COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 48 (2024) e13477

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13477

A Hierarchical Bayesian Model of Adaptive Teaching

Alicia M. Chen,[a] Andrew Palacci,[b] Natalia Vélez,[c]
Robert D. Hawkins,[d,*] Samuel J. Gershman[b,e,*]

[a]*Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology*
[b]*Department of Psychology, Harvard University*
[c]*Department of Psychology, Princeton University*
[d]*Department of Psychology, University of Wisconsin–Madison*
[e]*Center for Brains, Minds, and Machines, Massachusetts Institute of Technology*

Received 9 February 2023; received in revised form 5 June 2024; accepted 8 June 2024

Abstract

How do teachers learn about what learners already know? How do learners aid teachers by providing them with information about their background knowledge and what they find confusing? We formalize this collaborative reasoning process using a hierarchical Bayesian model of pedagogy. We then evaluate this model in two online behavioral experiments ($N = 312$ adults). In Experiment 1, we show that teachers select examples that account for learners' background knowledge, and adjust their examples based on learners' feedback. In Experiment 2, we show that learners strategically provide more feedback when teachers' examples deviate from their background knowledge. These findings provide a foundation for extending computational accounts of pedagogy to richer interactive settings.

Keywords: Pedagogy; Theory of mind; Communication; Social cognition; Bayesian modeling

1. Introduction

Humans do not just passively absorb knowledge from the world; we actively teach one another culturally specific skills and knowledge (Csibra & Gergely, 2009; Kline, 2015). We can be taught to distinguish edible mushrooms from poisonous ones, to build more effective

*Robert D. Hawkins and Samuel J. Gershman contributed equally to this work.
Correspondence should be sent to Alicia M. Chen, Department of Brain and Cognitive Sciences, MIT, 43 Vassar St., Cambridge, MA 02139, USA. E-mail: aliciach@mit.edu

tools, and to solve differential equations, all of which would be difficult or dangerous to acquire through trial and error alone. However, the cognitive mechanisms enabling effective pedagogy are not fully understood. Teaching is especially challenging because learners differ in their prior knowledge, learning goals, and cognitive abilities (Fyfe & Rittle-Johnson, 2016; Kalyuga, 2007; Vygotsky, 1978). Given such variability, how do teachers know what to teach? And when learners receive information from teachers, how do they interpret that information in light of what they already know?

Bayesian cognitive models offer one framework for understanding the mechanisms by which people teach and learn. The most influential class of these models characterizes teaching as an application of theory of mind, our capacity to reason about others' mental states and to predict how others will respond to the information we provide. Formally, in these models, a teacher selects what to teach by anticipating how the learner's beliefs will change under different kinds of information, and a learner tries to recover the intended concept by considering how a helpful teacher would have selected information for them (Shafto & Goodman, 2008; Shafto, Goodman, & Frank, 2012; Yang, Vong, Yu, & Shafto, 2019). These models belong to a broader class of *rational communication* approaches, where humans are modeled as rational agents that choose actions proportion to their expected utility (Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016), and consider how informative an utterance would be by reasoning about how it will change the beliefs of the listener in the intended direction (Frank & Goodman, 2012; Goodman & Frank, 2016).

This basic probabilistic model has successfully accounted for a wide range of empirical phenomena in psychology. For example, it has been used to explain how teachers induce rule-based concepts from limited pedagogical examples (e.g., Shafto, Goodman, & Griffiths, 2014) and how inferences differ across pedagogical and nonpedagogical contexts (Bonawitz et al., 2011). It has also formed a basis for a number of successful applications in robotics and artificial intelligence (Ho, Littman, MacGlashan, Cushman, & Austerweil, 2016; Yang, Vong, Sojitra, Folke, & Shafto, 2021; Zhu, Singla, Zilles, & Rafferty, 2018). Importantly, however, this class of models has typically treated teaching as a one-way street. Teachers are assumed to already know the learners' background knowledge (often complete naivety) and provide information in a single batch. In other words, the teacher models do not adapt their instruction based on feedback from the learner, and the learner models do not influence the teaching process. Every student gets the same lecture.

Rather than treating students as passive listeners, modern educators are encouraged to adopt more adaptive, personalized approaches and specialize their teaching to each student's knowledge and backgrounds (Cronbach & Snow, 1977; Staub & Stern, 2002; Hoy, Davis, & Pape, 2006; Goodboy & Myers, 2008; Slavich & Zimbardo, 2012; van de Pol, Volman, & Beishuizen, 2012; Lindsey, Shroyer, Pashler, & Mozer, 2014; but see Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001; VanLehn, Siler, Murray, Yamauchi, & Baggett, 2003; Narciss, 2008; Carless, 2022, for discussions regarding the limitations of personalized teaching for improving student learning). Contrary to the predictions of existing cognitive models, sequential interaction helps teachers pinpoint what students do and do not know, and allows students to give feedback to the teacher (Rocca, 2010; Zhu, 2015). Thus, over repeated interactions, teachers may, over time, *learn* about the learner so they can provide information that is tailored to them. Teaching may thus be similar to other forms of communication where learning about our partners helps us establish common ground between them, and adapt and communicate better over time (Clark & Marshall, 1981; Hawkins et al., 2022; Hao, Jhaveri, & Shafto, 2024; Isaacs & Clark, 1987; Kleinschmidt & Jaeger, 2015).

The goal of this paper is to formalize the processes used by teachers and learners to adapt to each other during pedagogical interactions. Specifically, we propose a hierarchical Bayesian model of pedagogy in which teachers and learners interact to resolve uncertainty at two levels: (1) the learner's uncertainty about the target concept; and (2) the teacher and learner's higher-order uncertainty about *what the other knows*. Our model predicts that teachers and learners use rational Bayesian inference to update their beliefs about each other in response to observed communication from their partner. To test the predictions of this model, we present results from two proportion-estimation experiments where participants interact with agents simulated from our model. In Experiment 1, participants are teachers who have uncertainty about the learner's background knowledge and may resolve this uncertainty using feedback from the learner. In Experiment 2, participants are learners who have uncertainty about what their teachers already know about them and what the correct concept is. They may resolve this uncertainty by providing feedback to the teacher. By manipulating what information is available and the value of that information, these experiments provide evidence for how teachers and learners can be sensitive to background knowledge and feedback, consistent with key predictions distinguishing our hierarchical model from previous models.

1.1. Open practices

All data, materials, and analysis code are publicly accessible at https://osf.io/ubxjr. All experiments and confirmatory analyses were preregistered; the preregistration for Experiment 1 is accessible at https://osf.io/7xbcg and the preregistration for Experiment 2 is accessible at https://osf.io/utmxq.

2. A hierarchical Bayesian model of adaptive teaching

In earlier nonadaptive models of teaching, teachers compute the utility of selecting an example (i.e., communicating a piece of information) by using theory of mind: They invert a model of the learner to reason about how the learner's beliefs might change, given a hypothetical example (Shafto & Goodman, 2008; Shafto et al., 2014). The teacher then selects examples in proportion to the likelihood that the learner assigns to the target concept, based on their prior beliefs and on the information presented so far. This inference is derived from Bayes' rule: The teacher must combine whatever they know about the learner's background knowledge (the prior) with, however, they think the learner would evaluate the (hypothetical) information provided (the likelihood). However, in adaptive and interactive settings, both teachers and learners act under uncertainty at multiple levels of abstraction. The teacher may be uncertain about the learner's background knowledge, which in turn affects their uncertainty about the learner's belief in the target concept. The learner, in turn, may not know whether the teacher is aware of their background knowledge, which in turn affects their uncertainty about what the teacher is trying to tell them.

