Modeling Cognitive Strategies in Teaching

Sevan K. Harootonian
Department of Psychology
Princeton University

Thomas L. Griffiths
Department of Psychology
Department of Computer Science
Princeton University

Yael Niv
Department of Psychology
Princeton Neuroscience Institute
Princeton University

Mark K. Ho
Department of Psychology
New York University

Abstract

Teaching is a complex social behavior that sometimes results from goal-directed processing. However, goal-directed teaching is cognitively demanding since it requires actively assessing and correcting gaps in a learner's knowledge. When do people teach using such mentally effortful strategies versus falling back on more cognitively frugal ones? Here, we investigated this question using a combination of novel behavioral experiments and computational theory. We found robust individual differences in people's teaching strategies: some participants spontaneously teach using high-effort processing (e.g., Bayesian theory of mind and model-based planning) while others engage in low-effort processing (e.g., model-free heuristics). Our results and analyses provide a novel demonstration of how people engage in planning versus heuristics when teaching, as well as how people adapt processing to avoid mental effort in social interactions.

1 Introduction

Humans are unmatched in their achievements as a species, from acquiring language and solving complex problems to creating sophisticated technologies like microprocessors and modern societies. These accomplishments are attributed to our extraordinary capacity for flexible learning, which enables abstract thinking and advanced reasoning—skills that modern machine-learning models strive to replicate [12]. However, much of our knowledge is actually acquired through social interactions, and behind every exceptional person, there is often a teacher or mentor figure who has significantly influenced their thinking [21, 16]. Despite the prevalence of teaching and mentoring, from indigenous tribes passing down traditions to university classrooms, the cognitive strategies that teachers employ to facilitate effective learning remain largely unexplored[4].

Teaching is challenging because it requires teachers to infer which information will most benefit the learner's understanding and performance. This process often relies on theory of mind—the ability to attribute mental states such as beliefs, desires, and intentions to others. Effective teaching requires the teacher to infer the learner's knowledge state and predict how the learner will interpret and respond to the provided information [3, 5, 17]. This has been modeled through frameworks like Bayesian Theory of Mind [8, 1, 9, 10] and Rational Speech Act models [18, 7, 2], which formalize how teachers anticipate learners' responses by considering how their beliefs change with different inputs. These models view teaching as rational communication, where teachers act as rational agents, choosing actions based on expected utility, including the learner's belief updates [2]. They suggest that goal-directed teaching strategies, relying on mental state inference, enable flexibility and adaptability, allowing teachers to tailor methods to the learner's progress [14].

Accepted at the Intrinsically Motivated Open-ended Learning workshop at NeurIPS 2024.

On the other hand, teaching can also be guided by more habitual and automatic behaviors. The habitual and automatic approach relies on ingrained routines or heuristics, where actions that have been previously effective are performed with minimal conscious thought or adaptation to the specific context. Heuristic strategies, less emphasized in teaching research, are often modeled through simpler, rule-based approaches that capture how teachers might default to well-practiced patterns of instruction without actively considering the learner's current knowledge state [6, 20, 11, 19]. Heuristics offer a lower cognitive-cost alternative by relying on simple decision rules or "shortcuts" that can be executed rapidly and with minimal mental effort, making them efficient in many teaching scenarios. However, this efficiency comes at the cost of reduced flexibility. Habitual strategies, driven by heuristics, may persist even in situations in which they are not the most effective approach, particularly in novel teaching environments [22].

Both goal-directed and habitual systems can be considered optimal and rational within their respective contexts, aligning with the broader concept of resource-rationality. According to this framework, individuals optimize their cognitive strategies not just to achieve their goals but also to manage their cognitive resources efficiently [13]. For example, to conserve cognitive resources, a teacher might rely on habitual strategies in familiar or low-stakes situations, while engaging in more goal-directed behavior when the stakes are higher or the learning context is novel or complex.

Building on these insights, here we investigate the cognitive strategies that individuals employ when teaching. We developed a novel Graph Teaching Task to measure individual variability in teaching strategies, and to distinguish between heuristic-based and mental-state inference-based approaches. Our findings reveal significant individual differences in teaching strategies, with some participants favoring heuristic-based methods while others engage in more theory of mind reasoning.

