Learning functions actively

Angela Jones

Max Planck Institute for Human Development

Eric Schulz

Harvard University

Björn Meder

Max Planck Institute for Human Development

University of Erfurt

Azzurra Ruggeri

Max Planck Institute for Human Development

Technical University Munich

Author Note

Angela Jones, Max Planck Research Group iSearch, Max Planck Institute for Human Development.
Eric Schulz, Department of Psychology, Harvard University.
Azzurra Ruggeri, Max Planck Research Group iSearch, Max Planck Institute for Human Development and School of Education, Technical University, Munich.
Björn Meder, Max Planck Research Group iSearch, Max Planck Institute for Human Development and Department of Psychology, University of Erfurt.
AJ and ES contributed equally to this work and share first authorship.
Corresponding author: Eric Schulz, E-mail: ericschulz@fas.harvard.edu

Abstract

How do people actively learn functional rules, i.e. a mapping of continuous inputs onto a continuous output? We investigate information search behavior in a multiple-feature function learning task in which participants either actively select or passively receive observations. We find that participants benefit from actively selecting information, in particular in their function extrapolation performance. By introducing and comparing different models of active function learning, we find that participants are best described by a non-parametric function learning model that learns about both the underlying function and inputs that are likely to produce high outputs. These results enrich our understanding of active function learning in complex domains.

Keywords: active learning; function learning; self-directed learning; search

Learning functions actively

Introduction

In every day life, we often have to learn functional relationships between different variables. How far can I drive with my new electric vehicle when the battery is fully charged? How much breading do I need for the perfect schnitzel? How many rhetorical questions should I pose to make my introduction compelling?

Traditionally, function learning behavior has been studied in passive information-processing paradigms. In these paradigms, participants are sequentially confronted with continuous inputs, for example the height of a bar, followed by a continuous response, such as the height of another bar. Participants' task is to learn the underlying function relating inputs to outputs. Learning success can be tested, for instance, by asking participants to make predictions about the outcome variable given previously unobserved input values (i.e., function extrapolation). These experiments have focused on *passive* function learning, where the provided inputs are either randomly determined or selected by the researcher. However, we often actively decide for which inputs we want to observe the outcome in the real world. For instance, to learn about how far one can drive an electric vehicle with a full charge, one could measure the maximum distance covered when driving at different speeds. How can and should an agent actively learn about functional relations among continuous variables? And what models describe human active function learning best?

In this paper, we implement a multiple-feature function learning task to investigate how adult participants actively select inputs for which they want to observe the resulting output. Our behavioral results show that people's understanding of the underlying function is more accurate when learning actively compared to passively observing randomly selected inputs and corresponding output. The advantage of active over passive learning is particularly pronounced when participants have to make judgments about new inputs (i.e. extrapolation judgements). To better characterize participants' search behavior, we evaluate several combinations of function learning models and active sampling strategies. The best-performing model is a Gaussian Process function learning model combined with an Upper Confidence Bound sampling strategy. This indicates that participants learn functions in a flexible way and can adapt to different underlying functional rules instead of assuming only one particular rule (e.g., a linear function). Moreover, the fact that this model fits best when combined with an Upper Confidence Bound sampling strategy suggests that participants care about both learning the function and finding inputs that produce high outputs.

Function learning

Studies on function learning usually present participants with several input-output pairs (e.g., two bars of different heights), and then test their learning of the underlying function by asking them to infer the output for inputs that have not been observed before (e.g., to predict the height of a second bar, given the first), either included in the range of the training inputs (*interpolation* task; e.g., the height of the first bar is very similar to one previously observed) or outside the range of training inputs (*extrapolation* task; e.g., the height of the first bar is different from any previously observed).

Studies using interpolation tasks have shown that linear, increasing functions, are easier to learn than non-linear, decreasing functions (Brehmer, 1974; Brehmer, Alm, & Warg, 1985; Byun, 1996; McDaniel & Busemeyer, 2005). Studies using extrapolation tasks (DeLosh, Busemeyer, & McDaniel, 1997; McDaniel & Busemeyer, 2005) have demonstrated that participants tend to extrapolate in a linear fashion (Kalish, Lewandowsky, & Kruschke, 2004; Kwantes & Neal, 2006), even when the underlying function is nonlinear (DeLosh et al., 1997). However, people are capable of non-linear extrapolation, for example when the underlying function is cyclical (Bott & Heit, 2004). They therefore have a strong linear bias when learning functional relationships, but remain flexible learners, able to adapt to the type of function being learned.

Different theories have been developed to explain these findings and account for human function learning. The most prominent are similarity-based and rule-based theories. *Similarity-based* theories (e.g., Busemeyer, Byun, Delosh, & McDaniel, 1997; DeLosh et al., 1997) assume that people associate similar inputs with similar outputs, without learning an explicit representation of the underlying function. Similarity-based theories successfully capture some aspects of the observed performance, for instance that some functions are easier to learn than others. However, they fail to explain participants' systematic extrapolation patterns.

Rule-based theories (Carroll, 1963; Koh & Meyer, 1991) assume that participants learn explicit parametric representations, for example linear or power-law functions. Rule-based theories of function learning can successfully predict linear function extrapolation performance, for example by simply assuming that participants learn linear rules. However, they fail to explain that some rules are more difficult to interpolate than others (McDaniel & Busemeyer, 2005).

Hybrid models of function learning contain a similarity-based learning process that acts on explicitly-represented rules. They assume similarity-based interpolation, but extrapolate using simple linear models (Bott & Heit, 2004; Busemeyer et al., 1997; McDaniel & Busemeyer, 2005). Some hybrid models are able to capture both extrapolation and interpolation patterns (McDaniel, Dimperio, Griego, & Busemeyer, 2009). One such hybrid model has been proposed by Griffiths, Lucas, Williams, and Kalish (2009), who have put forward a rational theory of function learning based on Gaussian Process regression. Gaussian Process (GP) regression is a non-parametric method to perform Bayesian regression. Moreover, GP regression exhibits an inherent mathematical duality that makes it both a rule-based and a similarity-based model of function learning. Gaussian Processes generate predictions based on the similarity between different input values as expressed through a kernel, reminiscent of similarity-based models, and every kernel can be considered the result of performing a Bayesian regression, echoing rule-based models, as each kernel corresponds to a particular prior over functions. Lucas, Griffiths, Williams, and Kalish (2015) and Schulz, Tenenbaum, Duvenaud, Speekenbrink, and Gershman (2017) showed that GP regression can account for a wide range of human interpolation and extrapolation patterns.

