; see also Figure 5, left). Each function family had three free parameters, $\mu$, $\lambda$, and $\iota$, which were varied on each trial. The following equations were used for the three function blocks:

$$f(x) = \frac{1.25\lambda}{e^{-|100(x-\mu)|} + e^{-|\frac{12.5x}{\alpha}|} + e^{-|12.5(\alpha x-\alpha))|} + 1} + \iota$$
$$\text{where} \begin{cases} \alpha = 2m & \mu \leq 0.5 \\ \alpha = \frac{1}{2(1-m)} & \mu > 0.5 \end{cases} \tag{1}$$

$$f(x) = \lambda \sin(x-\mu) + \iota \tag{2}$$

$$f(x) = \lambda(x-\mu)^2 + \iota \tag{3}$$

As we anticipated that participants might be more likely to change their sampling strategy when moving between less and more adversarial functions, the order of block presentation was partially counterbalanced in 4 conditions: the quadratic block always appeared first or last, and the order of the two unusual blocks were randomized. Critically, participants were unaware of the blocking, as the trials were presented without breaks between blocks.

In each of the 12 trials, participants followed the same procedure as in the warm-up trial, in which they chose eight locations to sample one-by-one, drew the distribution graph, and learned about the true graph.

# Results

## Sampling Task

**Strategy Clustering** To characterize participants' sampling strategies, we employed Gaussian mixture modelling (GMM). GMMs assume that observed features are drawn from a set of $K$ Gaussian distributions, with the probability of an observation belonging to a particular distribution corresponding to:

$$p(\mathbf{x}) = \sum_{i=1}^{K} \alpha_i \cdot \mathcal{N}(x|\mu_i, \Sigma_i) \tag{4}$$

where $\mu_i$, $\Sigma_i$ are the mean and covariance matrix for the $i$th Gaussian, respectively, and $\alpha_i$ is the probability that $x$ belongs to the $i$th Gaussian.

We implemented a nonparametric GMM using the package `mclust` in R (Scrucca et al., 2016). This model used individual participants' ordered samples to simultaneously infer the cluster that best matched each participants' sampling behaviour and the number of total strategies. The analysis revealed that a model with 8 clusters of participant sampling strategies outperformed all others, including models with 9 ($\Delta$BIC = -127.4) and 7 ($\Delta$BIC = -219.7) clusters.

Subjectively, we observed that the 8 clusters tended to fall into five main categories (Figure 2). One category (clusters 2 and 7) involved participants evenly sampling the domain space in a monotonically increasing fashion, while a second group of clusters (3 and 5) involved omitting later points in order to sample earlier points more densely. A third category (clusters 6 and 8) involved participants selecting the lowest and highest values, and then sampling remaining points in the

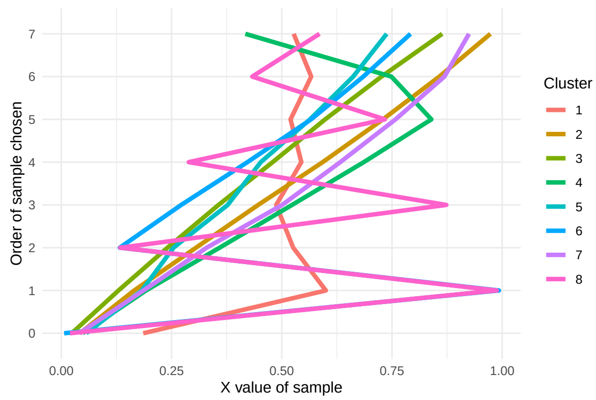


Figure 2: Sampling strategy clusters revealed by nonparametric GMM. The X axis indicates the location of the sample chosen, while the Y axis indicates the sample number, i.e., the first, second, third, fourth sample chosen, and so on. Most participants began by sampling relatively evenly along the X axis from left to right—sampling steadily higher X values on later samples—or sampling the lowest and highest points first before partitioning the remaining search space. Some participants (Cluster 4) began to sample evenly, but then returned to earlier points, possibly to re-sample areas of high uncertainty.

space between, either by partitioning the space into smaller areas or sampling evenly in the remaining area. Participants in Cluster 4 evenly sampled for most of the domain, before returning to sample earlier areas further. Finally, Cluster 1 included the remainder of participants who engaged in less consistent strategies that varied more broadly. Due to the similarity between these clusters, we collapse the clusters we identify into these 5 categories in subsequent analyses.

