The Benefits of Staying Local: Bounded Adaptive Control by Modeling Linearly and Acting Locally

Jiong Xiao (4773019xj@gmail.com)[1] **& Neil R. Bramley**[2]
[1]University College London
[2]Department of Psychology, University of Edinburgh

Abstract

We present a novel experiment and perspective on human adaptive control. In our experiment, participants repeatedly adjust, zero, one or two variables with the goal of controlling a third variable, targeting a moving reward region. Across tasks, we vary the function that maps the control variables to the target variable, and use computational modeling to examine how participants represent and solve the tasks. While broadly successful, we find evidence suggesting that participants fall back on projecting a locally linear monotonic relationships, while also taking control actions that are conservative, preferring to adjust one variable rather than both relative to their previous action. We suggest that this allows for robust performance even when interacting with nonlinear non-monotonic functions.

Keywords: function learning; control; resource rationality; explore–exploit; incrementalism

Introduction

Humans face an ongoing challenge of adaptive control: We must act in goal-based ways in an environment whose structure is complex, changeable, and perennially under-determined by our past experience. Successful adaptive control requires combining exploration with function learning and model-based planning. As an everyday example, baking a successful cake requires a specific combination of ingredients and cooking steps. Baking a loaf of bread requires similar ingredients and steps, but in different quantities. Past experience informs us about particular set points within this combinatorial mapping of possible bakes, but this function is complex with various nonlinear and interacting elements—e.g. cakes may need more sugar but less flour than breads; but the yeast needed might depend on the quantity of sugar; and perceived sweetness may increase in one range, and saturate in another. Just as becoming a successful baker relies on gaining familiarity with this mapping, numerous everyday control problems, from medicine to motor control, admit complex functional forms that are, in principle, learnable and exploitable through experience, yet challenging to model and explore completely. Unfortunately, natural environments and action spaces are high-dimensional and open-ended, making normative solutions intractable in general, and successful control a matter of bounded rationality, balancing the benefits against the costs of model-based reasoning (Lieder and Griffiths, 2020).

We contribute to this research topic by examining how people control a variety of artificial systems involving two pseudo-continuous input variables and a pseudo-continuous output variable, across a sequence of time steps in which the rewarding target state continuously changes. To foreshadow, we find that participants are broadly successful in controlling these systems, yet their predictions reflect inductive biases toward linearity and monotonicity, while their control actions are strongly conservative. Participants favor actions that, relative to their previous action, adjust one rather than both variables and are more likely to hit targets that can be achieved by locally linear extrapolation of the underlying function. We use computational modeling to characterize participants' control as based on a rolling, locally linear approximation to the target function, supported by actions that are "rationally" conservative—since staying local reduces the discrepancy between the linear approximation and the true target function.

Control

A wealth of research has examined how humans and machines use experience to learn to maximize rewards (Dayan and Daw, 2008; Sutton and Barto, 2018) and achieve adaptive control (Osman and Speekenbrink, 2012). In well-studied tasks participants select from a discrete or continuous (Davis et al., 2018; Schulz et al., 2017) action space, acting in environments that may be contextual (Schulz, Konstantinidis, and Speekenbrink, 2018), unstable (Speekenbrink and Konstantinidis, 2015), or nonlinear (Berry and Broadbent, 1987; Wu et al., 2018). In these experiments, actions provide both rewards and information about the control problem, guiding future actions, leading to an explore–exploit tradeoff. Our contribution focuses on capturing human control behavior in a multivariate (specifically bi-variate) pseudo-continuous action space while systematically varying the underlying function. This setting allows us to focus on how human control strategies mitigate function and action space complexity.

Function Learning

Since every control problem has some ground-truth function, inferring and modeling this function is, in general, a precondition for optimal control (Conant and Ashby, 1970). However, the space of possible functions is infinite. Mitigating this, humans have strong inductive biases favoring simpler and more familiar functions such as those that are linear, or monotonic (Sanborn et al., 2010). People are quicker to learn and more accurate in extrapolating relationships that are truly linear (Brehmer, 1987) and monotonic (DeLosh et al., 1997). While well studied in prediction, the role of these inductive biases

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

remains underexplored in control settings. This is potentially because learning and control drive people's focus differently. A controller is orientated toward minimizing the error between the realized and the goal outcomes, while a learner focuses on the error between the predicted and realized outcomes (Gibson et al., 1997; Osman and Speekenbrink, 2012). Moreover, control tasks involve intervening on variables, while the function learning literature often studies learning from passive observations of input-output pairs. The current study directly links the affordances of a control task to the function induction problem, thereby investigating how inductive biases toward simple functions interact with strategic control.