2 Methods

Task: Participants acted as teachers in the experiment. They were shown a graph with possible edges connecting nodes with rewards to be collected, and a single path from top to bottom already taken by the learner. Participants were instructed that the learner may not know all the possible edges (paths) between nodes and that they had taken a path that collected as many points as they could, given their knowledge (Figure 1). The participant's task was to teach the learner one edge to help the learner improve their performance as much as possible. Critically, there was no feedback presented, meaning that the teachers did not know whether their teachings had helped the learner's performance. The task consisted of 40 trials, each featuring a new learner with a new graph.

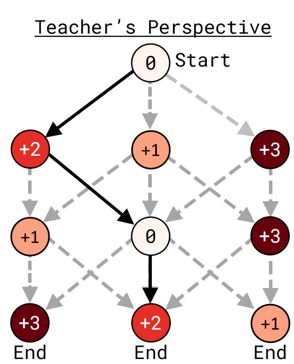

Figure 1: An example trial from experiment 1 where the nodes represent reward values, grey edges are possible paths, and dark edges indicate the learner's previous path attempting to collect the most points. This is a directed deterministic graph where the learner can only move downwards.

Participants: One hundred participants (61 female, 31 male; mean age: 34 ± 10.38 years) were recruited from Prolific. Participants earned an average of \$4.02 ± \$1.26 and spent 19.16 ± 10.39 minutes on the experiment. One participant who expressed confusion about the task was removed from the analysis, and no participant performed below chance level. This study was approved by the Institutional Review Board of Princeton University, and all participants provided informed consent

3 Modeling Approach

We used two types of models to explain participant behavior: Bayesian inference models and heuristic-based models. The Bayesian models aim to infer the learner's knowledge and maximize the utility of teaching an edge, while the heuristic models rely on task features rather than learner features.

3.1 Bayesian Mentor Models

The Bayesian Mentor models consist of three variants: the *Optimal Bayesian Mentor* (OBM), the *Naive Bayesian Mentor* (NBM), and the *Prior-only Mentor* (POM). Each model follows the same general structure: (1) modeling the learner's behavior, (2) inferring the learner's knowledge, and (3) calculating the utility of teaching each edge.

Learner Model: We model the learner as an agent navigating a directed deterministic graph, represented as a Markov Decision Process $M_L = \langle S, \mathcal{T}_L, R \rangle$, where S is the set of nodes (states), $\mathcal{T}_L$ is the learner's transition model (a subset of the true transitions $\mathcal{T}$), and R is the known reward function, which assigns rewards to each node. The learner follows an optimal policy, resulting in a trajectory $\zeta = \{s_0, s_1, ..., s_T\}$.

Utility Calculation: For each Bayesian model, the utility of teaching an edge $e = (s, s')$ is the improvement in the learner's performance, defined as the difference in cumulative reward between the learner's current trajectory ζ and the new trajectory ζ_e:

$$U(e \mid \zeta) = \sum_{\mathcal{T}_L} P(\mathcal{T}_L \mid \zeta) \left(G(\zeta_e) - G(\zeta) \right) \tag{1}$$

where $G(\zeta)$ is the cumulative reward, and $P(\mathcal{T}_L \mid \zeta)$ is the probability distribution over the learner's transition model.

Inference Variants: The main difference between Bayesian models is in how $P(\mathcal{T}_L \mid \zeta)$ is computed.

Optimal Bayesian Mentor (OBM): The OBM infers the learner's transition model $\mathcal{T}_L$ by considering the likelihood of the observed trajectory ζ under the learner's optimal policy:

$$P_{\text{OBM}}(\mathcal{T}_L \mid \zeta) \propto P(\zeta \mid \mathcal{T}_L) P(\mathcal{T}_L) \tag{2}$$

Here, $P(\zeta \mid \mathcal{T}_L)$ is computed via inverse planning, representing the likelihood of the trajectory ζ given $\mathcal{T}_L$, and $P(\mathcal{T}_L)$ is a uniform prior over valid transition models.

Naive Bayesian Mentor (NBM): The NBM uses a simplified likelihood function that assigns probability 1 if the learner's transition model $\mathcal{T}_L$ is consistent with the observed trajectory ζ, and 0 otherwise:

$$P_{\text{NBM}}(\mathcal{T}_L \mid \zeta) \propto \begin{cases} 1, & \text{if } \zeta \subseteq \mathcal{T}_L \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

This means the NBM considers all $\mathcal{T}_L$ that include ζ, without evaluating the optimality of the trajectory.

The *Prior-only Mentor (POM)* does no inference but rather assumes a uniform prior over all possible transition models.