To understand how participants may have shifted their strategy across trials, we analyzed the frequency of strategy shifting between trials, using a generalized linear mixed model. We considered a shift to have occurred if a participant used a strategy from one strategy cluster on trial $t$, and a strategy from a different cluster on trial $t+1$. We considered the function family of the trial, the counterbalanced order in which function families were presented, and the presentation order of the trials in the model.

Change in strategy use was high overall, with participants switching strategies between trials 40.9% of the time. Shifting did not increase or decrease significantly across trials within blocks, $\chi^2(1) = 2.64, p = .10$, and including trial order as a factor did not improve model fit, ΔBIC = 6.5, so it was not included in the subsequent analysis.

Participants showed different patterns of strategy switching across function families, $\chi^2(2) = 11.25,\ p = .004$. As we predicted, participants were more likely to shift their strategy after trials in the cyclic block than in the quadratic block, $b = 0.46,\ SE = 0.18,\ z = 2.60,\ p = .025$. Participants were also less likely to switch overall when the quadratic function block was presented before the cyclic and hurricane blocks, $b = -0.87,\ SE = 0.34,\ z = -2.56,\ p = .01$ (Figure 3).

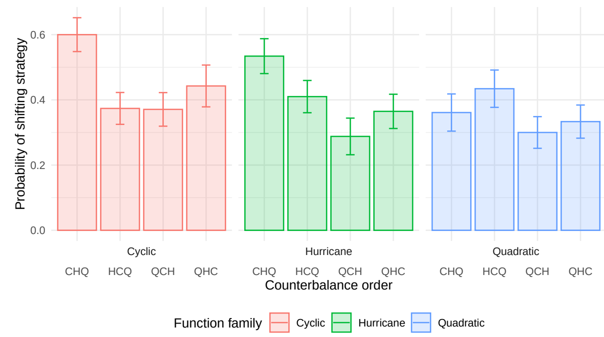


Figure 3: Probability of switching samples by counterbalance order; each bar on the $x$ axis reflects a particular order in which the function families were presented, e.g. "CHQ" meant that participants encountered the cyclic functions, then hurricane functions, then quadratic functions. Participants were most likely to switch sampling strategies when encountering the cyclic function (red bars), and were least likely to switch when the quadratic function was encountered first (two rightmost bars of each colour).

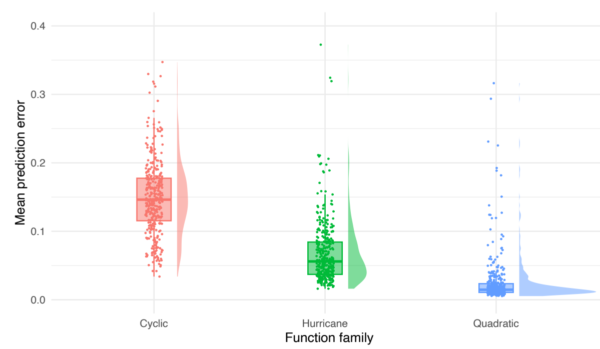


Figure 4: Average prediction error for the cyclic (red), hurricane (green) and quadratic (blue) function families.

Notably, including the random effects of individual participants in the model substantially improved fit, ΔBIC = 47.9, $\chi^2(1) = 54.78,\ p < .001$, suggesting that participants also varied significantly in the degree to which they shifted. Supporting this, some participants used the same strategy throughout—11 participants did not change their strategy throughout the experiment—while others shifted with much more frequency. Although no participants shifted after every single trial, 35 participants shifted on more than 50% of trials.

## Prediction Task

To measure participants' predictions as well as their sampling behaviour, we calculated the mean-squared error (MSE) of participants' prediction lines relative to the ground truth functions. As participants' errors were positively skewed (skew = 2.24), errors were log transformed before analysis.

Using a linear mixed-effects regression with random intercepts per participant, we tested participants' performance

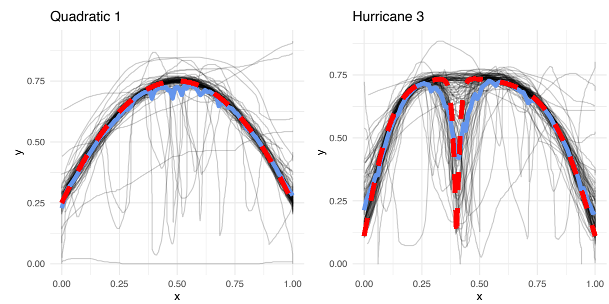


Figure 5: Individual prediction drawings (black lines) made by participants for two sample functions (left: quadratic, right: hurricane), with mean learned function (blue) and ground truth function (red dashed) overlaid. Participants who encountered a hurricane or cyclic function right before they encountered the first quadratic function (left) were more likely to predict the presence of a sinusoidal relationship. Conversely, some participants failed to sample in the region of the local minimum of the hurricane function (right), leading them to infer a function that was a closer fit for the quadratic function family.