Resource Rationality Meets Localism

Cognitive modeling frequently works "downward" from analysis of the ideal solution to a given task (Anderson, 2013; Marr, 1982). Humans, however, have finite time and computational resources to discover or implement ideal solutions, meaning behaviors that appear suboptimal or biased from a normative perspective, can be principled or even resource rational under closer scrutiny (Griffiths, 2020).

One approach to managing computational costs of model-based inference is to constrain the space of possibilities under consideration. Prioritizing simpler functional forms and considering a restricted subset of actions is one example. When learning actively about causal systems, people will often focus narrowly on pairs of variables at a time (Btesh et al., 2025; Fernbach and Sloman, 2009; Markant et al., 2016)—implicitly assuming they act independently—or test only one hypothesis at a time, generating local alternatives, and reusing proven action strategies (Bramley et al., 2017; Gong et al., 2025).

Vollmeyer et al. (1996) found that varying one variable at a time serves as an effective strategy for learning about dynamic systems as it facilitates falsification of alternative hypotheses (Kuhn and Brannock, 1977; Tschirgi, 1980). These suggest that people develop a local focus in learning about environments and infer their knowledge of systems from their interventions. Meanwhile, humans are capable of greater sophistication, with even young children being able to manipulate multiple variables when it is more efficient or necessary to do so (Bramley et al., 2022). Moreover, everyday control tasks, such as driving a car or baking a cake, inherently require the simultaneous manipulation of multiple variables.

Task

We examine adults in a bi-variate control task involving three pseudo-continuous variables and a variety of functional relationships. Participants must repeatedly set two variables in order to control a third, moving between shifting reward regions, and must do so while figuring out the functional relationships between the variables (see Figure 1).

Formal framework

At each time step $t \in 1 : T$, the controller takes action a setting $x_1, x_2 \in \mathbb{R}$. The generative function f maps the inputs

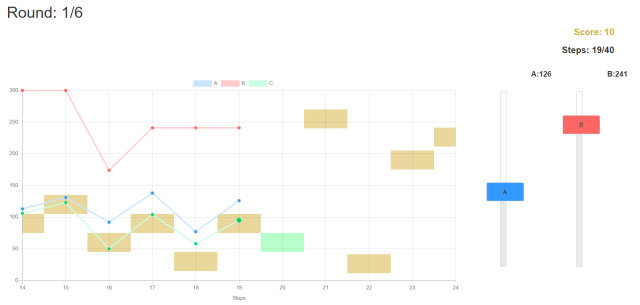


Figure 1: Task Interface: Left: Visualised history of control variables A (blue), B (red) and target variable C (green) control targets (yellow). At each time step, participants set A and B (red and blue sliders) and press spacebar to reveal C.

to the output $y \in \mathbb{R}$ at t such that $y^t = f(x_1^t, x_2^t)$.An action is rewarded if the resulting output falls within a tolerance band of the target value (±15 scale increments).

We consider 3×2 functional forms (see Figure 2a) such that the individual functional relationships are either:

1. *Linear:* Specifically of the form $g(x) = ax + b$
2. *Exponential:* Specifically $g(x) = a^x + b$
3. *Quadratic:* Specifically $g(x) = (ax + b)^2$

while the combination function is either:

1. *Additive*: $f(x_1, x_2) = g(x_1) + h(x_2)$, or
2. *Multiplicative*: $f(x_1, x_2) = g(x_1) \times h(x_2)$.

At each time, a reward can be achieved by a set of actions; we call this set the valid action set V.

Hypotheses

We had two core hypotheses:

1. Other things being equal, participants will favor control actions that alter fewer variables (i.e. 1 rather than 2), and move them less, relative to their previous position.
2. Errors controlling nonlinear and non-monotonic functions will be partially explained by assuming participants assume linearity and/or monotonicity.