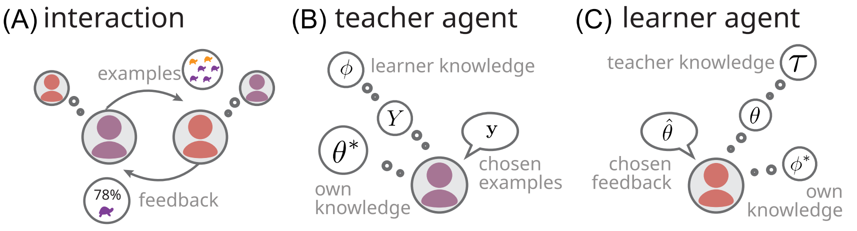

Fig. 1. Model schematic. (A) Teachers communicate by sending examples Y, and learners communicate by sending feedback $\hat{\theta}$. (B) A teacher reasons about the learner's background knowledge ϕ and which examples Y would be helpful for leading the learner to the true concept θ^*. (C) A learner reasons about what the teacher knows about them τ, what the right concept θ is, and whether it would be helpful to send feedback $\hat{\theta}$ to the teacher.

To take an example, suppose you just started tutoring a high-school student in math. How do you approach the tutoring session? If you know that there are three common misconceptions people have about solving systems of equations, you might be able to pin down whether a particular student is falling into misconception C rather than misconception A or B. If you knew which misconception they have, then you could adapt your teaching to clarify that misconception. In other words, if you know how learners' background knowledge could be structured—such as the existence of three common misconceptions—how do you teach to tease apart which misconceptions your student has? And how does your student communicate with you, to help you further narrow down what misconceptions they might have?

We formalize these intuitions in a computational framework that generates quantitative predictions by extending Bayesian models of pedagogy (Shafto et al., 2014) with *structured uncertainty* over the partner's background knowledge, which may be updated across interactions. Thus, we formalize our proposed theory as a hierarchical Bayesian model (Gelman, Carlin, Stern, & Rubin, 1995). Hierarchical Bayesian models have been widely used in cognitive science to explain inductive reasoning (Goodman, Ullman, & Tenenbaum, 2011; Gershman & Niv, 2015; Kemp, Perfors, & Tenenbaum, 2007). In the context of communication, they have been used to model how speakers and listeners coordinate on partner-specific meanings while also learning about which meanings may be shared as conventions in the broader community (Hawkins et al., 2022).

We begin by assuming teachers and learners are rational agents that choose actions proportional to their expected utility (Bridgers, Jara-Ettinger, & Gweon, 2020; Jara-Ettinger et al., 2016; Vélez, Chen, Burke, Cushman, & Gershman, 2023). Here, the actions available to teachers and learners are communicative acts—teachers can communicate with learners by choosing examples, and learners can communicate with teachers by providing feedback (Fig. 1). At a high level, both teachers and learners compute the utility of actions by considering their downstream costs and rewards. We model the teacher and learners' goals as aligned: The learner's goal is to learn a target concept, and the teacher's goal is to provide examples that will lead the learner to the target concept.

Calculating the expected utility of an action involves multiple steps of reasoning. First, in order to select examples that will be most helpful to the learner, the teacher considers what background knowledge the learner may have prior to their interaction; the teacher's beliefs about the learner's background are informed both by the teacher's world knowledge, and by any feedback that the student may have provided. Then, for each possible set of background knowledge, the teacher anticipates how the learner's beliefs will change after the teacher presents a new example; they weigh these potential benefits of the new example against the costs of providing it.

The learner, in turn, uses an analogous process to decide whether to send feedback to the teacher. They weigh the cost of sending feedback against the potential benefits of giving the teacher information that will help them teach better in the future. The learner's reasoning is as follows. First, after the learner receives information from the teacher, the learner uses Bayesian inference to jointly infer what the correct concept is and what the teacher knows about them. For example, if the teacher already knows everything about the learner's background knowledge, then giving feedback would not actually help the teacher provide more informative examples. Then, for each possible ground-truth concept and each possible representation the teacher might have of the learner's knowledge, the learner anticipates how that teacher might change their downstream teaching behavior based on their feedback, and how their own beliefs about the correct concept may change in response to this downstream behavior. In this way, both agents are able to teach and learn more effectively as they interact.

2.1. *Proportion estimation task*

To evaluate our models (full mathematical description under the Experiment 1 and Experiment 2 sections) directly against human behavior, we designed a simple proportion estimation task (see Frank & Liu, 2018, for the use of a similar task to model a variety of phenomena in classroom teaching, including the effects of ability grouping and assessment). Participants, interacting with an agent simulated from our model, are told that they are either tour guides (Experiment 1) or tourists (Experiment 2) on an island famous for its colorful turtles. Some turtles are orange, and some are purple. Participants in the role of tour guide (i.e., teachers T) know the true proportion θ^* of different colors on the island, and are tasked with transmitting that information to tourists (i.e., learners L). The tour guide is able to bring the tourists to a region of the island and show them the set of turtles there (a finite set of turtles of each color $Y = \{n_{\text{orange}}, n_{\text{purple}}\}$). Critically, each tourist has some *background knowledge*, operationalized as their immediate prior experience about the island. They have previously visited a small region of the island where they saw a few orange and purple turtles. This background knowledge defines the learner's prior ϕ, which may be skewed toward orange ($\phi_o = \{n_{\text{orange}}, n_{\text{purple}}\} = \{9, 1\}$) or purple ($\phi_p = \{n_{\text{orange}}, n_{\text{purple}}\} = \{1, 9\}$). The teacher has uncertainty $P_T(\phi)$ over the learner's background knowledge.

Learners have the option to send feedback in the form of an estimate $\hat{\theta}$ of the true proportion to communicate their beliefs about what they think the correct island proportion is. At the end of multiple rounds of teaching, learners are asked to estimate the proportion of orange and purple turtles on the island. In the experiments below, we test quantitative predictions derived from our model about how teachers and learners adjust their communication in response to information from each other.

2.2. Implementation details

2.2.1. Model implementation details

We ran simulations to (1) generate actions for the agents participants interact with in the experiment and (2) evaluate human behavior against our model. We implemented our simulations using the probabilistic programming language WebPPL (Goodman & Stuhlmüller, 2014). There are four free parameters. Two of the free parameters are the teacher's inverse temperature parameter α_T and the learner's inverse temperature parameter α_L. These parameters modulate the amount of probability weight assigned to each utility value: As the temperature decreases, agents are more likely to choose the higher-utility actions. In practice, we set these parameters to the same value $\alpha = \alpha_T = \alpha_L = 4$. The two other free parameters are the teacher's action cost w and the learner's action cost c. These parameters capture the intuition that communicating lengthier amounts of information is costlier. In practice, we also set these parameters to the same value, $w = c = 0.01$.