Active learning

In the past years, a strong interest in human information search and active learning has emerged, with several studies finding beneficial effects of active compared to passive learning (see Coenen, Nelson, & Gureckis, 2018, for a review). For instance, Lagnado and Sloman (2004) found that learners who were given the opportunity to actively intervene on a causal system made more accurate inferences than passive learners who could not freely decide which information to obtain (also see Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003). In category learning, D. B. Markant and Gureckis (2014) found that active learners sampled more along the line of the category boundaries, thereby selecting more informative inputs, which in turn increased their categorization performance. Furthermore, recent studies have demonstrated that active control of the study experience leads to enhanced recognition memory in both children and adults (D. B. Markant, Ruggeri, Gureckis, & Xu, 2016; Ruggeri, Markant, Gureckis, Bretzke, & Xu, 2019), compared to conditions lacking this control, and that this benefits persists over time. Whether the opportunity to learn functions actively results in similar performance enhancements is an open question.

A critical question discussed in research on active learning is how to define the *usefulness* of pieces of information (see Nelson, 2005; Settles, 2010, for reviews). Different formal measures have been put forward, with the most prominent ones including the reduction in uncertainty measured via Shannon (1948) entropy (Lindley et al., 1956), the increase in the probability of making a correct classification decision (Nelson, McKenzie, Cottrell, & Sejnowski, 2010), and obtaining information for improving payoffs (Meder & Nelson, 2012; Wu, Schulz, Speekenbrink, Nelson, & Meder, 2018). Crupi, Nelson, Meder, Cevolani, and Tentori (2018) demonstrated that several of these measures can be unified into a coherent mathematical framework, thereby connecting formerly competing models of the value of information.

It is still unclear which measure best accounts for how human learners select information. For instance, probability gain consistently best described human search decisions in experienced-based category learning, where the goal is to maximize overall classification accuracy (Meder & Nelson, 2012; Nelson et al., 2010). In other tasks, however, information gain (expected reduction in Shannon entropy) is a better predictor for human search behavior (Bramley, Lagnado, & Speekenbrink, 2015; D. Markant & Gureckis, 2012; Meder, Nelson, Jones, & Ruggeri, 2019; Nelson, Divjak, Gudmundsdottir, Martignon, & Meder, 2014). Moreover, search behavior can vary depending on how information about the structure of the environment is communicated (Nelson et al., 2010; Wu, Meder, Filimon, & Nelson, 2017). These findings suggest that there might not be one single measure of usefulness that can account for behavior across all paradigms. Generally, it is still debated which measure of usefulness best describes active learning behavior in more complex domains such as function learning, which require combining a model of learning and a sampling strategy for evaluating and selecting queries (Bramley et al., 2015; Wu et al., 2018).

The present study: Active function learning

In this paper, we investigate for the first time the impact of active control over the function learning process on performance. To do that, we propose a novel experimental and theoretical framework for studying function learning that marries research on human function learning with recent advances in psychological theories of active learning.

Next, we describe the paradigm we developed to investigate active function learning. We then report analyses of the behavioral data, complemented by a computational analysis of participants' learning and search behavior, in which we compare different models of active function learning.

Experiment

Participants

Participants were 720 adults (mean age=36.34, $SD = 10$, 294 females), recruited from Amazon Mechanical Turk. Average task duration was 11.83 minutes ($SD = 10.71$). Participants received a participation fee of $2.00 and a bonus of up to $1.40 (mean bonus=$0.97, SD=$0.23). Study approval was obtained from the Max Planck Institute Ethical Review Board and participants gave informed consent prior to participating.

Materials and Procedure

Participants played a browser-based card game, in which each card showed a different monster with values for its three features ("friendly," "cheeky," and "funny"; see Figure 1). The instructed goal was to learn to predict the number of "magic fruits" monsters picked (criterion), based on their feature values (inputs). Participants were told they would receive a basic participation fee and a performance-dependent bonus.

Methods and design

Participants had to learn the underlying function from sequentially obtaining information on the criterion value for different monsters' feature values. The learning phase card set consisted of 27 cards generated by factorially combining all feature values between 2 and 4, such that participants observed only a restricted range of the function. All 27 cards were initially displayed with the feature values visible and the criterion value hidden (Figure 1).

Participants were randomly assigned to one of $2 \times 2 \times 4$ between-subject conditions, where we manipulated how people learned about the function (*learning type*, i.e. active or passive), the function underlying the relationship between the monsters' feature values and