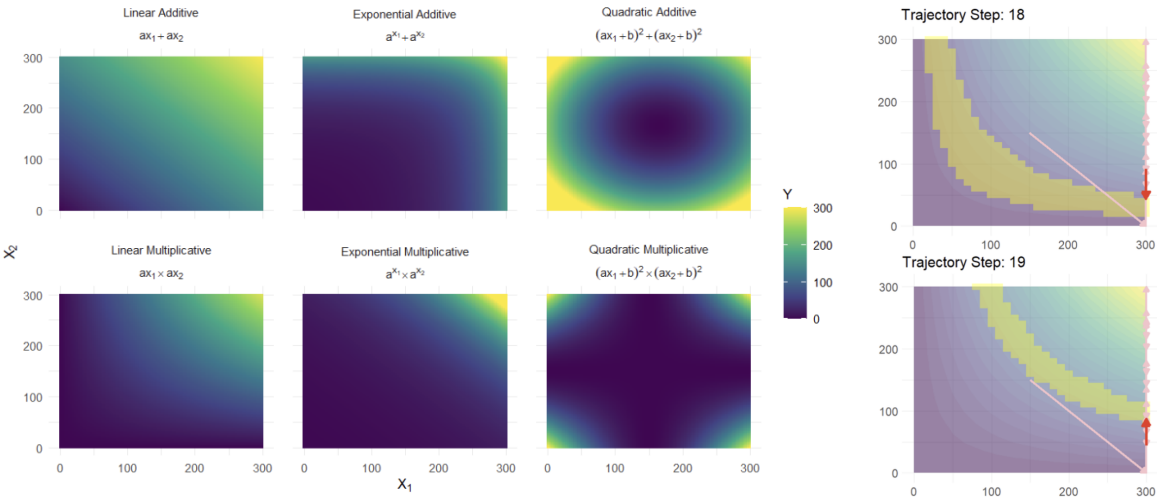
Models

We test our hypotheses using several computational models:

Normative Controller This model assumes knowledge of the ground truth function and selects indifferently among the set of valid control actions. Formally:

$$P(a^t) = \frac{\exp((\mathbb{I}_{V^t}[a^t]/\tau)}{\Sigma_{a' \in A} \exp((\mathbb{I}_{V^t}[a']/\tau)} \quad (1)$$

where $\mathbb{I}_{V^t}[a]$ is an indicator variable for membership of the valid action set V (i.e., the actions that will land y in the target region) at time t introduced above. Temperature parameter τ accounts for the noise in decisions.



(a) Action space.

(b) Example trajectory.

Figure 2: (a)Action spaces across six conditions. The x and y axes indicate values of input variables x_1 and x_2. Therefore, each coordinate corresponds to a possible action a setting $x_1, x_2 \in \mathbb{R}$, and the colour indicates the output value y. The colour indicates the value of y for a given action. We set a and b such that the full range of y fits exactly to the range of x_1 and x_2(b)Example participant trajectory over two steps in the Linear × Multiplicative condition. Valid action set V is highlighted in yellow. Red arrow source = the input on the previous time step, target = input on the current time step. In both cases, the participant alters only x_2, travelling the shortest distance to the new target. The pink lines depict the participant's past action trajectory.

The Conservative Controller As a refinement of the normative model, this model assumes participants tend to select from the set of successful actions that can be computed, but are *also* biased toward actions that stay closer to the previous setting. We parametrize this with two free cost parameters, biasing selection c_N penalizes for the number of variables $\in \{0, 1, 2\}$ that are moved to positions different from their position at $t-1$ and c_D penalizing the total "city block" distance the variables are moved $\in (0, 600)$. The model assigns a probability to a given action $a \in \mathbb{R}^n$ at time t as:

$$P(a^t) = \frac{\exp((\mathbb{I}_{V^t}[a^t] + c_N N_{a^t} + c_D D_{a^t})/\tau)}{\Sigma_{a' \in A} \exp((\mathbb{I}_{V^t}[a'] + c_N N_{a'} + c_D D_{a'})/\tau)} \quad (2)$$

and , N_a represents the number of input variables changed by a relative to a^{t-1}, and D_a represents the amount of changes made by a relative to a_{t-1} (the sum of absolute values of changes in input variables). Formally,$\Sigma_{n' \in N} |\Delta(x_{n'}^{t-1}, x_{n'}^{t})|$). The Conservative Controller embodies hypothesis 1.

The Conservative Controller assumes the valid actions are at least noisily available to the agent, implicitly that the participants have inferred the true functional form and can anticipate the outcomes of any action, so it always favours truly rewarding actions over truly non-rewarding ones.