For all simulations, we used exhaustive enumeration for exact inference. For Experiment 1, we discretized the space of island weights $\theta \in [0, 1]$ in increments of 0.01 for computational tractability. For Experiment 1, our dependent measure was the total number of turtles sent (in any combination of orange and purple), so we simulated 193 total participants (the number of participants in our experiment) using the model. For Experiment 2, our dependent measure was whether the participant sent a guess, so the model outputs were guess probabilities. See the Supplementary Appendix for simulations depicting model behavior over a wider range of parameter values.

2.2.2. Experiment implementation details

We implemented both experiments using the jsPsych library (De Leeuw, 2015). All statistical analyses were conducted in R using the lme4 and lmerTest packages (Bates, Mächler, Bolker, & Walker, 2014; Kuznetsova, Brockhoff, & Christensen, 2017).

3. Experiment 1: Teachers are sensitive to learners' priors and feedback

In Experiment 1, participants played the role of tour guides who had to teach a "tourist" (a simulated agent) the true proportion of turtles on the island (see Fig. 2). We investigated our model's predictions about how teachers maintain uncertainty about learners' prior beliefs when they teach, and how they may "personalize" their teaching as they learn more about individual learners. In accordance with the predictions of our model, we predicted that participants would both (1) take into account explicit information about learners' background knowledge to modulate the information they initially select, and (2) adjust their teaching based on learners' subsequent feedback.

3.1. Model overview for Experiment 1

Each trial unfolds in two phases. First, the teacher chooses a set of examples, Y_1, to teach the proportion, and then the simulated learner updates their beliefs and gives feedback $\hat{\theta}$.

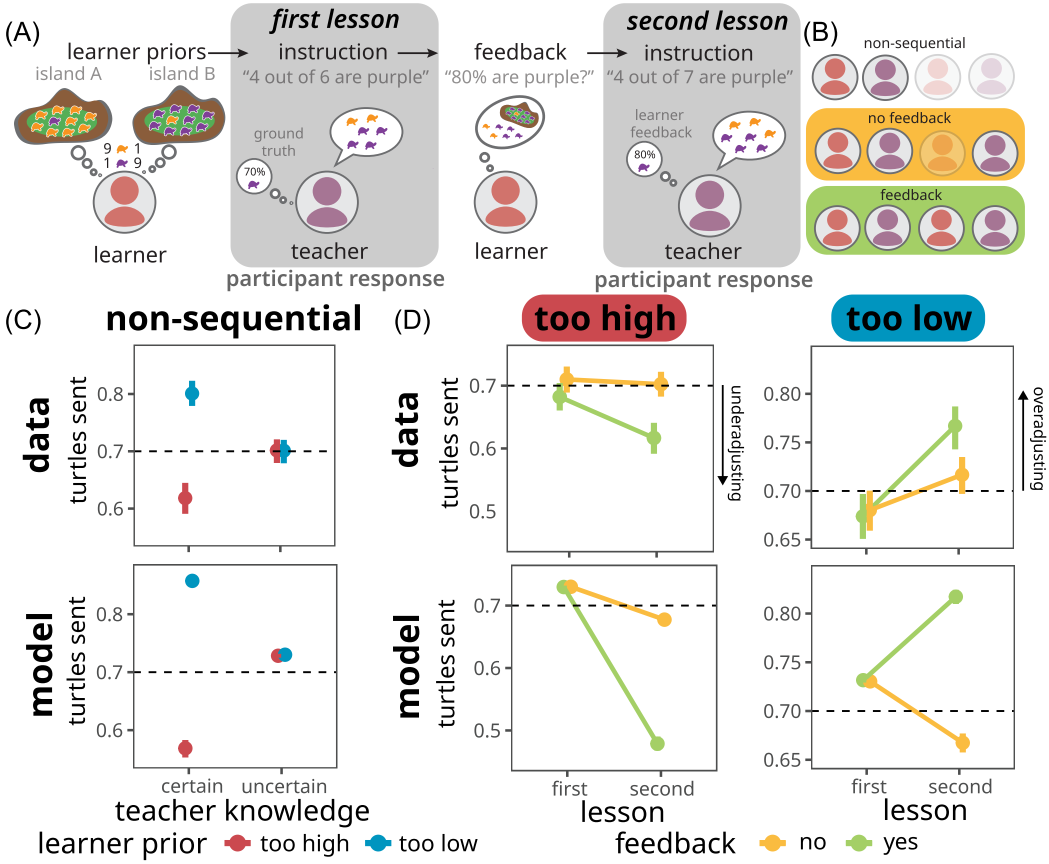

Fig. 2. Experiment 1. (A) Experiment 1 schematic. (B) Conditions. (C) Results for nonsequential conditions. The y-axis is the proportion of turtles sent in the majority color of turtles on the island, and the dotted line is the true island proportion. Participants are sensitive to learners' priors (i.e., they underadjust in the "too high" condition and overadjust in the "too low" condition), but only when they know about them. (D) Results for sequential conditions. In the "feedback" condition, participants receive feedback from the learner after the first lesson. In response to feedback, participants underadjust in the "too high" condition and overadjust in the "too low" condition. Error bars for all plots are bootstrapped 95% confidence intervals.

Our primary interest is the second phase, where the teacher uses the feedback to update their beliefs about the learner background knowledge ϕ and select a second set of examples, Y_2. We proceed through each phase in turn.

3.1.1. *Selecting the first set of examples Y_1*

The teacher balances the informativity of selecting a certain example with the cost of providing it, to evaluate which examples to select. If the teacher already perfectly knows the learner's background knowledge ϕ, then their utility function is:

$$U(Y_1; \theta^*, \phi) = (1 - w) \cdot \underbrace{\log P_{L_0}(\theta^* | Y_1, \phi)}_{\text{reward (informativity)}} - w \cdot \underbrace{|Y_1|}_{\text{cost}}. \tag{1}$$

This utility function follows the speaker utility function in Rational Speech Act models of pragmatic communication (Frank & Goodman, 2012; Goodman & Frank, 2016). The reward term $\log P_{L_0}(\theta^* | Y_1, \phi)$ describes how helpful the examples are for updating the learner's beliefs, and the cost term $|Y_1|$ corresponds to the number of examples that the teacher sends to the learner. The cost term captures the intuition that speakers balance the helpfulness of their examples with a tendency to be parsimonious (i.e., provide just as many examples as is helpful); for experimental control, we fixed this cost across all participants in our experiments and in our model. The free parameter $w \in [0, 1]$ controls the relative weighting of cost and reward.[1]

To compute the reward term $\log P_{L_0}(\theta^* | Y_1, \phi)$, the teacher reasons about a base learner L_0 that is assumed to use Bayesian inference to estimate the true proportion that explains the evidence they have seen (i.e., equivalent to inferring a coin weight from a series of flips):

$$P_{L_0}(\theta^* | Y_1, \phi) \propto P(Y_1 | \theta^*, \phi) P(\theta^* | \phi) \tag{2}$$

The teacher selects examples using a softmax decision rule on the utility function given above, with softmax inverse temperature parameter $\alpha_T \in [0, \infty)$. The larger α_T is, the more likely the teacher is to select the single highest-utility example set:

$$P_T(Y_1 | \theta^*, \phi) \propto \exp(\alpha_T \cdot U(Y_1; \theta^*, \phi)). \tag{3}$$

So far, this teacher model is equivalent to prior accounts. But what should the teacher do when they do *not* know the learner's background knowledge? In a Bayesian framework, the teacher ought to *marginalize* over $P_T(\phi)$, their prior beliefs about what the learner already knows:

$$P_T(Y_1 | \theta^*) \propto \exp\left(\alpha_T \cdot \int_\phi U(Y_1; \theta^*, \phi) P_T(\phi)\, d\phi\right). \tag{4}$$

This model yields predictions about what the teacher will do in the first phase of the task.