3.2 Heuristic Mentor Models

Heuristic Mentor models rely on simple task features to select edges, offering a low-cost alternative to Bayesian models that is adjusted to the task at hand, but not to the specific learner.

Reward Heuristic: This heuristic prioritizes edges based on the sum of rewards at connected nodes: $F_{\text{Reward}}(e_{i,j}) = R(s_i) + R(s_j)$, with higher reward sums indicating higher utility.

Level Heuristic: This heuristic considers the vertical position of edges. Top edges (closer to the start) can be riskier but more impactful, while bottom edges are typically safer. The feature is defined as $F_{\text{Level}}(e_{i,j}) = j$, with the regression model capturing individual differences in risk tolerance.

4 Results

We observed significant individual differences in Teaching Scores[1], indicating that participants may be using different teaching strategies (Figure 2A). The dotted vertical lines represent the performance of different models, with the Bayesian mentor models outperforming the heuristic models. The bimodal distribution of participant scores in Figure 2 suggests the dominance of two primary teaching strategies: a heuristic-based strategy (perhaps a combination of the Reward and Level heuristics) and a mentalizing-based strategy, which involves either a noisy optimal inference (OBM), a suboptimal inference (NBM), or no inference at all (POM).

To further understand the teaching strategies employed by participants, we computed Bayesian Information Criterion (BIC) scores for each model for each subject. Figure 2B shows the mean BIC scores with 95% confidence intervals for the different models (black circles with whiskers), together with individual BIC scores represented by colored points, where the color reflects each participant's teaching score. Lower BIC scores indicate a better fit, and the two best-fitting models are the mixture heuristic model of Level & Reward and OBM. As can be seen by the colors of the dots, the mixture heuristic model fit the lower-performing individuals best (blue hues) while OBM fit the higher-performing individuals best (yellow hues). This finding strengthens our hypothesis that different individuals are using distinct mental strategies to teach. Although the mixture heuristic model provided a significantly better fit for the dataset as a whole (Mann-Whitney U test: $U = 4060, n_1 = 99, n_2 = 99, p = 0.037$), Figure 2C shows a higher number of subjects were best fit by the OBM model (with at least a $\Delta BIC \geq 10$, which is considered "very strong" evidence against models with higher BIC [15]). Again, we see that the best fit individuals for OBM are the higher-performing ones, further supporting our hypothesis.

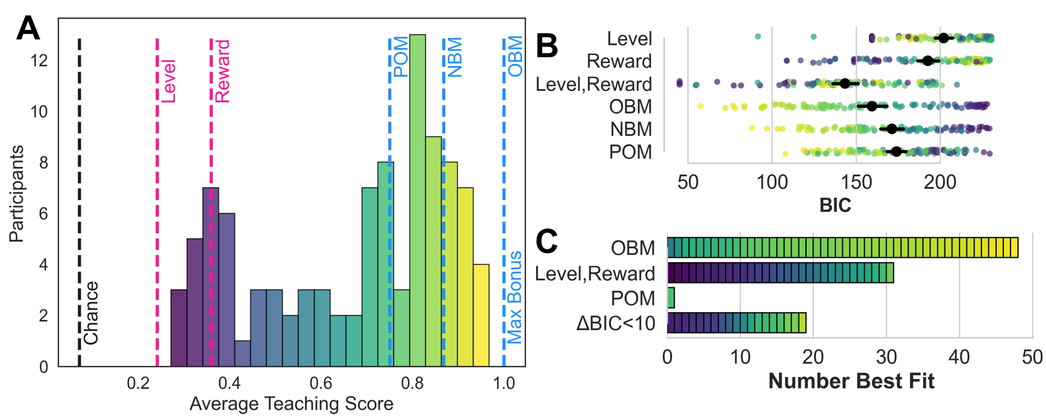

Figure 2: A) Histogram of participants' average Teaching Score across 40 trials. The dotted vertical lines represent the performance of each model as a pure teaching strategy. B) BIC scores for different models across subjects.The x-axis show the 95% CI of the mean BIC. The colored points are BIC scores for individual and the color reflecting their Teaching Score. C) Bar plot of the number of subjects best fit by each model with at least a $\Delta BIC \geq 10$. Colors in B,C represent individuals' average performance as per A.