Local Linear Controller We also consider a model variant that relaxes the strong assumption that the ground truth function is known. Local Linear Controller (LLC) instead infers an action set by assuming the true function has a linear functional form. It further re-estimates this function on a rolling basis, using the most recent 2 action–outcome pairs. This model thus embodies hypothesis 2 by incorporating three assumptions: locality, linearity and independence. First, the model is *temporally* local in the sense that it estimates the function from recent data, concretely inferring a slope s from $\Delta(a^{t-2}, a^{t-1})$ and $\Delta(y^{t-2}, y^{t-1})$. We assume the first action is selected at random, while for $t = 2 : T$, s of the nth input variable x_n is computed as follows:

$$s_n^t = \begin{cases} \frac{\Delta(y^{t-2}, y^{t-1})}{\Delta(x^{t-2}, x_n^{t-1})} \cdot \frac{|\Delta(x_n^{t-2}, x_n^{t-1})|}{\sum_{n' \in N} |\Delta(x_{n'}^{t-2}, x_{n'}^{t-1})|} & \text{if } \Delta(x_n^{t-2}, x_n^{t-1}) \neq 0 \\ s_n^{t-2} & \text{otherwise} \end{cases} \quad (3)$$

Formally, $\frac{\Delta(y^{t-2}, y^{t-1})}{\Delta(x_n^{t-2}, x_n^{t-1})}$ captures the slope of the change of y in respect to x_n. In cases where multiple variables change, the controller must determine how much of the change each variable is responsible for to determine the individual effects. We make the simplifying assumption that the controller assigns responsibility in proportion to the changes in the individual variables. This is instantiated by the part $\frac{|\Delta(x_n^{t-2}, x_n^{t-1})|}{\sum_{n' \in N} |\Delta(x_{n'}^{t-2}, x_{n'}^{t-1})|}$. When a variable is not changed at the time step, the controller's estimate of the unchanged variable's effect s remains unchanged from the last time step.

The LLC's prediction for the change in the outcome y given a^t is computed as:

$$\Delta(y^{t-1}, y^t) = \Sigma_{n \in N} \Delta(x_n^{t-1}, x_n^t) \times s_n^t \quad (4)$$

Based on the above, Local Linear Controller has an action set L that the controller predicts will land y in the target region. The action set L coincides with V exactly in the case that the functional form is, in fact, independent and linear. In other

scenarios, L may overlap partially or diverge entirely from V. Crucially, L depends on the most recent action taken.

Furthermore, given the desired difference $\Delta(y^{t-1}, y^t)$, without iterating over all possible actions and deciding which belongs to the set L, the controller computes the action it needs to take by focusing on one variable. Specifically, by fixing other variables, the desired change for a specific variable x_n can be computed as:

$$\Delta(x_n^{t-1}, x_n^t) = \frac{\Delta(y^{t-1}, y^t)}{s_n^t} \tag{5}$$

If changing one variable is insufficient to achieve the desired outcome, the controller adjusts the other variable to compensate for the remaining difference. This captures the preference to change the fewest number of variables within the environment because focusing on changing one variable allows for quick computation using Equation 5. Essentially, while LLC considers all actions within L as rewarding actions, it tends to compute actions that involve fewer variables.

Experiment

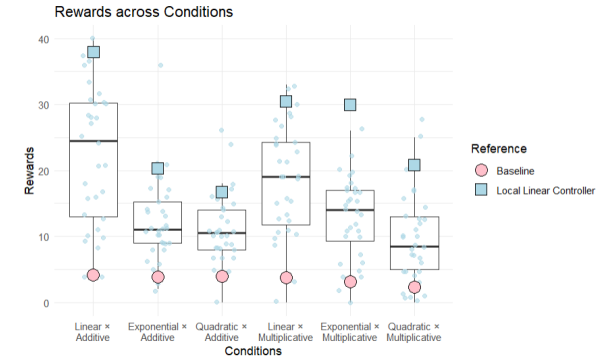
Participants

We recruited thirty-six participants (15 female, age M±SD=20.4±2.1) from the subject pool at University of Edinburgh for course credit. We excluded 2 participants who held both sliders constant throughout at least one round, and 2 whose data were improperly recorded. Ethical approval was granted by Psychology Research Ethics Committee at the University of Edinburgh. The task took 38.3±20.7 minutes.