3.1.2. Feedback from learner $\hat{\theta}$ and subsequent examples Y_2

When the learner gives the teacher feedback in the form of an estimate $\hat{\theta}$, the teacher can use this information to update their beliefs about what the learner knows:

$$P_T(\phi | Y_1, \hat{\theta}) \propto P_L(\hat{\theta} | \phi, Y_1) P_T(\phi), \tag{5}$$

where $\hat{\theta} | \phi, Y_1$ is treated as a sample from the learner's sampling distribution (Eq. 10).

The teacher then selects more examples Y_2 to send to the learner, again selecting examples using a softmax rule. To calculate utilities, they now marginalize over their updated *posterior* beliefs about the learner's background knowledge:

$$U(Y_2; Y_1, \hat{\theta}, \theta^*) = \int_\phi U(Y_2; \theta^*, Y_1, \phi) P_T(\phi | Y_1, \hat{\theta}) d\phi. \tag{6}$$

This model yields predictions about what the teacher will do in the second phase of the task.

3.1.3. Calculating base learner probabilities

How does the teacher calculate $P_{L_0}(\theta|Y,\phi)$ (Eq. 2) in the context of our experiment? In our experiments, each learner has one of two possible priors $\phi \in \{\phi_o, \phi_p\}$. Before encountering the teacher, learners with the prior ϕ_o saw nine orange turtles and one purple turtle, while learners with the prior ϕ_p saw nine purple turtles and one orange turtle. Thus:

$$\phi \sim \text{Categorical}([\phi_o, \phi_p], \mathbf{p} = [0.5, 0.5]) \tag{7}$$

$$\theta|\phi \sim \begin{cases} \text{Beta}(9, 1) & \text{if } \phi = \phi_o \\ \text{Beta}(1, 9) & \text{if } \phi = \phi_p \end{cases} \tag{8}$$

$$n_{\text{orange}}|\theta, \phi \sim \text{Bin}(n, \theta) \tag{9}$$

where n_{orange} is the number of orange turtles out of n total turtles in the teacher's examples. The base learner reasons about the true island weight θ given examples from the teacher. By conjugacy, the learner's sampling distribution is then

$$\theta|Y, \phi \sim \begin{cases} \text{Beta}\big(9 + n_{\text{orange}}, 1 + n_{\text{purple}}\big) & \text{if } \phi = \phi_o \\ \text{Beta}\big(1 + n_{\text{orange}}, 9 + n_{\text{purple}}\big) & \text{if } \phi = \phi_p \end{cases} \tag{10}$$

The teacher uses this base model of the learner for the "informativity" of their examples in their utility function.

3.2. Methods

3.2.1. Participants

To test the predictions of our model, we recruited a convenience sample of 193 adult participants[2] from the United States from Amazon Mechanical Turk using the CloudResearch interface (Litman, Robinson, & Abberbock, 2017). In all experiments, we recruited participants with an approval rating above 95% and used the CloudResearch approved list to ensure high data quality (Hauser et al., 2022). Participants received a base pay of $5 with a bonus of up to $12 contingent on their performance in the task. The task took roughly 30 min, so participants received an estimated compensation of $15/h. After reading the instructions, participants had to pass a comprehension check before proceeding to the experiment. We did not exclude any participants.[3] Participants gave informed consent, and the Harvard Committee on the Use of Human Subjects approved the experiment.

3.2.2. Procedure

3.2.2.1. Conditions: Participants completed four within-subject conditions. Our primary condition of interest is the "sequential + feedback" condition (Fig. 2A). In this condition, teachers provided two "lessons" (i.e., two different sets of turtles). Before the first lesson, participants were told the ground-truth composition of turtles on the island (e.g., that the island was 70% orange) and given distributional information about the learner's possible prior beliefs (i.e., that the learner had *either* seen 9 out of 10 turtles in the majority color, or 1 out of 10 turtles in the majority color). Thus, teachers knew that the learner's beliefs were skewed somehow, but they did not know exactly how. After teachers provided their first lesson, the learner then provided feedback (in the form of an estimate of the true proportion simulated from the posterior mean derived from the simulated learner model; Eq. 2) before teachers provided a second lesson. This condition allowed us to analyze both the teacher's initial decision without knowledge of the learner's prior as well as how teachers adjusted their decision after receiving feedback from the learner.

Additionally, we studied three control conditions (Fig. 2B). As a baseline for what teachers would do on a second lesson without receiving feedback, the "sequential + no feedback" condition was the same as the primary condition but without the feedback phase in between the two lessons. The other two, nonsequential conditions examined one-shot teaching behavior under different knowledge settings, to check whether participants' behavior corresponds to the model when the prior is explicitly known. In the "non-sequential + certain" condition, teachers were explicitly given the learner's prior (e.g., that they had previously seen 9 out of 10 turtles in the majority color). The "non-sequential + uncertain" condition was equivalent to the first trial of the sequential condition; teachers knew that learners had seen nine turtles of one color and one turtle of the other color, but they did not know which color the learner had seen more of.

3.2.2.2. Experiment design: Within each of the four conditions described above, we also manipulated the learners' priors. In the "too high" condition, participants were told that the learner has previously seen 9 out of 10 turtles in the majority color (relative to the true proportion of 70%). In the "too low" condition, the learner had previously seen only 1 out of 10 turtles in the majority color (relative to the true proportion of 70%). Importantly, the learner's background knowledge was only explicitly shown to participants in one condition (non-sequential + certain); however, the simulated learner's background knowledge influenced the feedback that they provided in sequential conditions, and their ultimate guess in all conditions. Finally, we also counterbalanced the true dominant color of turtles on the island: A 70% of turtles were orange in one island, compared to 30% in the other island. All told, participants taught for a total of 16 trials (4 teaching conditions x 2 learner priors x 2 islands). The trial sequence was fully randomized. Participants were told that they would receive a bonus of $0.70 for each trial if at the end of each trial the simulated learner guessed an island proportion of 8% within the true island proportion, minus a cost of $0.01 for each additional example the participant provided (fixed across all participants, as in the model). No feedback was given at the end of each trial, in order to minimize order effects across the course of the trial sequence and prevent participants from explicitly "learning" to do what the model predicts.

3.3. Results and discussion

Our main question was whether our experimental results are consistent with the predictions of our computational model. A critical feature of our model, which distinguishes it from a nonhierarchical version, is that it represents a teacher's uncertainty about what learners know. It predicts that teachers will not only be sensitive to information about a learner's background knowledge, but also adjust their beliefs based on feedback.