5 Conclusion

Our study reveals significant individual differences in teaching strategies, ranging from cognitively demanding goal-directed approaches to simpler heuristic-based methods. Using a novel Graph Teaching Task, we demonstrated that while some participants engage in complex mental-state reasoning, others rely on "quicker", model-free heuristics. [2] Those using goal-directed strategies

[1]As a normalized performance metric, we calculated a "Teaching Score" for each trial by dividing the additional points the average learner (given the current graph) would have gained from the participant's teaching by the points the learner would have gained from the optimal edge, as determined by the Optimal Bayesian Mentor model. This normalization ensured that trials with graphs yielding more points would not be overweighted compared to other trials, providing a normalized measure of performance across different graph configurations.

[2]The task was not designed as a reaction time task, hence we can't confirm if the heuristics were faster to use.

performed better, as shown by their higher Teaching Scores and closer fit to the Optimal Bayesian Mentor model, highlighting the adaptive nature of teaching strategies based on cognitive effort and task complexity.

These findings suggest that understanding how humans adaptively balance heuristic and goal-directed teaching strategies can inform the development of autonomous systems that learn and interact with humans in complex environments. By modeling this flexibility, machines could better adapt to evolving contexts and engage more effectively by adjusting their strategies based on human behavior.

References

[1] Chris L Baker, Julian Jara-Ettinger, Rebecca Saxe, and Joshua B Tenenbaum. Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, 1(4):0064, 2017.

[2] Alicia M Chen, Andrew Palacci, Natalia Vélez, Robert D Hawkins, and Samuel J Gershman. A hierarchical bayesian model of adaptive teaching. *Cognitive Science*, 48(7):e13477, 2024.

[3] Michelene TH Chi, Stephanie A Siler, Heisawn Jeong, Takashi Yamauchi, and Robert G Hausmann. Learning from human tutoring. *Cognitive science*, 25(4):471–533, 2001.

[4] Christopher M Clark. *Teachers' thought processes*. Number 72. Institute for Research on Teaching, Michigan State University, 1984.

[5] Chris D Frith and Uta Frith. The social brain: Allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361 (1467):2201–2214, 2006.

[6] Gerd Gigerenzer, Peter M Todd, and the ABC Research Group. *Simple heuristics that make us smart*. Oxford University Press, 1999.

[7] Noah D Goodman and Michael C Frank. Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, 20(11):818–829, 2016.

[8] Julian Jara-Ettinger. Theory of mind as inverse reinforcement learning. *Current Opinion in Behavioral Sciences*, 29:105–110, 2019.

[9] Julian Jara-Ettinger, Hyowon Gweon, Laura E Schulz, and Joshua B Tenenbaum. The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in cognitive sciences*, 20(8):589–604, 2016.

[10] Alan Jern and Charles Kemp. A decision network account of reasoning about other people's choices. *Cognition*, 142:12–38, 2015.

[11] Paul A Kirschner, John Sweller, and Richard E Clark. Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2):75–86, 2006.

[12] Brenden M Lake, Tomer D Ullman, Joshua B Tenenbaum, and Samuel J Gershman. Building machines that learn and think like people. *Behavioral and Brain Sciences*, 40:e253, 2017.

[13] Falk Lieder and Thomas L Griffiths. Strategy selection as rational metareasoning. *Psychological Review*, 124(6):762–794, 2017.

[14] Donald A Norman and Tim Shallice. Attention to action: Willed and automatic control of behavior. In *Consciousness and Self-Regulation*, volume 4, pages 1–18. Springer, 1986.

[15] Adrian E Raftery. Bayesian model selection in social research. *Sociological methodology*, pages 111–163, 1995.

[16] Barbara Rogoff. *Apprenticeship in thinking: Cognitive development in social context*. Oxford university press, 1990.

[17] Rebecca Saxe. Uniquely human social cognition. *Current Opinion in Neurobiology*, 16(2): 235–239, 2006.

[18] Patrick Shafto and Noah D Goodman. Teaching and the bayesian learner. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, pages 1632–1637, 2008.

[19] John E Taplin, Anna Schwenk, and Larisa V Shavinina. Heuristic and analytic processes in reasoning. *Thinking and Reasoning*, 10(4):433–455, 2004.

[20] Peter M Todd and Gerd Gigerenzer. Ecological rationality: Intelligence in the world. *Oxford Handbook of Thinking and Reasoning*, pages 197–214, 2012.

[21] Lev Semenovich Vygotsky and Michael Cole. *Mind in society: Development of higher psychological processes*. Harvard university press, 1978.

[22] Wendy Wood and David T Neal. A new look at habits and the habit-goal interface. *Psychological Review*, 114(4):843–863, 2007.