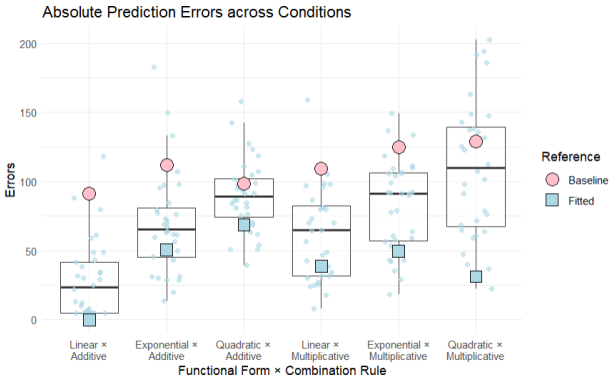
Design and Materials We used a 3 *Functional Form* × 2 *Combination Rule* within-subject design (see Figure 2a). In each condition, the participants interacted with the interface depicted in Figure 1. The interface displayed a reactive line chart with three colored lines corresponding to x_1, x_2, and y labeled "A", "B", and "C" for participants, and color-coded in red, blue and green. The chart also showed the target locations, which are highlighted in yellow and vary across time steps. In the first 20 time steps, reward regions were handpicked to be non-overlapping and to alternate between upward and downward shifts (mean adjacent difference ≈ 45; see osf.io/3e5pr). The last 20 regions were defined by a sinusoidal function that produced larger, less regular gaps (mean adjacent difference ≈ 149). There were two movable colored sliders labeled A and B (corresponding to x_1 and x_2). Participants interacted with the two sliders to control the target variable C (corresponding to y). The values of the input variables were bounded between 0 and 300. They may choose to move zero(i.e., taking no action), one or both sliders up or down at each time step. Participants press the spacebar to proceed to the next time step. If no slider is moved at the previous time step, the sliders stay at the same location.

Procedure

Participants first completed instructions and comprehension checks before facing the 6 trials in random order. Each trial consisted of 40 time steps. Participants earned a point at the time step where they land the target variable C in the target region.[1] After each trial, participants were asked to make five predictions without feedback to assess their beliefs about the functional form. The prediction phase displayed frozen slider values for A and B, and participants had to set a third slider to guess the value of C. Finally, participants provided text feedback and difficulty and engagement scores.



(a) Reward by trial type



(b) Prediction Errors by Trial type.

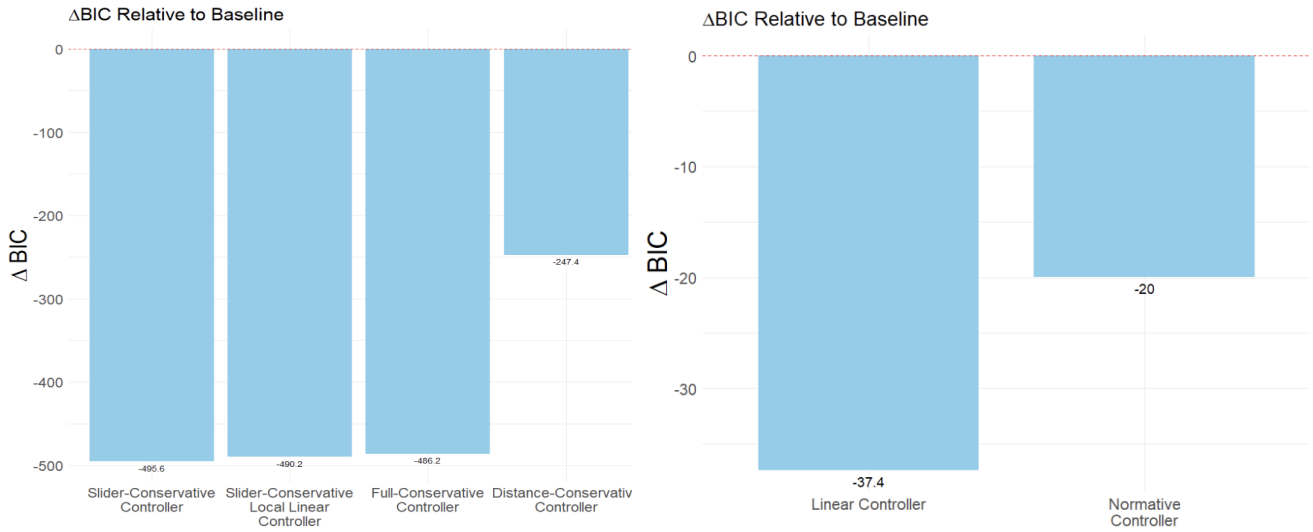
Figure 3: Reward and prediction error by trial type. See osf.io/3e5pr for full data.