3.3.1. *Teachers are sensitive to learner priors*

Our first hypothesis was that participants would be sensitive to learner priors if they are explicitly stated. To test this hypothesis, we examined the one-shot *non-sequential* conditions (Fig. 2C), where we manipulated the learner's prior biases and displayed them to teachers. Because learner priors were symmetric with respect to the majority color (i.e., having seen 9 out of 10 orange on a 70% orange island is equivalent to having seen 9 out of 10 purple on a 70% purple island), we collapsed across "majority-purple" and "majority-orange" islands by inverting the scale and mapping all numbers to the majority color. We used a logistic mixed-effects regression model to predict the proportion of turtles in the majority color in the teacher's examples (using a binomial linking function to account for the total number in the set). We included categorical fixed effects of the learner's prior ("too high" vs. "too low") and the teacher's knowledge condition ("certain" vs. "uncertain"), along with their interaction, and random intercepts and slopes for each participant. We found evidence for an interaction effect ($b = -0.27$, $z = -8.05$, $p < .001$), as predicted by our computational model (Fig. 2C; Eq. 3). Participants systematically adjusted their examples away from the true proportion to counteract the learner's priors when they were known (the "certain" condition), as shown by "underadjusting" in the "too high" condition, and "overadjusting" in the "too low" condition. For robustness, we also ran simulations showing that the magnitude of adjustment is modulated by both the temperature parameter and the strength of the learner's prior, with higher adjustment in response to stronger learner priors (Fig. A1). These findings indicate that teachers adapt their examples to the learner's background knowledge when it was available.

3.3.2. *Teachers are sensitive to learner feedback*

Our second hypothesis was that participants would be sensitive to learner feedback because it reveals information about learners' priors. To test this hypothesis, we examined the sequential conditions (Fig. 2D). We analyzed each of the learner prior conditions ("too high" and "too low") with separate models for interpretability, because we predicted effects to go in opposite directions (i.e., adjusting up in one case and adjusting down in the other). We used a logistic mixed-effects regression model predicting the proportion of turtles in the majority color in the teacher's examples (again using a binomial linking function to account for the total size of the set), including fixed effects of lesson number ("first" vs. "second") and feedback condition ("feedback" vs. "no feedback"), along with their interaction. We also included random intercepts and slopes for each participant. As predicted by our computational model (Eq. 6), we found evidence of an interaction effect in the "too high" condition ($b = 0.04$, $z = 1.98$, $p = .048$; Fig. 2D, left) as well as in the "too low" condition ($b = -0.11$, $z = -4.38$, $p < .001$; Fig. 2D, right). Without feedback, participants provided similar information on the first and second lessons. The key predictions of our hierarchical model consider how teachers adjust their examples upon receiving information from the learner: Upon receiving feedback, which suggested the direction in which the learner's

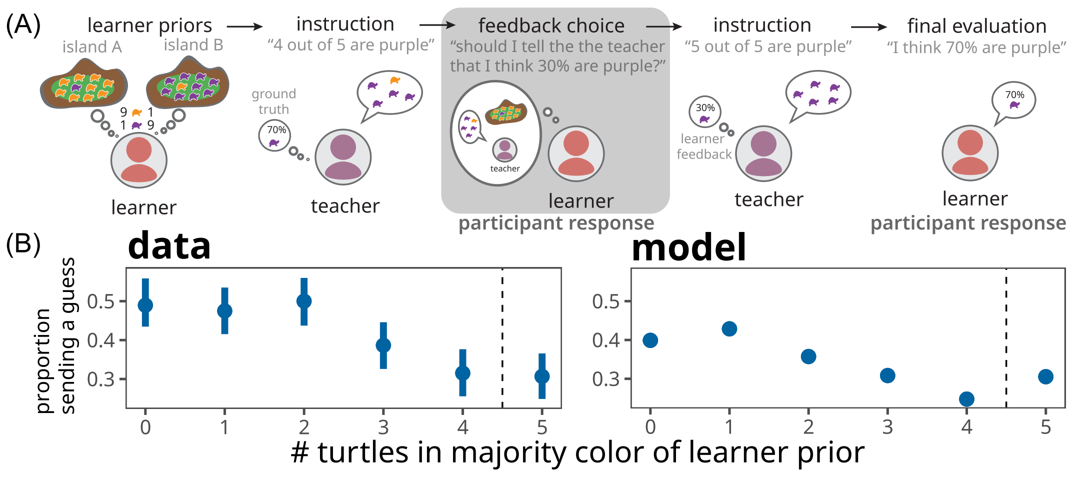

Fig. 3. Experiment 2. (A) Experiment 2 schematic. (B) Experiment 2 results. Teachers send five turtles in total, in any proportion of orange and purple. The proportion of participants (learners) sending a guess decreases as the ratio of turtles in the teacher's examples approaches the ratio corresponding to the learner's prior. The dotted line is at the ratio of the learner's prior (9 turtles of one color, out of 10 turtles). Error bars are bootstrapped 95% confidence intervals.

beliefs were miscalibrated, participants adjusted their examples to correct for the learner biases that were exposed by the feedback, as indicated by "underadjusting" behavior in the "too high" condition and "overadjusting" behavior in the "too low" condition. Similar to in the nonsequential conditions, the simulated magnitude of adjustment is also modulated by both the temperature parameter and the strength of the learner's prior, with higher adjustment in response to stronger learner priors (Fig. A2). These findings indicate that teachers adjust their teaching to learners when they receive feedback from them, in contrast to the predictions of a nonhierarchical model of teaching, which, instead of adjusting to learners' knowledge, would send the same information in the first and second lessons.

4. Experiment 2: Learners adjust their feedback based on teachers' beliefs

In Experiment 1, we found that, in accordance with our model, participants consider learners' priors when they teach and also adjust their teaching based on learners' feedback. In other words, teachers are *learning* from learners' feedback so that they can teach better. Our next experiment investigates how learners may actively decide to *teach* teachers about themselves, to help their teacher teach better.

In Experiment 2 (Fig. 3A), participants played the role of a tourist, evaluating examples from a tour guide (i.e., a simulated teacher). Our primary question concerns how learners evaluate whether to share information about themselves to a teacher. In accordance with the predictions of our model, we predicted that participants would use examples selected by the teacher to evaluate what the teacher knows about them and to decide whether to send feedback.

4.1. Model for Experiment 2

For Experiment 2, we consider the case where the participant is the learner L. Their goal is to get to the concept that the teacher is trying to teach them. They choose whether to give feedback to the teacher after seeing the first set of examples from the teacher. While in Experiment 1 the simulated learner's decision to give feedback is fixed, in Experiment 2, our primary interest is how the learner adaptively decides whether to send feedback in the first place, based on what they have seen from their teacher.

4.1.1. Receiving examples from the teacher

The participant receives examples from a teacher and interprets these examples as coming from a teacher who pedagogically chooses examples for a base learner. Let τ be what the learner knows about the teacher's uncertainty $P_T(\phi)$. After receiving a set of examples Y_1 from the teacher, the learner jointly reasons about how much the teacher knows about them τ and what the true concept θ is. By Bayes' rule:

$$P_L(\theta, \tau | Y_1, \phi) \propto P_T(Y_1 | \theta, \phi, \tau) P(\theta) P(\tau), \tag{11}$$

where $P_T(Y_1|\theta, \phi, \tau)$ is the learner's model of the teacher. In our experiments, the teacher either knows ($\tau = \texttt{certain}$) or does not know ($\tau = \texttt{uncertain}$) what the learner knows:

$$P_T(Y|\theta, \tau, \phi) \propto \begin{cases} \exp(\alpha_T \cdot U(Y; \theta, \phi)) & \text{if } \tau = \texttt{certain} \\ \exp(\alpha_T \cdot \int_\phi U(Y; \theta, \phi) P_T(\phi))\, d\phi & \text{if } \tau = \texttt{uncertain} \end{cases} \tag{12}$$

The two possible learner priors $\phi \in \{\phi_o, \phi_p\}$ are the same as in Experiment 1. In both experiments, we set uniform priors over θ and ϕ.