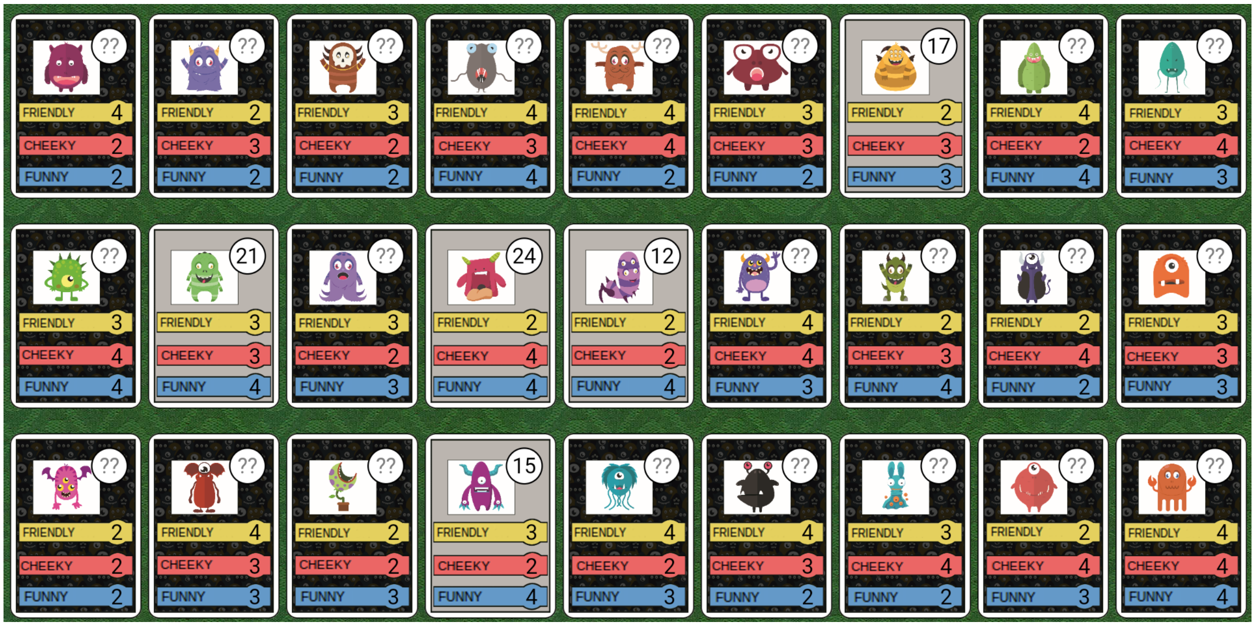

Figure 1. Screenshot of the multiple-feature function learning task (linear condition). Participants had to learn the relationship between the monsters' feature values ("friendly," "cheeky," and "funny") and the criterion (number of fruits picked, shown in the top right corner of selected cards). In this example, at this point in the game, the criterion value has been observed for five monster cards, each with a unique feature combination; the criterion values of the remaining cards are unknown.

the criterion (*function type*, i.e. linear or quadratic), and the amount of information participants received during learning (*number of observations*, between 0 and 27).

Learning type. In the active learning condition, participants could choose for which cards to observe the criterion value. Participants in the passive learning condition had to reveal the criterion value of randomly selected cards, one at a time, until the learning horizon was exhausted. Thus, participants in both conditions received the same number of data points, but while active learners could freely decide which data to observe, passive learners received randomly selected data points. Once revealed, the criterion value remained visible throughout the learning phase (Figure 1).

Function type. To test how a possible advantage of active learning might depend on the complexity of the underlying function participants were assigned to either a linear or a quadratic function. The linear function was

$$y = f(x) = 6x_1 + 3x_2 + x_3 - 10, \tag{1}$$

where y is the criterion value and x_1, x_2 and x_3 are the feature values. The weights for each feature were decaying, to ensure that participants had to attend to all features to achieve good performance and could not easily use simpler strategies, such as tallying.

The quadratic function was

$$y = f(x) = -x_1^2 + 3x_2 + x_3 + 21. \tag{2}$$

We set the weights of the different features such that the range of output values was similar[1] to that experienced by participants in the linear function condition. For all participants, the features were randomly assigned to x_1, x_2 or x_3.

Number of observations. To test how the amount of learning data impacts participants' function learning under passive vs. active learning regimes, we varied the length of the learning horizon. Participants observed either 0, 1, 5, 22, or 27 input-output pairs (cards) during the learning phase. The group with 0 observations was added to assess how participants would perform when they were not given the chance to gain any information about the underlying function.

Test phase

The test phase consisted of two tasks: a *criterion estimation* task and a *pair comparison* task (order counterbalanced across participants). No feedback was given during the test phases; the final bonus was determined based on participants' overall performance in the test phase (see below).

In the criterion estimation task, participants had to infer the criterion of a given monster (card) from its feature profile. This task included three types of trials: five recall

[1] The range of outputs for the learning set was 10-30 for the linear function and 13-33 for the quadratic function. The outputs for the extrapolation trials were the same values for both conditions and varied between 0 and 40.

trials, five interpolation trials, and eight extrapolation trials (18 cards in total; task order was randomized block-wise across participants). In the *recall* trials, the cards presented new monsters but with feature profiles for which participants had already observed the criterion in the learning phase.[2] In the *interpolation* trials, the cards presented new monsters with feature profiles corresponding to the five cards that had not been observed during the learning phase.[3] In the *extrapolation* trials, the cards showed new monsters with feature values of 1 or 5, representing a part of the function space that participants had not been trained on during the learning phase.

For each card, participants were asked to provide their criterion estimates by moving a slider horizontally between 0 and 40 (in increments of 1) until it reached the desired criterion value. Estimates within 5 of the true criterion value were rewarded with $0.06; estimates within 10 were rewarded with $0.04; estimates within 20 were rewarded with $0.02; estimates further than 20 away from the criterion were not rewarded.

In the *pair comparison* task, participants were shown eight card pairs whose feature values ranged between 1 and 5, such that these profiles contained both known and unknown feature values. For each pair, they had to decide which monster had gathered more fruits. This task assessed how well participants could judge the relative weights of each feature in the function they had to learn. For three of these trials, the card pairs were assembled such that one of the three features differed between cards, while the values for the other two features were held constant. For the other five trials, card pairs were assembled so that the value of the first, second or last feature outweighed the combined value of the two other features on each card, so that this feature was the main determinant of the number of fruits collected. Every correct selection was awarded with $0.04.

[2]Note that this means there were no recall trials for the group who only observed one card or no cards at all.

[3]Note that this means there were no interpolation trials for the group who observed all 27 cards during the learning phase.

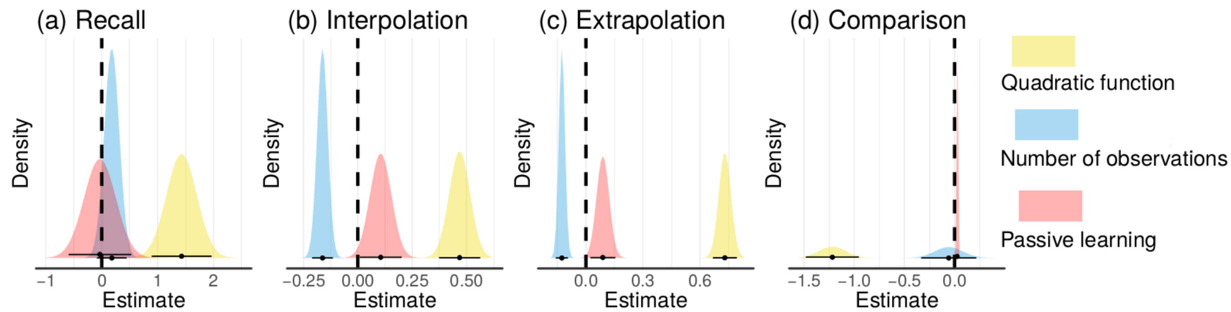

Figure 2. Posterior effects of conditions onto participants' performance. Since performance was measured by participants' absolute error, larger estimates indicate worse performance. *(a)* Effect on absolute recall error. *(b)* Effect on absolute interpolation error. **(c)** Effect on absolute extrapolation error. *(d)* Effect on pair comparison tasks. Distributions show posterior densities of effects when standardized regression estimates were entered into a Bayesian hierarchical model. Black dots indicate the posterior mean and error bars show the 95% highest posterior density interval.