Results

We first examine performance by trial type. Participants were reasonably successful in all trial types, scoring between 9.03 ± 6.7 for the *Quadratic* × *Multiplicative* function and 23.3 ± 10.9 for the *Linear* × *Additive* function, in all cases outperforming random responding(see Figure 3a). To capture condition differences, we fit a linear mixed-effects model predicting reward rate as a function of function type and combination type, with random intercepts for participants. Participants scored higher in Linear (β = 5.99, $t(156)$=4.28, $p < .001$), Monotonic (β = 2.27, $t(156)$=2.14, $p = .034$), and Additive (β = 4.31, $t(156)$=2.88, $p = .004$) trials. These suggest that inductive biases may play a role in control.

For the prediction phase, we computed the mean abso-

[1] Maximum score was therefore 40 points per trial or 240 overall.



(a) BICs for Models Encoding for Cost parameters.

(b) BICs for Models Inferring Valid Actions.

Figure 4: Model comparison results.

lute differences between participants' predictions and ground truths by trial type (see Figure 3b). We again fit a linear mixed-effects model predicting prediction errors from the properties of the functional forms and combination rules, with random intercepts for participants. Predictions deviated less from the true values in conditions where generative functions were linear (β = -109.92, $t(156)$=-2.93, p = .003), monotonic (β = -116.31, $t(156)$=-4.10, $p < .001$), and additive (β = -82.75, $t(156)$=2.92, p = .004). These demonstrate the presence of inductive biases in the prediction task. We fit a linear regression model with the intercept (i.e., the output variable's value when the input variables are 0) set to 0 for each functional form. We then use the fitted regressions to predict values based on the input values presented in the prediction phase. The human errors are not significantly correlated with the errors of fitted regressions (r = .63, p = 0.18). This shows that people may have formed a representation of the task environment that cannot be captured by assuming only additive linear functions with an intercept of 0.

We next assess our hypothesis that participants' actions were conservative, in the sense of changing a minimal number of variables. Binomial tests at the individual level reveal that every participant is classified as preferring to change one rather than both sliders relative to the previous time step.[2]

Model-based Analyses

We next use our computational models to further investigate our hypotheses. We start with assessing the qualitative prediction of Local Linear Controller. Specifically, we simulated LLC 100 times for each condition to compare its performance by trial type against humans' (see Figure 3a), finding it well correlated with human data ($r(4) = .87, p = .025$). Concretely, like participants, LLC performs better when the functional relationship is linear and monotonic.

We next fit our Conservative and Local Linear Controller models to participants' actions. We also included a Baseline model that selects actions at random. We first ask how likely participants select actions from a locally inferred action set V, and then ask whether conservative costs improve model fit.

We fit all models by minimizing negative log-likelihood using R's optim function (implementing Nelder-Mead). Given the prohibitively large action space (90601 actions), we opted to coarse-grain the model predictions to $\times$5-unit bins. To account for noise and natural imprecision in participants' actions, we used a Gaussian filter smoothing with a σ value of 2 to "blur" the model likelihoods such that actions near likely actions become likelier.

We computed BICs to compare the model fits while accounting for the number of parameters (Schwarz, 1978) and present ΔBIC for the Conservative, LLC and Normative Controllers relative to the baseline model in Figure 4.

First, we compare Normative and LLC against the baseline. As noted above, the two controllers have different "valid" action sets, V and L. Here, how likely the model takes an action a solely depends on whether a belongs to the model's "valid" action set. In other words, we are interested in how likely people are to select actions from the action set L and the action set V. This allows us to further assess whether LLC captures people's error patterns via its simplifying assumptions. We directly use Equation 1 for fitting the Normative Controller. For Local Linear Controller, we replaced the indicator variable

[2]There are 90601 possible actions people can take at each time step, out of which 600 (0.66 %) involve moving one slider. If the proportion of the one-slider action taken is significantly higher than the expected proportion, the participant is classified as preferring moving a single slider.