4.1.2. Decision to send feedback

If the learner gives feedback, the teacher can use this information to update their beliefs about the learner's priors, which can help the learner get to the true concept after the teacher's second set of examples. So the learner evaluates the utility of giving feedback, weighing its costs and rewards:

$$U(\texttt{feedback}; Y_1, \phi) = (1 - c) \cdot \log P_L(\texttt{correct}|Y_1, \phi, \texttt{feedback}) - c \cdot \mathbb{I}(\texttt{feedback}) \tag{13}$$

where c controls the relative weighting of cost and reward. The learner then chooses to give feedback using a softmax rule, modulated by the learner's inverse temperature parameter α_L:

$$P_L(\texttt{feedback}|Y_1, \phi) \propto \exp(\alpha_L \cdot U(\texttt{feedback}; Y_1, \phi)). \tag{14}$$

To calculate the "informativity" of feedback, the learner calculates the benefit of feedback for each possible true concept (by comparing it with their final judgment of the island proportion θ_{final}), and marginalizes over what the concepts can be:

$$P_L(\texttt{correct}|Y_1, \phi, \texttt{feedback}) = \int_{\theta^*} P_L(\theta_{\text{final}} = \theta^*|Y_1, \phi, \texttt{feedback}, \theta^*) P(\theta^*)\, d\theta^*. \quad (15)$$

To calculate the benefit of feedback for each concept, the learner marginalizes over all the possible knowledge levels of the teacher, all the possible learners that each of those teachers may have in mind, and all of the possible second set of examples Y_2 that each of those teachers might provide:

$$P_L(\theta_{\text{final}}|Y_1, \phi, \texttt{feedback = yes}) =$$
$$\sum_{Y_2} \left(P_L(\theta_{\text{final}}|Y_1, Y_2, \phi) \int_\tau \int_\phi P_T(Y_2|Y_1, \phi, \tau) P_T(\phi|\hat{\theta}_L, Y_1, \tau) P_L(\tau|Y_1, \phi)\, d\tau\, d\phi \right). \quad (16)$$

4.2. Methods

4.2.1. Participants

For Experiment 2, we recruited a convenience sample of 119 adult participants[4] from the United States from Amazon Mechanical Turk using the CloudResearch interface (Litman et al., 2017). Participants received a base pay of \$3 with a bonus of up to \$6 contingent on their performance in the task. The task took roughly 15 min, so participants received an estimated compensation of \$15/h. As in Experiment 1, participants had to pass a comprehension check before proceeding to the experiment. We did not exclude any participants. Participants gave informed consent, and the Harvard Committee on the Use of Human Subjects approved the experiment.

4.2.2. Procedure

Participants were placed in the role of a tourist learning about the proportion of purple and orange turtles on the island (Fig. 3A). Each trial consisted of two lessons, where the (simulated) teacher sent a set of up to five turtles in any combination of orange and purple colors. At the beginning of each trial, participants were explicitly shown their prior knowledge (i.e., how many orange and purple turtles they had already seen on the island). Participants were informed that some teachers knew about this prior knowledge, and some teachers did not. Critically, after the first lesson, participants could send feedback to the teacher for a cost of \$0.05. If participants opted to send feedback, they could send teachers a message before the second lesson to indicate whether they thought it was more likely that the island had 30% orange or 70% orange turtles. For the second lesson, the simulated teacher used the learner's feedback (if available) to update their beliefs about the learner's prior knowledge, and then selected a second set of examples to maximize the learner's belief in the true island proportion (Eq. 16). After all examples had been presented, we asked participants to report whether they thought the island's turtle population was composed of 30% orange and 70% purple turtles or 70% orange and 30% purple turtles. Participants were awarded a bonus of $0.50 if their final guess was correct.

We manipulated two factors within a fully within-subjects design. First, we manipulated the examples that teachers provided in the first lesson. Teachers sent learners five turtles total, in a combination of orange and purple. There were six possible sets of examples that teachers could provide in all, ranging from 0 orange and 5 purple turtles to 5 orange and 0 purple turtles. Additionally, we manipulated the learner's prior. Before interacting with the teacher, learners were shown a sample of nine purple turtles and one orange turtle in one condition, and nine orange turtles and one purple turtle in the other.

Altogether, each session contained 12 trials (2 learner priors x 6 sets of examples), presented in a a fully randomized sequence. Within each trial, we randomly sampled how much teachers knew about the learner. The simulated teacher either knew precisely which turtles the learner saw (akin to the "non-sequential" + "certain" condition in Experiment 1), or knew that the learner saw nine turtles of one color and one turtle of another color but were unsure whether the majority color in the learner's prior aligned with the true majority color in the population (akin to the "non-sequential" + "uncertain" condition in Experiment 1). While this manipulation did not affect what participants saw before the critical trial (their decision for whether to send feedback), it did affect the simulated teacher's second set of examples. Additionally, we randomly sampled the true island proportion, which was either 30% orange turtles or 70% orange turtles. Thus, both of these manipulations affected the examples given to the learner on the second lesson, which were generated from the model based on participants' guess decisions and the guesses they provided. Like in Experiment 1, participants did not receive information about their accumulated bonus until the end of the experiment.

4.3. *Results and discussion*

Our model, unlike a nonhierarchical model, predicts that learners will be sensitive to how much they think the teacher knows about them, and use this information to strategically decide when to send feedback.

4.3.1. *Exploratory analyses*

First, we report the results of some exploratory analyses relating to learners' performance on the task. These analyses were not preregistered.

4.3.1.1. Learners successfully inferred the true proportion after receiving the second set of examples: In a logistic mixed-effects regression model predicting final guess accuracy from lesson number, with random intercepts and slopes for each participant, we found that learners made more accurate inferences about the true proportion of turtles after the second set of examples (65.8% correct) compared to their guess after only the first set of examples (baseline; 49.7% correct; $b = 0.35$, $z = 6.65$, $p < .001$), suggesting that our model-generated simulated teacher was providing helpful examples during the second phase of the task.

4.3.1.2. Learners performed better when they provided a guess: In a logistic mixed-effects regression predicting final guess accuracy, with categorical fixed effects of the teacher's knowledge ("full" vs. "partial") and whether the participant sent a guess after the first set of examples, with random intercepts for each participant, we found a main effect of whether the participant provided a guess ($b = 0.71$, $z = 5.89$, $p < .001$). We did not find evidence of a main effect of the teacher's knowledge ($b = 0.04$, $z = 0.60$, $p = .550$) or evidence of an interaction effect ($b = -0.17$, $z = -1.40$, $p = .159$). These results suggest that participants who provided a guess did better on the task, regardless of the knowledge level of the teacher. Additionally, when learners did provide guesses, they provided guesses in accordance with what the model predicts (Fig. A3).