Behavioral results

We calculated the effect of each manipulation by performing Bayesian multi-level regressions of the conditions' main effects[4] onto participants' absolute errors in the criterion estimation task (see Appendix for details and additional analyses using maximal random effects structures). For the learning type condition, we created an indicator variable that was set to 1 if learning was passive and 0 if it was active. Function type was coded as 1 if it was quadratic and 0 if it was linear. The number of available observations was entered as continuous variable into the regression. We z-standardized this variable to get a standardized estimate of its effect size. All regressions were performed using a random-intercept over participants and tests are reported based on a comparison with models not containing the tested variable (see Appendix for details).

Figure 2 shows the effects of the different manipulations onto participants' performance (their absolute estimation error) for the different tasks included in the test phase (Figure 2a). In the recall trials of the criterion estimation task, we found no evidence for either the horizon ($\beta = 0.18$, $HPD_{95} = [-0.09, 0.44]$, $BF = 0.9$) or the learning type ($\beta = 0.03$, $HPD_{95} = [-0.60, 0.53]$, $BF = 0.5$) being beneficial for participants'

[4]There was no evidence for interaction effects with all $BF < 1$.

performance. However, there was a strong effect of function type, with linear functions being easier to recall than quadratic functions ($\beta = 1.43$, $HPD_{95} = [0.89, 1.96]$, $BF > 100$).

In the interpolation trials, participants performed better when given a longer learning horizon ($\beta = -0.16$, $HPD_{95} = [-0.21, -0.11]$, $BF > 100$) and when learning a linear function ($\beta = -0.47$, $HPD_{95} = [-0.56, -0.37]$, $BF > 100$). We also found moderate evidence for an advantage of the active learning condition ($\beta = 0.11$, $HPD_{95} = [0.01, 0.20]$, $BF = 3$).

In the extrapolation trials, we found that participants were better given a longer learning horizon ($\beta = -0.13$, $HPD_{95} = [-0.16, -0.09]$, $BF > 100$) and a linear function ($\beta = 0.73$, $HPD_{95} = [0.67, 0.79]$, $BF > 100$). Additionally, we found evidence for an advantage of the active learning condition ($\beta = 0.09$, $HPD_{95} = [0.02, 0.15]$, $BF > 100$).

Since there was also a group of participants who did not observe any outputs before doing the criterion estimation task (labeled as missing value for the condition variable, since 0 observations are neither active nor passive), we also compared those participants to searchers who had actively sampled only 1 card. This showed that participants who had observed only 1 card already performed better than participants who observed no card at all in the interpolation trials ($\beta = -1.24$, $HPD_{95} = [-2.29, -0.20]$, $BF = 4$) but not in the extrapolation trials ($\beta = 0.33$, $HPD_{95} = [-1.06, 1.77]$, $BF = 0.08$). Thus, we found some evidence that even small amounts of information can improve participants' performance.

To assess participants' performance in the pair comparison task, we calculated the number of correct choices per participant and regressed the different conditions onto this number in a Bayesian linear regression without any random effects[5]. The results revealed that participants did not benefit from actively learning the function ($\beta = -0.06$, $HPD_{95} = [-0.33, 0.22]$, $BF = 0.4$), but performed better in the linear than in the quadratic condition ($\beta = -1.22$, $HPD_{95} = [-1.48, -0.95]$, $BF > 100$) and after having

[5]Note that we obtain the same results if we treat each response individually as in the analyses for participants' criterion estimation performance. Additionally, the results did not change when performing the analysis for the different types of pairs, i.e. pairs where one feature differed and pairs where two features outweighed another feature.

observed more cards ($\beta = 0.03$, $HPD_{95} = [0.01, 0.04]$, $BF = 79$).

Modeling active function learning

Active function learning requires an agent to build up a model of the underlying function and to sample the most useful inputs according to their beliefs. Thus, the building blocks for a computational analysis of active function learning are a model of participants' function learning and of their sampling strategies (e.g., to measure the usefulness of their selection, akin to information gain or probability gain), used to match the model's expectations onto informative actions. We compared two models of function learning, each combined with three different sampling strategies, to see which combination best accounted for participants' behavior.

Models for function learning

Linear Regression. A linear regression assumes that the outputs at time point t are a linear function of the inputs plus some added noise:

$$y_t = f(x_t) + \epsilon_i = \beta_0 + \sum_{i=1}^{k} \beta_i x_{t,i} + \epsilon_t, \tag{3}$$

where the noise term ϵ_t follows a normal distribution $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$ with mean 0 and variance σ_ϵ^2, β_0 is the intercept term and β_i are the slopes for the different features. Within a Bayesian framework, we can compute the posterior distribution over the weights and use this distribution to generate predictions about new observations, given their feature values (see Appendix for details).

Gaussian Process Regression. A GP regression is a non-parametric Bayesian way to model regression problems that can theoretically learn any stationary function by the means of Bayesian inference (Schulz, Speekenbrink, & Krause, 2018). If f is a function over input space $\mathcal{X}$ that maps to real-valued scalar outputs, then this function can be modeled as a random draw from a GP:

$$f \sim \mathcal{GP}(m, k). \tag{4}$$

Here, m is a mean function that is commonly set to 0 to simplify computations. The kernel function k specifies the covariance between outputs.

$$m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})] \tag{5}$$

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}\left[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))\right]. \tag{6}$$

The kernel function k encodes prior assumptions about the underlying function. A common choice is the *radial basis function* (RBF) kernel to model the underlying functional dependencies:

$$k_{\mathrm{RBF}}(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{||\mathbf{x} - \mathbf{x}'||^2}{\lambda}\right). \tag{7}$$

The length-scale λ governs the amount of correlation between inputs $\mathbf{x}$ and $\mathbf{x}'$. Importantly, whereas a linear regression makes explicit assumptions about the underlying functional form (i.e., linear), GP regression makes predictions for new observations based on their similarity to previously observed features and their outputs via the the kernel.

Active Sampling Strategies

Both function learning models generate predictions about the expected mean and associated uncertainties of outputs produced by different inputs. However, active function learning also requires a sampling strategy that maps models' predictions onto utilities to guide data selection. We compared three such sampling strategies.