$\mathbb{I}_{V^t}[a]$ in Equation 1 with $\mathbb{I}_{L^t}[a]$, such that:

$$P(a^t) = \frac{\exp(\mathbb{I}_{L^t}[a^t]/\tau)}{\Sigma_{a' \in A} \exp(\mathbb{I}_{L^t}[a']/\tau)} \quad (6)$$

where $\mathbb{I}_L[a]$ indicates the membership for L. Both controllers outperform the baseline. LLC has a lower BIC than both other Controllers. This suggests that LLC captures people's error patterns to some extent(see Figure 4b).

Next, we assessed our hypothesis that people are conservative. As shown in Equation 2, the Conservative Controller has two additional parameters, C_N, which minimizes the number of sliders moved, and C_D, which minimizes the city-block distance of movements. We evaluated here ablated variants of this full model, including the slider-conservative (without the parameter C_D) and distance-conservative controllers (i.e., without the parameter C_N). To note, as shown in Equation 2, the default Conservative Controller is constructed based on the Normative Controller. Therefore, it relies on the simplifying assumption that the controller distinguishes between ground-truth valid actions and invalid actions. While this assumption does not account for systematic errors in human decisions, it reflects aspects of the constraints imposed by the task's goal.

In addition, to understand the full behaviour of LLC, we make LLC slider-conservative by including the cost term c_N in Equation 6. This increases the likelihood of actions that are consistent with the tendency of LLC to compute single-slider actions while admitting other actions from L.

Crucially, the default Slider-Conservative Controller performs the best, supporting our first hypothesis that people prefer to change minimal numbers of variables (see Figure 4a). Surprisingly, the Slider-Conservative Local Linear Controller ($\Delta BIC = -490.2$) has a higher BIC than the default Slider-Conservative Controller ($\Delta BIC = -495.6$). This suggests that Local Linear Controller do not fully capture the complexity of people's internal models.

Discussion

By introducing a bi-variate shifting-target control task, we show that people prefer to control with one variable at a time through conservative models. The fit between the Distance-Conservative Controller and human data suggests that people prefer making minimal changes. However, the model considering both the distance and the number of sliders did not fit the human data better than the model only considering the number of sliders manipulated. This could be because people aimed to land the variable in the center of the target region to minimize chance of undershooting, which requires more changes than minimally necessary for gaining the reward.

Crucially, our task presents people with a pseudo-continuous action space where actions are not tied to fixed rewards. These properties distinguish our task from other well-established sequential decision-making tasks, such as contextual bandit problems (Schulz, Konstantinidis, and Speekenbrink, 2018; Wu et al., 2018), where people seek to select the most rewarding actions from a bounded set of discrete decisions. Gaussian process function learning (Schulz, Speekenbrink, and Krause, 2018), together with the upper confidence bound strategy (Auer, 2002), has been found predictive of people's exploratory behavior in such contexts. In our task, the shifting requirement to achieve different outputs and the relatively unbounded action space may push people to acquire a model or an algorithm that accurately and efficiently computes actions. In the future, we may consider how a Gaussian process model or other function learning models may support exploration (e.g., selecting informative actions) in a shifting-target control task like ours.

Our results show that participants' inductive biases for simple functional forms interact with performance: participants achieved higher rewards and made more accurate predictions on linear, monotonic, and additive problems. One interpretation is that such environments facilitate the acquisition of internal models of the system, consistent with model-based control. Alternatively, participants may rely on linear approximations regardless. Piecewise linear approximation is a common approach to modeling complex functions in engineering (Dunham, 1986), and local approximations can be formed on the fly, especially if control actions are also auto-regressive. The limit of this strategy might be that pure gradient-based adjustment does not form an explicit representation of the underlying generative function, resembling generic model-free optimization control strategies. However, participants' ability to respond sensibly in the prediction phase, and the overall mixed pattern of behavioral and modeling results suggest an interplay between model-based representations and simpler local heuristics. Future work could investigate how interventions support the transition from local approximations to taking actions based on more structured internal models. Meanwhile, it should be noted that our prediction phase included a limited set of probes. Therefore, it is not a comprehensive test of internal representations. Future work could employ alternatives, such as asking participants to identify the potential function from a set of functions visualized in plots.