4.3.2. Confirmatory analysis

Now that we have verified that participants indeed perform better after seeing the teacher's second set of examples (and after providing a guess), and that their final guesses matched the model's predictions, we turn to our key predictions about when participants should choose to provide feedback in the first place.

4.3.2.1. Learners were less likely to send a guess when teachers' examples corresponded to their priors: Our model (learners sending a guess; Eq. 14) predicts that the probability of sending a guess decreases as teachers provide more turtles in the majority color of learners' priors (Fig. 3B). Thus, we used a logistic mixed-effects regression model to predict learners' guess decisions as a function of the number of turtles teachers provided in the majority color of learners' priors, with random intercepts and slopes for each participant. Because our model predicts a nonmonotonic pattern for some parameter ranges (Figs. A4 and A5), we coded number of turtles as a quadratic term. As in Experiment 1, we collapsed across "majority-purple" and "majority-orange" learner priors by mapping all numbers to the majority color. We found a main effect of both polynomial terms ($b_x = -35.83$, $z = -5.18$, $p < .001$; $b_{x^2} = -14.33$, $z = -3.00$, $p = .003$), suggesting that the probability of sending a guess decreases as teachers provide more turtles in the majority color of learners' priors. Thus, unlike the predictions of a nonhierarchical model (which would not be able to use the teacher's examples to plan ahead, reason about the teacher's knowledge, and anticipate whether to send a guess to them), participants in our experiment used their teachers' first set of examples to decide how to communicate with the teacher, to help the teacher teach better so that participants could get to the right answer.

5. General discussion

In this paper, we formalized a hierarchical Bayesian model of pedagogy that integrates (1) how a teacher and a learner reason about each others' minds to decide whether and how to share information, and (2) how a teacher and a learner communicate to resolve uncertainty not only about the target concepts, but also about each other. We designed two experiments to test the predictions of this model. Experiment 1 tested the prediction that teachers are sensitive to learners' priors and adapt their teaching based on learners' feedback. Experiment 2 tested the prediction that learners decide whether to give feedback to their teachers by using their teachers' examples to reason about how much the teacher knows about them and what the right answer might be. For both experiments, we found that our model accounted for the patterns in participants' responses.

Past work has modeled adaptive teaching as a sequential decision-making process (often without theory of mind), where teachers can trade off between "teaching" and "probing" to make inferences about the learner's knowledge state, which the teacher may not know anything about (Çelikok, Murena, & Kaski, 2023; Fan, Tian, Qin, Li, & Liu, 2018; Frank & Liu, 2018; Johns, Mac Aodha, & Brostow, 2015; Koedinger, Brunskill, Baker, McLaughlin, & Stamper, 2013; Peltola, Çelikok, Daee, & Kaski, 2019; Popp & Gureckis, 2020; Rafferty, Brunskill, Griffiths, & Shafto, 2016; Whitehill & Movellan, 2017; Wang, Wu, & Goodman, 2022; Yuan et al., 2021). Our hierarchical Bayesian approach is complementary to these sequential approaches. We aim to explain how teachers and learners maintain and resolve *structured uncertainty* about each other. A teacher may know something about the overall distribution of learners without knowing exactly what an individual learner knows. Nonetheless, the teacher's priors about different kinds of learners *in general* helps them more quickly narrow down what individual learners need. Incorporating such reasoning processes into machine teaching systems can aid the construction of novel pedagogical tools (Chandra, Chen, Li, Ragan-Kelley, & Tenenbaum, 2024; Chandra, Li, Nigam, Tenenbaum, & Ragan-Kelley, 2024; Ross & Andreas, 2024).

In the current work, we model the cognitive mechanisms underlying an idealized teaching scenario, where teachers and learners are trying to reach the same goal (getting the learner to the correct concept) by rationally incorporating the information they receive from their partners. However, in practice, real-world teaching may not always follow these assumptions. Teachers' and students' goals may not be aligned (Lemos, 1996), and students may not be aware of teachers' learning objectives (Wijngaards-de Meij & Merx, 2018). It may be hard for teachers to accurately monitor students' understanding (Chi et al., 2001) or diagnose what students know and do not know in the first place (Van de Pol, Volman, & Beishuizen, 2011). Socioeconomic status, immigrant background, gender, and race all affect how teachers perceive students' motivation and skills (Brandmiller, Dumont, & Becker, 2020; Redding, 2019), and may affect how teachers and learners interpret and process information from each other (Narciss, 2008).

While the proportion estimation task allowed us to directly evaluate our model predictions on exactly the same task, we highlight additional challenges facing future work scaling up this framework to more naturalistic scenarios. First, the structure of the learner population and real-world concepts are more complex than the 2-component mixture we assumed. For example, some domains may be well-captured by a dependency graph (Kemp & Tenenbaum, 2008): observing a student solve a particular differential equation implies that it is likely they already know how to do basic algebra and arithmetic.

Second, real-world interactions are longer and richer than a single exchange of examples and feedback. Misunderstandings are only revealed across patterns of errors unfolding in a multistep process and require integrating information sequentially over time (e.g., Rafferty, Jansen, & Griffiths, 2020). Teaching with language rather than examples or demonstrations enables richer strategies for probing abstract knowledge and providing feedback (Sumers, Ho, Hawkins, Narasimhan, & Griffiths, 2021; Sumers, Ho, Hawkins, & Griffiths, 2023). In practice, teachers are encouraged to communicate conceptual or procedural knowledge, or meta-cognitive strategies (Narciss, 2008).

Finally, our framework models knowledge states as beliefs in the right concept, and learners' "background knowledge" as their immediate prior experience with the concept. However, background knowledge can also come in a variety of other forms, including visual access (Hawkins, Gweon, & Goodman, 2021), latent motives such as greediness or risk aversion (van Baar, Nassar, Deng, & FeldmanHall, 2022), or salient points in the hypothesis space itself (Aboody, Velez-Ginorio, Santos, & Jara-Ettinger, 2023). In addition to "background knowledge," teachers are also encouraged to learn about properties of their students like learner engagement or affect (Arroyo et al., 2014), and engage in strategies like observing learners apply what they have been taught (Okita & Schwartz, 2013). Tailoring to learners' speed of learning may even be more helpful than tailoring to explicit knowledge (Yudelson, Koedinger, & Gordon, 2013), and learners may adapt their inferences based on teachers' teaching styles (Bass, Shafto, & Bonawitz, 2018). Thus, knowledge of a variety of different cognitive mechanisms involved in learning may aid teachers to better help and adapt to students (Koedinger, Corbett, & Perfetti, 2012). Future work building tools to scale up adaptive reasoning processes (e.g., by incorporating natural language and more complex teacher–learner interactions) should incorporate a diverse range of data sources, classroom interventions, and collaborations with teachers and education researchers to enhance our understanding of these processes in real-world settings (e.g., Demszky, Wang, Geraghty, & Yu, 2024; Wang, Wirawarn, Goodman, & Demszky, 2023; Wang, Zhang, Robinson, Loeb, & Demszky, 2024).