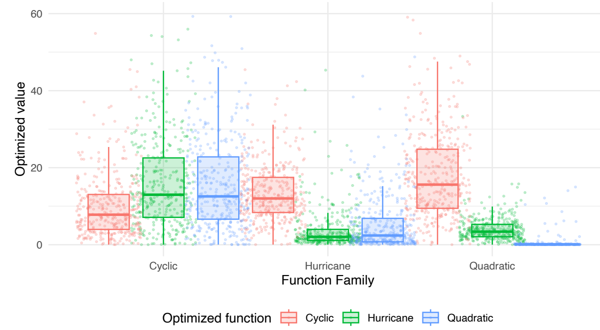


Figure 6: Optimization values for participant predictions by true function family. Lower values indicate predictions are a better fit for the optimized function family.

across individual trials and specific function families. As we were also interested in analyzing whether particular strategies were more effective in improving participants' predictions, we included the cluster characterizing the participants' sampling strategy for each trial as a predictor of their accuracy.

Participants' prediction accuracy varied significantly across functions, $F(2,1018) = 564.27,\ p < .001$. Specifically, participants were more inaccurate on the cyclic functions than both the hurricane functions ($b = 0.88,\ SE = 0.05,\ t(1006) = 16.59,\ p < .001$) and quadratic functions ($b = 1.99,\ SE = 0.05,\ t(1009) = 37.94,\ p < .001$; participants were also more inaccurate on the hurricane functions than the quadratic functions, $b = 1.11,\ SE = 0.05, t(1009) = 21.94, p < .001$. There were no significant improvements across all trials, $F(1,1002) = 1.11,\ p = .29$. Although sampling strategy did not directly predict performance on the task, $F(4,985) = 0.61,\ p = .65$, the analysis revealed a two-way interaction between the effect of strategy use and function family on prediction error, $F(8,1031) = 3.09,\ p = .002$.

To investigate this interaction, we computed pairwise contrasts comparing prediction accuracy between all pairs of strategies within each function family. We found that participants who sampled more densely in the earlier portion of the domain (Clusters 3 and 5) on quadratic functions had a lower average error than those whose sampling strategy was more diffuse (Cluster 1), $b = -0.29,\ SE = 0.09, t(1081) = -3.33,\ p = .008$. This may have been the result of more careful or vigilant sampling by participants who had previously encountered unusual or adversarial functions; by sampling points more closely together, such participants might have been able to more conclusively determine that the function did not belong to one of these adversarial functions. This could have contrasted with participants who, having previously seen an adversarial function, may have continued to infer an incorrect underlying function. For example, many participants interpolated a cyclic or hurricane function rather than quadratic, for the first function in the quadratic block (Figure 5, left).

Lastly, we aimed to understand whether participants, regardless of the accuracy of their predictions, were more likely to correctly draw predictions that matched the correct function family. For example, a participant who had inferred that the functions being learned were cyclic in nature could infer a cyclic function with an incorrect or out-of-sync wavelength, relative to the ground truth function. This could lead to high prediction errors, but nevertheless reflect real learning about the abstract function family.

To investigate this, we investigated whether participants' predictions were best fit by a function from the correct function family. We used Nelder-Mead optimization to find the best fitting parameter values of all participant predictions for all three function families, for each individual participant prediction line. Our implementation of this algorithm searched for parameter values of each function family (cyclic, hurricane, quadratic) that minimized the mean squared error (MSE) between the computed $y$ value and the participant's actual predictions. The optimization value for a particular function family reflected this difference after having searched for the best-fitting parameter values.

Thus, when participants' predictions included abstract characteristics of the correct function family (e.g., multiple waves for a cyclic function, a sharp local minimum for a hurricane function), the optimization algorithm typically resulted in a lower value for the correct function family, indicating a better fit, and higher values for the incorrect function families. Conversely, if a participant's prediction more closely resembled an incorrect function family on a particular trial (e.g., drawing a sharp local minimum for a quadratic function, or omitting this feature for a hurricane function, as in Figure 5), then the optimization value for an incorrect function family would be lower.

Consistent with participants learning the appropriate function family, optimization values were significantly lower for the true function family than for others across all three function families. A linear mixed-effects regression testing the effect of the true function family and the optimized function revealed significant interactions across all terms (all $b > -14.41$, all $t > 13.8$, all $p < .001$), suggesting that participants' drawings were overall a much better fit for the true function family (Figure 6).