Our conservative model assumed control was generically costly, favoring least-action solutions. Future studies should clarify the nature of these costs. For example, conservative action might reflect a rational strategy for dealing with an incompletely explored world. Reflecting the control-of-variables principle (Kuhn & Brannock, 1977), making more changes makes input-output relationships difficult to learn by making it unclear how each input variable influences the outcome. Moreover, isolating one dimension dramatically reduces the action space. Thus, minimal interventions and maintaining local focus may be favored because it is cognitively feasible and aligns with computational rationality. More broadly, extending the investigation of this framework to richer multidimensional tasks could illuminate when attending to additional dimensions, or switching between them, becomes adaptive.

AI Statements

ChatGPT was used for proofreading an early draft and debugging code for the experiment.

References

Anderson, J. R. (2013). *The adaptive character of thought*. Psychology Press.

Auer, P. (2002). Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, *3*(Nov), 397–422.

Berry, D. C., & Broadbent, D. E. (1987). The combination of explicit and implicit learning processes in task control. *Psychological Research*, *49*(1), 7–15.

Bramley, N. R., Dayan, P., Griffiths, T. L., & Lagnado, D. A. (2017). Formalizing neurath's ship: Approximate algorithms for online causal learning. *Psychological Review*, *124*(3), 301.

Bramley, N. R., Jones, A., Gureckis, T. M., & Ruggeri, A. (2022). Children's failure to control variables may reflect adaptive decision-making. *Psychonomic Bulletin & Review*, *29*(6), 2314–2324.

Brehmer, B. (1987). Note on subjects' hypotheses in multiple-cue probability learning. *Organizational Behavior and Human Decision Processes*, *40*(3), 323–329.

Btesh, V., Bramley, N., Speekenbrink, M., & Lagnado, D. (2025). Less is more: Local focus in continuous time causal learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Conant, R. C., & Ashby, W. R. (1970). Every good regulator of a system must be a model of that system. *International Journal of Systems Science*, *1*, 89–97.

Davis, Z. J., Bramley, N. R., Rehder, B., & Gureckis, T. (2018). A causal model approach to dynamic control. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *40*, 40.

Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, affective & behavioral neuroscience*, *8*, 429–453.

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*(4), 968.

Dunham, J. G. (1986). Optimum uniform piecewise linear approximation of planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1), 67–75.

Fernbach, P. M., & Sloman, S. A. (2009). Causal learning with local computations. *Journal of experimental psychology: Learning, memory, and cognition*, *35*(3), 678.

Gibson, F. P., Fichman, M., & Plaut, D. C. (1997). Learning in dynamic decision tasks: Computational model and empirical evidence. *Organizational Behavior and Human Decision Processes*, *71*(1), 1–35.

Gong, T., Pacer, M., Griffiths, T. L., & Bramley, N. R. (2025). Rational causal induction from events in time. *Psychological Review*.

Griffiths, T. L. (2020). Understanding human intelligence through human limitations. *Trends in Cognitive Sciences*, *24*(11), 873–883.

Kuhn, D., & Brannock, J. (1977). Development of the isolation of variables scheme in experimental and" natural experiment" contexts. *Developmental Psychology*, *13*(1), 9.

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

Markant, D. B., Settles, B., & Gureckis, T. M. (2016). Self-directed learning favors local, rather than global, uncertainty. *Cognitive Science*, *40*, 100–120.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Henry Holt; Co., Inc.

Osman, M., & Speekenbrink, M. (2012). Prediction and control in a dynamic environment. *Frontiers in Psychology*, *3*, 68.

Sanborn, A. N., Griffiths, T. L., & Shiffrin, R. M. (2010). Uncovering mental representations with Markov chain Monte Carlo. *Cognitive Psychology*, *60*(2), 63–106.

Schulz, E., Klenske, E. D., Bramley, N. R., & Speekenbrink, M. (2017). Strategic exploration in human adaptive control. *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*.

Schulz, E., Konstantinidis, E., & Speekenbrink, M. (2018). Putting bandits into context: How function learning supports decision making. *Journal of experimental psychology: learning, memory, and cognition*, *44*(6), 927.

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. https://www.sciencedirect.com/science/article/pii/S0022249617302158

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.

Speekenbrink, M., & Konstantinidis, E. (2015). Uncertainty and exploration in a restless bandit problem. *Topics in Cognitive Science*, *7*, 351–367.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction (2nd ed.)* The MIT Press.

Tschirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development*, *51*(1), 1–10.

Vollmeyer, R., Burns, B. D., & Holyoak, K. J. (1996). The impact of goal specificity on strategy use and the acquisition of problem structure. *Cognitive Science*, *20*(1), 75–100.

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.