Uncertainty sampling selects at each step the combination of feature values for which the predicted output is most uncertain, i.e., shows the highest predictive posterior standard deviation.

$$a_t(\mathbf{x}) = \arg\max \ \sigma_{t-1}(\mathbf{x}) \tag{8}$$

This strategy reduces the uncertainty over the input space quickly, and is mathematically related to focusing on the information gain of each observation (Krause, Singh, & Guestrin, 2008).

Mean sampling selects at each step the input values that currently promise to produce the highest output:

$$a_t(\mathbf{x}) = \arg\max \ \mu_{t-1}(\mathbf{x}) \tag{9}$$

This strategy does not attempt to learn efficiently but rather learns about the function serendipitously by trying to produce high outputs (i.e., here, higher numbers of magic fruit).

Finally, *upper confidence bound sampling* (UCB) tries to both reduce uncertainty and achieve high outcomes by sampling the input that currently shows the highest upper confidence bound

$$a_t(\mathbf{x}) = \arg\max \ \mu_{t-1}(\mathbf{x}) + \beta\sigma_{t-1}(\mathbf{x}), \tag{10}$$

where β is a free parameter governing the extent to which participants sample uncertain options. UCB sampling will, on average, converge to both high knowledge about the underlying function and sampling the highest possible outcomes. It has been found to describe human behavior well in exploration-exploitation paradigms where a global value function governs outcomes (Schulz, Wu, Ruggeri, & Meder, 2019; Wu et al., 2018).

Model Comparison Results

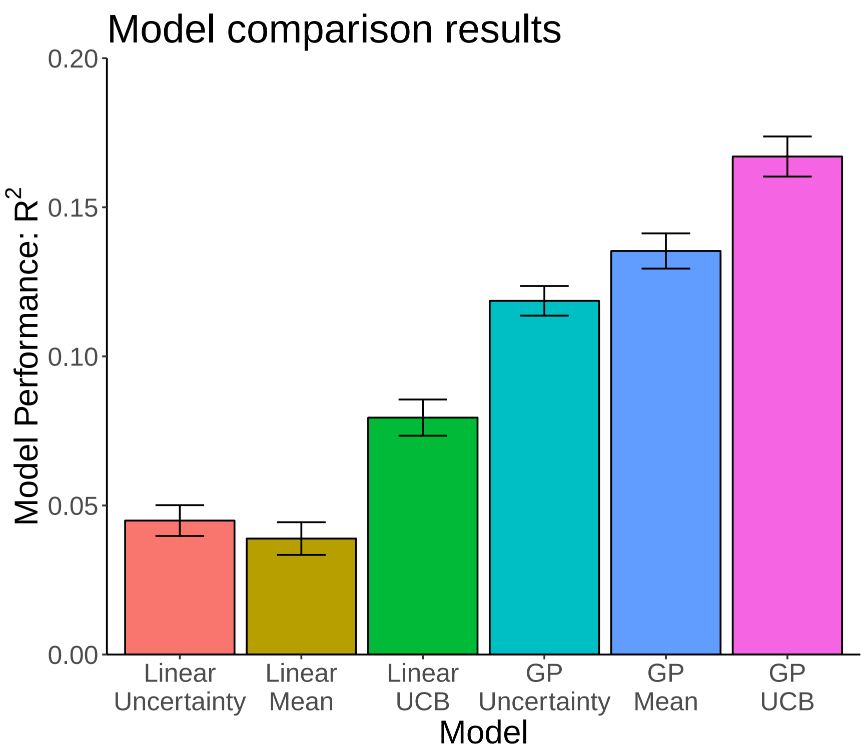

Figure 3. Model comparison results. Average descriptive performance (mean performance over subjects, R^2) for every model and sampling strategy combination. Error bars indicate the standard error of the mean.

We combined all of the above-described function learning models and sampling strategies and compared how well they described individual active learners' card choices in the learning phase(see Fig. 3). Since assessing model accuracy requires more than one data point, we only consider learners who observed more than one card (i.e., learning horizon>1, $N = 194$).

We used all models to generate means and uncertainties for trial $t+1$ by feeding them with participants' observations up until trial t, repeating this procedure for every participant over all trials. These means and uncertainties were then converted into utilities by different sampling strategies. Afterwards, we submitted the resulting utilities into a softmax function to convert them into choice probabilities

$$P(x) = \frac{\exp(a_t(\mathbf{x}))/\tau)}{\sum_{j=1}^{N} \exp(a_t(\mathbf{x})/\tau)} \tag{11}$$

where τ is a free temperature parameter estimated for each subject from the data. For each participant, we calculated a model's $\text{AIC}(\mathcal{M}) = -2\log(L(\mathcal{M})) + 2k$, where L is the models log-likelihood and k the number of free parameters. Afterwards, we standardized model performance using a pseudo-R^2 measure that compared each model to a random baseline (i.e., a model that randomly chooses input combinations (cards) to learn about the function):

$$R^2 = 1 - \frac{\text{AIC}_i}{\text{AIC}_{\text{random}}} \tag{12}$$

Figure 3 shows that the GP function learning model outperformed the linear model for each sampling strategy. The best overall model was a GP regression model combined with a UCB sampling strategy (GP-UCB), which showed an average performance of $R^2 = .17$ and best described 89 participants. The second best model was a Gaussian Process regression model combined with a mean sampling strategy, which showed an average performance of $R^2 = .13$ and best described 23 participants. This model performed significantly worse than the GP-UCB model, $t(193) = 9.38$, $p < .001$, $d = 0.67$, $BF > 100$. A Gaussian Process regression model combined with an uncertainty sampling strategy led to an average performance of $R^2 = .12$, and described 60 participants best. This model also performed worse than the GP-UCB model, $t(193) = 6.41$, $p < .001$, $d = 0.46$, $BF > 100$.