Nevertheless, we found that on some functions, participants' drawings were more consistent with a different function. For example, in the third hurricane function (Figure 5, right), 48 participants were best fit by a quadratic rather than a hurricane function, and one participant was best fit by a cyclic function. This pattern was most common across the hurricane function family; when participants' samples did not reveal the trough in the centre of the function, this may have led some participants to infer a smooth function throughout.

## Discussion

Across three function families, including two highly unusual function families designed to be adversarial to simple heuristics, people deployed a number of different strategies and strategy types. Many of the strategies appear expressly designed for their ability to sample the entire domain at a relatively low cost, with both the evenly-spaced sampling strategies and the binary-partition strategies resulting in relatively equally spaced samples after all eight were chosen. Other participants reserved one or two remaining samples to resolve points of high uncertainty in the middle of the function.

Participants were most likely to change their strategies for the cyclic function, which also proved to be the most difficult; although not significantly so, we also observed that participants shifted their strategies slightly more often on trials when they were learning about the hurricane functions as well. This may be the result of "default" sampling strategies being relatively well-suited to the quadratic function, a relatively smooth function that could be captured by many common Gaussian process kernels (Duvenaud, 2014). The sharp discontinuity of the hurricane function and the short period of many of the cyclic functions may have posed a larger challenge to participants; in this case, the unexpected value of some of the collected samples might have prompted people to search for alternative strategies.

Perhaps due to the difficulty of the functions themselves, we did not find evidence that any strategies improved participants' subsequent performance on the adversarial functions. This may suggest that the task of searching for strategies that are well-suited for particular unusual function families may be too challenging to resolve in a short time frame, or that stronger incentives (e.g. providing larger rewards to participants with lower prediction errors) may be necessary to motivate participants to discover more effective strategies.

Similarly to Villagrá et al. (2018), we observed that across most trials, where participants' extrapolations were shaped by previously observed functions, most participants' interpolated predictions best fit by a parameterization from the same family, even when their predictions were inaccurate for the specific function. This suggests that participants made broadly appropriate inferences about the high-level abstract characteristics of the function.

A limitation of our approach with Gaussian mixture models is that strategies that may have involved a highly variable choice of points could not be definitively clustered. As GMMs rely on the assumption that clusters have a specified mean and variance, highly variable data points can be clustered into a single, highly diffuse cluster, a phenomenon we observed in our analysis. Conversely, clusters with very low variance can be clustered separately, even when they display high subjective similarity. The presence of grid lines on our tasks could have led different participants to consistently align their responses to slightly different points, leading to low variance in some sampled points. Abstract labels such as tick marks can improve participants' performance on passive function learning tasks (e.g. Kalish, 2013); such labels might have also facilitated learning in our active learning task by prompting participants to anchor their responses to a particular grid line, albeit resulting in multiple subjectively similar clusters. Nevertheless, we plan to consider alternative methods for clustering participants' sampling strategies.

For example, although we preserved the order of participants' samples, some participants may still have sampled evenly across the domain, but simply did so in an unusual order. Conversely, some participants may have lingered on or returned to resample a point of higher uncertainty, resulting in a final distribution of sampled points similar to the strategy we observed in Cluster 4, but in a different order. Alternately, samples could be considered based on their normalized distance from previous samples: if one could potentially sample a distant point, but samples a nearby point instead, this could reflect a belief that there is substantial uncertainty about a point in the vicinity of another. This could be helpful in some scenarios—for example, finding the trough of a hurricane function—but unhelpful in others. In future work, we plan to consider other contingencies between samples, such as their relationship to each other spatially independent of order or how they partition the remaining area to be sampled, to allow for a more nuanced understanding of people's ability to dynamically shift between strategies in active learning tasks.

## Acknowledgments

This research was supported by a CGS-D Scholarship to R.A.G. by the Social Sciences and Humanities Research Council of Canada.

## References

Bonawitz, E., Denison, S., Gopnik, A., & Griffiths, T. L. (2014). Win-stay, lose-sample: A simple sequential algorithm for approximating bayesian inference. *Cognitive Psychology*, *74*, 35–65.

Box, G. E., & Draper, N. R. (1987). *Empirical model-building and response surfaces.* John Wiley & Sons.

Bramley, N. R., Lagnado, D. A., & Speekenbrink, M. (2015). Conservative forgetful scholars: How people learn causal structure through sequences of interventions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 708.

Bruner, J. S., et al. (1966). *Toward a theory of instruction* (Vol. 59). Harvard University Press.

Chierchia, G., Soukupová, M., Kilford, E. J., Griffin, C., Leung, J., Palminteri, S., & Blakemore, S.-J. (2022). Confirmatory reinforcement learning changes with age during adolescence. *Developmental Science*, e13330.

Coenen, A., Rehder, B., & Gureckis, T. M. (2015). Strategies to intervene on causal systems are adaptively selected. *Cognitive Psychology*, *79*, 102–133.

Cohn, D. A., Ghahramani, Z., & Jordan, M. I. (1996). Active learning with statistical models. *Journal of Artificial Intelligence Research*, *4*, 129–145.

Dubey, R., & Griffiths, T. L. (2020). Reconciling novelty and complexity through a rational analysis of curiosity. *Psychological Review*, *127*(3), 455.

Duvenaud, D. (2014). *Automatic model construction with gaussian processes* (Doctoral dissertation). University of Cambridge.

Gelpi, R., Saxena, N., Lifchits, G., Buchsbaum, D., & Lucas, C. G. (2021). Sequential heuristics for active function learning. *Proceedings of the 19th International Conference on Cognitive Modeling*, 80–86.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*, 451–482.

Giron, A. P., Ciranka, S., Schulz, E., van den Bos, W., Ruggeri, A., Meder, B., & Wu, C. M. (2022). Developmental changes in learning resemble stochastic optimization.

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, bayesian learning mechanisms, and the theory theory. *Psychological Bulletin*, *138*(6), 1085.

Gureckis, T. M., & Markant, D. B. (2012). Self-directed learning: A cognitive and computational perspective. *Perspectives on Psychological Science*, *7*(5), 464–481.

Hirsh-Pasek, K., Golinkoff, R. M., Berk, L. E., & Singer, D. (2009). A mandate for playful learning in preschool: Applying the scientific evidence.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures*. Routledge.

Jones, A., Schulz, E., Meder, B., & Ruggeri, A. (2018). Active function learning. *Biorxiv*, 262394.

Kalish, M. L. (2013). Learning and extrapolating a periodic function. *Memory & Cognition*, *41*(6), 886–896.

Kruschke, J. K. (2008). Bayesian approaches to associative learning: From passive to active learning. *Learning & Behavior*, *36*(3), 210–226.

Lieder, F., & Griffiths, T. L. (2017). Strategy selection as rational metareasoning. *Psychological Review*, *124*(6), 762.

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, e1.

Lucas, C. G., Griffiths, T. L., Williams, J. J., & Kalish, M. L. (2015). A rational model of function learning. *Psychonomic Bulletin & Review*, *22*(5), 1193–1215.

Markant, D. B., & Gureckis, T. M. (2014). Is it better to select or to receive? Learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General*, *143*(1), 94.

Markant, D. B., Ruggeri, A., Gureckis, T. M., & Xu, F. (2016). Enhanced memory as a common effect of active learning. *Mind, Brain, and Education*, *10*(3), 142–152.

Meder, B., Nelson, J. D., Jones, M., & Ruggeri, A. (2019). Stepwise versus globally optimal search in children and adults. *Cognition*, *191*, 103965.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*(4), 608.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, *93*(3), 223–231.

Roy, N., & McCallum, A. (2001). Toward optimal active learning through monte carlo estimation of error reduction. *ICML, Williamstown*, *2*, 441–448.

Schulz, E., Speekenbrink, M., & Krause, A. (2018). A tutorial on gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of Mathematical Psychology*, *85*, 1–16.

Settles, B. (2012). Active learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, *6*(1), 1–114.

Shrager, J., & Siegler, R. S. (1998). Scads: A model of children's strategy choices and strategy discoveries. *Psychological science*, *9*(5), 405–410.

Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In *Developing cognitive competence: New approaches to process modeling* (pp. 31–76).

Villagrá, P. L., Preda, I., & Lucas, C. G. (2018). Data availability and function extrapolation. *Proceedings of 40th Annual Meeting of the Cognitive Science Society*.

Wilson, A. G., Dann, C., Lucas, C., & Xing, E. P. (2015). The human kernel. *Advances in Neural Information Processing Systems*, *28*.

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018). Generalization guides human exploration in vast decision spaces. *Nature Human Behaviour*, *2*(12), 915–924.

Yi, M. S., Steyvers, M., & Lee, M. (2009). Modeling human performance in restless bandits with particle filters. *The Journal of Problem Solving*, *2*(2), 5.