The linear regression model combined with an upper confidence bound sampling strategy (Lin-UCB) achieved an average performance of $R^2 = 0.08$, describing 15 participants best overall. This model also performed worse than the GP-UCB model, $t(193) = 11.04$, $p < .001$, $d = 0.79$, $BF > 100$. However, the Lin-UCB model performed better than a linear regression model combined with uncertainty sampling, ($t(193) = 8.16$, $p < .001$, $d = 0.59$, $BF > 100$, which had a mean performance of $R^2 = 0.04$ and described 3 participants best overall. The Lin-UCB model also performed better than the a linear regression model combined with a mean greedy sampling strategy, ($t(193) = 10.41$, $p < .001$, $d = 0.75$, $BF > 100$, which had a mean performance of $R^2 = .04$ and described 4 participants best overall. Interestingly, the GP-UCB model performed better than the Lin-UCB model even when the underlying function was linear, ($t(85) = 7.97$, $p < .001$, $d = 0.86$, $BF > 100$, indicating that participants applied a Bayesian similarity-based learning strategy even if the underlying function could have been learned by linear rules.

Finally, we analyzed the parameter estimates of the winning GP-UCB model. The mean of the softmax temperature parameter was estimated to be $\hat{\tau} = 0.25$, suggesting that participants' sampling behavior corresponded closely to selecting the highest value option, once they had taken into account both the mean and uncertainty associated with the inputs, and that they did not simply sampled options randomly. The mean estimate for the exploration parameter was $\hat{\beta} = 5.73$, showing that participants valued the reduction of uncertainty positively, trying to learn more about uncertain parts of the underlying function.

Discussion

We investigated participants' function learning behavior and performance in a task where they had to learn about a function relating three continuous features to a continuous criterion. Our behavioral results showed that participants struggled more when having to learn a nonlinear as compared to a linear function. This replicates previous results on human function learning using single features (Brehmer, 1974; Carroll, 1963). Participants' judgments were also more accurate when they could make more observations during the learning phase, in particular when making interpolation and extrapolation judgments. Most importantly, participants benefited from actively learning about the underlying function. This effect was particularly pronounced for function extrapolation in the criterion estimation task. Since extrapolation is known to be a particularly challenging aspect of function learning and has been termed the "sine qua non" of function learning (DeLosh et al., 1997), our findings highlight the advantages of actively learning about functional relations. Participants in the active learning condition did not, however, show increased performance in the recall and the paired comparison tasks. Because recent results have shown that active learning leads to improved memory of encountered exemplars (Ruggeri et al., 2019), our results suggest that the specific learning goal (e.g., learning a function versus memorizing objects) might mediate the benefits of active learning. Follow-up experiments could further investigate the conditions under which active control over the learning experience benefits participants' recognition memory and functional recall.

By comparing several models of active function learning, we found that a GP regression combined with an upper confidence bound sampling strategy explained participants' behavior best (Lucas et al., 2015; Schulz et al., 2017). This means that participants learned about the underlying function in a flexible and adaptive way; it also shows that participants cared about both reducing uncertainty and finding out which inputs produce high outputs. This finding mirrors previous results obtained in contextual and spatially-correlated bandit tasks (Schulz, Konstantinidis, & Speekenbrink, 2016; Wu et al., 2018). In particular, recent studies showed that participants solve the exploration-exploitation dilemma in reinforcement learning problems in a similar fashion (Schulz, Wu, et al., 2019; Wu et al., 2018). This hints at the possibility of a universal sampling strategy underlying both information search and the search for rewards. Participants may not easily be able to turn off the exploitation part of their sampling strategy as they normally encounter a mix of exploration and exploitation problems in real life (Schulz, Bhui, et al., 2019).

Our results enrich our understanding of active learning in complex domains and pave the way for future studies on active, self-directed function learning. Active function learning is particularly crucial to effectively navigate the world by making accurate inferences and predictions, as many real-world phenomena depend on functional relationships. Mastering this ability can boost learning more generally, especially from a developmental perspective. We know from previous studies that children and adults can differ in their ability to generalize from past observations and their tendency to seek out uncertainty (Schulz, Wu, et al., 2019). An important question for future research is therefore to identify and precisely characterize the emergence and developmental trajectories of active function learning.

References

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected papers of Hirotogu Akaike* (pp. 199–213). Springer.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.

Bott, L., & Heit, E. (2004). Nonmonotonic extrapolation in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 38–50.

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*, 708-731.

Brehmer, B. (1974). The effect of cue intercorrelation on interpersonal learning of probabilistic inference tasks. *Organizational Behavior and Human Performance*, *12*, 397–412.

Brehmer, B., Alm, H., & Warg, L.-E. (1985). Learning and hypothesis testing in probabilistic inference tasks. *Scandinavian Journal of Psychology*, *26*, 305–313.

Busemeyer, J. R., Byun, E., Delosh, E. L., & McDaniel, M. A. (1997). Learning functional relations based on experience with input-output pairs by humans and artificial neural networks. In K. Lamberts & D. R. Shanks (Eds.), *Studies in cognition. Knowledge, concepts and categories* (pp. 408–437). Cambridge, MA: The MIT Press.

Byun, E. (1996). *Interaction between prior knowledge and type of nonlinear relationship on function learning.* (Unpublished doctoral dissertation). ProQuest Information & Learning.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*, 1–28.

Carroll, J. D. (1963). *Functional learning: The learning of continuous functional mappings relating stimulus and response continua.* Educational Testing Service.

Coenen, A., Nelson, J. D., & Gureckis, T. M. (2018). Asking the right questions about the psychology of human inquiry: Nine open challenges. *Psychonomic Bulletin & Review*, *26*, 1548–1587.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Crupi, V., Nelson, J. D., Meder, B., Cevolani, G., & Tentori, K. (2018). Generalized information theory meets human cognition: Introducing a unified framework to model uncertainty and information search. *Cognitive Science*, *42*, 1410–1456.

DeLosh, E. L., Busemeyer, J. R., & McDaniel, M. A. (1997). Extrapolation: The sine qua non for abstraction in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 968–986.

Griffiths, T. L., Lucas, C., Williams, J., & Kalish, M. L. (2009). Modeling human function learning with gaussian processes. In *Advances in Neural Information Processing Systems* (pp. 553–560).

Gronau, Q. F., Sarafoglou, A., Matzke, D., Ly, A., Boehm, U., Marsman, M., . . . Steingroever, H. (2017). A tutorial on bridge sampling. *Journal of Mathematical Psychology*, *81*, 80–97.

Kalish, M. L., Lewandowsky, S., & Kruschke, J. K. (2004). Population of linear experts: knowledge partitioning and function learning. *Psychological Review*, *111*, 1072-1099.

Koh, K., & Meyer, D. E. (1991). Function learning: Induction of continuous stimulus-response relations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 811–836.

Krause, A., Singh, A., & Guestrin, C. (2008). Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research*, *9*, 235–284.

Kwantes, P. J., & Neal, A. (2006). Why people underestimate y when extrapolating in linear functions. *Journal of Experimental Psychology: Learning, Memory, and*

Cognition, *32*, 1019-1030.

Lagnado, D. A., & Sloman, S. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 856–876.

Lindley, D. V., et al. (1956). On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, *27*, 986–1005.

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

Markant, D., & Gureckis, T. M. (2012). Does the utility of information influence sampling behavior? In *Cogsci.*

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*, 94–122.

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.

McDaniel, M. A., & Busemeyer, J. R. (2005). The conceptual basis of function learning and extrapolation: Comparison of rule-based and associative-based models. *Psychonomic Bulletin & Review*, *12*, 24–42.

McDaniel, M. A., Dimperio, E., Griego, J. A., & Busemeyer, J. R. (2009). Predicting transfer performance: A comparison of competing function learning models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 173–195.

Meder, B., & Nelson, J. D. (2012). Information search with situation-specific reward functions. *Judgment and Decision Making*, *7*, 119–148.

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

Mullen, K. M., Ardia, D., Gil, D. L., Windover, D., & Cline, J. (2009). DEoptim: An R package for global optimization by differential evolution.

Nelson, J. D. (2005). Finding useful questions: on Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, *112*, 979–999.

Nelson, J. D., Divjak, B., Gudmundsdottir, G., Martignon, L. F., & Meder, B. (2014). Children's sequential information search is sensitive to environmental probabilities. *Cognition*, *130*(1), 74–80.

Nelson, J. D., McKenzie, C. R., Cottrell, G. W., & Sejnowski, T. J. (2010). Experience matters information acquisition optimizes probability gain. *Psychological Science*, *21*(7), 960–969.

Rasmussen, C. E., & Nickisch, H. (2010). Gaussian processes for machine learning (gpml) toolbox. *Journal of Machine Learning Research*, *11*(Nov), 3011–3015.

Rouder, J. N., & Morey, R. D. (2012). Default Bayes Factors for Model Selection in Regression. , *47*, 877–903.

Ruggeri, A., Markant, D. B., Gureckis, T. M., Bretzke, M., & Xu, F. (2019). Memory enhancements from active control of learning emerge across development. *Cognition*, *186*, 82–94.

Schulz, E., Bhui, R., Love, B. C., Brier, B., Todd, M. T., & Gershman, S. J. (2019). Structured, uncertainty-driven exploration in real-world consumer choice. *Proceedings of the National Academy of Sciences*, *116*(28), 13903–13908.

Schulz, E., Konstantinidis, E., & Speekenbrink, M. (2016). Putting bandits into context: How function learning supports decision making. *bioRxiv*, 081091.

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.

Schulz, E., Tenenbaum, J. B., Duvenaud, D., Speekenbrink, M., & Gershman, S. J. (2017). Compositional inductive biases in function learning. *Cognitive Psychology*, *99*, 44 - 79.

Schulz, E., Wu, C. M., Ruggeri, A., & Meder, B. (2019). Searching for rewards like a child means less generalization and more directed exploration. *Psychological Science*, *30*, 1561-1572.

Settles, B. (2010). Active learning literature survey. *University of Wisconsin, Madison*, *52*(55-66), 11.

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

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E.-J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, *27*, 453–489.

Wu, C. M., Meder, B., Filimon, F., & Nelson, J. D. (2017). Asking better questions: How presentation formats influence information search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(8), 1274-1297.

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*, 915-924.

Appendix A

Statistical tests

We report all statistical tests using both frequentist and Bayesian formats. We present frequentist tests alongside their effect sizes, i.e. Cohen's d (Cohen's d; Cohen, 1988). Bayesian statistics are expressed by their Bayes factors (BFs). A Bayes factor quantifies the likelihood of the data under the alternative hypothesis H_A compared to the likelihood of the data under the null hypothesis H_0. For example, a BF of 5 indicates that the data are 5 times more likely under H_A than under H_0; a BF of 0.2 indicates that the data are 5 times more likely under H_0 than under H_A. We apply the "default" Bayesian t-test as proposed by Rouder and Morey (2012) when comparing two independent groups, using a Jeffreys-Zellner-Siow prior with its scale set to $\sqrt{2}/2$. We approximate the Bayes factor between two different mixed-effects regressions by applying bridge sampling (Gronau et al., 2017). For the Bayesian regression models, we postulate that the β-coefficients for each participant β_i are drawn from a normal distribution

$$\beta_i \sim \mathcal{N}(\mu_\beta, \sigma_\beta^2), \tag{13}$$

estimating the group-level mean μ_β and variance over participants σ_β^2. We use the following priors on the group-level parameters:

$$\mu_\beta \sim \mathcal{N}(0, 100) \tag{14}$$

$$\sigma_\beta \sim \text{Half-Cauchy}(0, 100). \tag{15}$$

Posterior inference is accomplished by using Hamiltonian Monte Carlo and `brms` package (Bürkner, 2017).

Appendix B

Detailed model implementation

We provide further mathematical details for the two models of active function learning.

Bayesian linear regression

The first function learning model is linear regression. We adopt a Bayesian perspective on linear regression, performing posterior inference over the weights. We assume a Gaussian prior over the weights $p(\mathbf{w}) = \mathcal{N}(0, \mathbf{\Sigma})$ and a Gaussian likelihood $p(\mathbf{y}_t|\mathbf{X_t}, \mathbf{w}) = \mathcal{N}(\mathbf{X_t}^\top \mathbf{w}, \sigma_\epsilon^{\mathbf{2}}\mathbf{I})$. The resulting posterior is

$$\begin{aligned} p(\mathbf{w}|\mathbf{y}_t, \mathbf{X}_t) &\propto p(\mathbf{y}_t|\mathbf{X}_t, \mathbf{w})p(\mathbf{w}) \\ &= \mathcal{N}\left(\frac{1}{\sigma_\epsilon^2}\mathbf{A}_t^{-1}\mathbf{X}_t\mathbf{y}_t, \mathbf{A}_t^{-1}\right) \end{aligned} \tag{16}$$

where $\mathbf{A}_t = \mathbf{\Sigma}^{-1} + \sigma_\epsilon^{-2}\mathbf{X}_t\mathbf{X}_t^\top$. This posterior can be used to generate predictions about different option's means and uncertainties, which can then be fed into different sampling strategies.

Gaussian Process regression

Gaussian process regression assumes that the output y of a function f at input $\mathbf{x}$ can be written as

$$y = f(\mathbf{x}) + \epsilon \tag{17}$$

with $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. In Gaussian process regression, the function $f(\mathbf{x})$ is distributed as a Gaussian process:

$$f(\mathbf{x}) \sim \mathcal{GP}\left(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')\right). \tag{18}$$

A Gaussian process $\mathcal{GP}$ is a distribution over functions and is defined by a *mean* and a *kernel* function. The mean function $m(\mathbf{x})$ reflects the expected function value at input $\mathbf{x}$: The kernel function k specifies the covariance between outputs.

$$m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})] \tag{19}$$

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}\left[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))\right]. \tag{20}$$

Conditional on observed data $\mathcal{D} = \{\mathbf{x}_n, y_n\}_{n=1}^N$, where $y_n \sim \mathcal{N}(f(\mathbf{x}_n), \sigma^2)$ is a noise-corrupted draw from the latent function, the posterior predictive distribution for a new input $\mathbf{x}_*$ is Gaussian with mean and variance given by:

$$\mathbb{E}[f(\mathbf{x}_*)|\mathcal{D}] = \mathbf{k}_*^\top(\mathbf{K} + \sigma^2\mathbf{I})^{-1}\mathbf{y} \tag{21}$$

$$\mathbb{V}[f(\mathbf{x}_*)|\mathcal{D}] = k(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{k}_\star^\top(\mathbf{K} + \sigma^2\mathbf{I})^{-1}\mathbf{k}_*, \tag{22}$$

where $\mathbf{y} = [y_1, \ldots, y_N]^\top$, $\mathbf{K}$ is the $N \times N$ matrix of covariances evaluated at each pair of observed inputs, and $\mathbf{k}_* = [k(\mathbf{x}_1, \mathbf{x}_*), \ldots, k(\mathbf{x}_N, \mathbf{x}_*)]$ is the covariance between each observed input and the new input $\mathbf{x}_*$. This posterior distribution can also be used to derive predictions about each option's mean and uncertainties, which can be fed into different sampling strategies.

These predictions depend crucially on the chosen kernel function. The kernel function $k(\mathbf{x}, \mathbf{x}')$ models the dependence between the function values at different input points $\mathbf{x}$ and $\mathbf{x}'$:

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}\left[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))\right] \tag{23}$$

We use a radial basis function kernel, which is defined as

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\lambda^2}\right). \tag{24}$$

The radial basis function provides an expressive kernel to model smooth and stationary functions. The two hyper-parameters λ (called the length-scale) and σ_f^2 (the signal variance) can be varied to increase or reduce the a priori correlation between points and consequentially the variability of the resulting function. We chose those parameters by maximizing the log marginal likelihood. For a GP with hyper-parameters θ, this likelihood is given by:

$$\log p(y|X,\theta) := -\frac{1}{2}y^\top(K+\sigma_n^2 I)^{-1}y - \frac{1}{2}\log|K+\sigma_n^2 I| - \frac{n}{2}\log 2\pi. \quad (25)$$

where the dependence of K on θ is left implicit. We optimize the hyper-parameters using gradient-based optimization as implemented in the `GPML` toolbox (Rasmussen & Nickisch, 2010).

Model comparison

We use both models of learning, the Bayesian linear regression and Gaussian Process regression, to model participants' active learning. We fit the models to a the data a participant has seen time point t and then make predictions about each options mean and choices at $t+1$. We then feed these means and uncertainties into the different sampling strategies. The resulting utilities are then parsed into a softmax choice rule. We optimize the β of the UCB sampling strategy as well as the temperature parameter τ for each participant using the log-likelihood L. Participant-wise optimization is performed by using differential evolution as implemented in `DEOptim` (Mullen, Ardia, Gil, Windover, & Cline, 2009). The resulting log-likelihood can be used to calculate Akaike's Information Criterion (AIC, Akaike, 1998)

$$\text{AIC}(\mathcal{M}) = -2\log(L(\mathcal{M})) + 2k, \quad (26)$$

where k indicates the number of optimized parameters (two for any model using UCB sampling, and one otherwise). We standardize the resulting AIC using a pseudo-R^2 measure which compares each model's AIC to a random baseline (without parameters):

$$R^2 = 1 - \frac{\text{AIC}_i}{\text{AIC}_{\text{random}}}. \tag{27}$$

Appendix C

Behavioral results without extreme numbers of observations

Because our data set also contained participants with either 0 or 30 number of observations, we also analyzed our main behavioral effects after excluding these participants. We therefore combined participants' recall, interpolation, and extrapolation performance into one data set of participants' criterion estimation performance. All of the variables were coded as in our main analyses. Additionally, we removed participants with either 0 or 30 observations. We then again estimated a linear Bayesian multi-level model, regressing the number of observations, the function type, and the learning type onto participants' absolute error during the test trials. The results of this analysis showed strong effects for all three of our main manipulations. In particular, participants performed better given a longer learning horizon ($\beta = -0.08$, $HPD_{95} = [-0.10, -0.06]$, $BF > 100$) and a linear function ($\beta = 4.05$, $HPD_{95} = [3.58, 4.51]$, $BF > 100$). Importantly, there was also a strong effect of learning condition, with participants in the active learning condition performing better than participants in the passive condition ($\beta = 0.65$, $HPD_{95} = [0.12, 1.12]$, $BF > 100$).

Next, we also analyzed the behavioral results using maximal linear mixed-effects models (Barr, Levy, Scheepers, & Tily, 2013). Although the overall model comparison suggested that only including a random intercept over participants was enough, it is nonetheless sometimes recommended to keep the comparison maximally. Thus, we repeated the analysis from above, this time entering all of the individual factors as random effects into the null model and comparing them to an alternative model that additionally included the tested factor as a fixed effect as well. This analysis also revealed that all three manipulations had a significant effect onto participants' criterion estimation performance. Specifically, participants performed better given more observations ($\beta = -0.08$, $HPD_{95} = [-0.10, -0.05]$, $BF > 100$), a underlying linear function ($\beta = 4.09$, $HPD_{95} = [3.59, 4.55]$, $BF > 100$) as well as in the active learning condition ($\beta = 0.61$,

$HPD_{95} = [0.16, 1.08]$, $BF > 100$).