Bayesian theory of mind models provide a common computational framework that can be used to describe social learning and teaching in a variety of contexts, from a variety of inputs. For example, by modifying teachers' and learners' utility functions, our computational framework can easily be extended to capture additional goals (e.g., reputational goals; Radkani, Tenenbaum, & Saxe, 2022; Yoon, Tessler, Goodman, & Frank, 2020) that can be at stake in educational settings. However, the downside of reasoning about other people's mental states and tracking these beliefs over time is that it is computationally costly (Apperly & Butterfill, 2009; Hawkins et al., 2021; Ho, Saxe, & Cushman, 2022; Wu, Vélez, & Cushman, 2022). For example, participants in Experiment 2 would need to reason about all the alternative beliefs their teacher might possibly have, all the alternative concepts each of those teachers might possibly want to teach, and all the pieces of information those teachers may possibly provide. In certain situations, people may be deploying more habitual or automatic strategies, instead of using recursive theory of mind (e.g., Bass, Espinoza, Bonawitz, & Ullman, 2024; FeldmanHall & Shenhav, 2019; Gershman, Gerstenberg, Baker, & Cushman, 2016).

Although we deliberately selected a simple class of concepts where Bayesian inference is tractable (due to conjugacy), we found evidence of more heuristic strategies. In Experiment 1 (Fig. 2C,D), participants uncertain about learner priors defaulted to sending the true proportion of 70% instead of precisely marginalizing over all the possible knowledge levels of the learner. Thus, if the teacher does not know what the learner knows, the teacher may assume that providing the answer directly may lead the learner to directly update to their stated proportion. For sequential interactions, this leads to downstream divergences between participants and the model: Because the model initially "overcompensates" in the "too low" condition due to exact marginalization, it then "undercompensates" for the second lesson, as it now has access to more examples to counteract learners' priors and teach the exact proportion of 70% (Fig. 2D, right). For more complex concepts, or for multiple rounds of interactions, doing exact Bayesian inference may be intractable.[5] An important direction for future work, then, is to develop resource-rational process models that approximate Bayesian inference in repeated, back-and-forth interactions.

5.1. Conclusion

Good teaching is good communication. In the past, theory of mind has been used to explain how humans teach and learn from one another in a variety of contexts, including demonstrations (Ho, Cushman, Littman, & Austerweil, 2021), verbal descriptions (Sumers, Hawkins, Ho, Griffiths, & Hadfield-Menell, 2022), and carefully chosen examples (Shafto et al., 2014). However, past work has largely studied these processes as one-sided interactions: Teachers either provide information to a naïve learner without receiving feedback (e.g., Shafto et al., 2014), or learners, who often already know the ground truth, evaluate teachers without communicating their evaluation back to the teacher (e.g., Aboody, Huey, & Jara-Ettinger, 2022; Bass et al., 2022; Gweon, Pelton, Konopka, & Schulz, 2014; Gweon & Asaba, 2018). In more realistic contexts, teaching requires a constant give and take. Teachers probe their students so that they can effectively teach, and learners communicate what they need so that they can effectively learn. By integrating Bayesian models of teaching with hierarchical models of communication, our work provides a theoretical and empirical foundation for understanding how teachers and learners effectively adapt to one another.

Acknowledgments

Thanks to Rebecca Saxe, Josh Tenenbaum, and members of the Computational Cognitive Neuroscience lab and Saxelab for helpful discussions and feedback. We are also grateful to the three anonymous reviewers for their thoughtful and detailed comments. This work was supported by the Center for Brains, Minds, and Machines (CBMM) funded by NSF STC award CCF-1231216 (S.J.G.), and an NIMH K00 award K00MH125856 (N.V.). We have no conflicts of interest to disclose.

Open Research Badges

This article has earned Open Data and Open Materials badges. All materials, code, and deidentified data are available at https://github.com/aliciamchen/hierarchical-teaching. Data and materials are available at https://osf.io/ubxjr. Both experiments were preregistered; the preregistration for Experiment 1 is available at https://osf.io/7xbcg and the preregistration for Experiment 2 is available at https://osf.io/utmxq.

Notes

1 A limitation of this setup that affects both humans and models, to be addressed in future work, is that the space of proportions that can be conveyed depends on the total number of (cost-incurring) examples sent (e.g., a teacher may only need to send 10 examples to precisely convey 70%, but needs to pay for more examples to convey 65%).

2 We note that this sample size deviated from our preregistered pre-exclusion sample size of 120 participants, which was estimated to achieve roughly 80% power based on a post-hoc power analysis on pilot data. Because we were not sure how many participants would pass the exclusion criteria, we erred on the size of a large sample size.

3 We included exclusion criteria in both preregistrations, and indicated that we would conduct our analyses with and without exclusions.

4 Our preregistered sample size was 120 participants, estimated to achieve roughly 80% power based on a post-hoc power analysis on pilot data.

5 We piloted experiments with more complex hypothesis spaces, including a version of Experiment 1 and an interactive human–human experiment. Our results were inconclusive. More work should be done to bridge the gap between idealized models of rational mental-state reasoning, and interactive teaching in more complex and naturalistic scenarios.

References

Aboody, R., Huey, H., & Jara-Ettinger, J. (2022). Preschoolers decide who is knowledgeable, who to inform, and who to trust via a causal understanding of how knowledge relates to action. *Cognition*, *228*, 105212.

Aboody, R., Velez-Ginorio, J., Santos, L. R., & Jara-Ettinger, J. (2023). When naïve pedagogy breaks down: Adults rationally decide how to teach, but misrepresent learners' beliefs. *Cognitive Science*, *47*(3), e13257.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*(4), 953.

Arroyo, I., Woolf, B. P., Burelson, W., Muldner, K., Rai, D., & Tai, M. (2014). A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. *International Journal of Artificial Intelligence in Education*, *24*, 387–426.

Bass, I., Bonawitz, E., Hawthorne-Madell, D., Vong, W. K., Goodman, N. D., & Gweon, H. (2022). The effects of information utility and teachers' knowledge on evaluations of under-informative pedagogy across development. *Cognition*, *222*, 104999.

Bass, I., Espinoza, C., Bonawitz, E., & Ullman, T. (2024). Teaching without thinking: Negative evaluations of rote pedagogy. *Cognitive Science*, *48*(6), e13470.

Bass, I., Shafto, P., & Bonawitz, E. (2018). That'll teach 'em: How expectations about teaching styles may constrain inferences. In *Proceedings of the Annual Conference of the Cognitive Science Society*.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, *120*(3), 322–330.

Brandmiller, C., Dumont, H., & Becker, M. (2020). Teacher perceptions of learning motivation and classroom behavior: The role of student characteristics. *Contemporary Educational Psychology*, *63*, 101893.

Bridgers, S., Jara-Ettinger, J., & Gweon, H. (2020). Young children consider the expected utility of others' learning to decide what to teach. *Nature Human Behaviour*, *4*(2), 144–152.

Carless, D. (2022). From teacher transmission of information to student feedback literacy: Activating the learner role in feedback processes. *Active Learning in Higher Education*, *23*(2), 143–153.

Çelikok, M. M., Murena, P.-A., & Kaski, S. (2023). Teaching to learn: Sequential teaching of learners with internal states. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37 (pp. 5939–5947).

Chandra, K., Chen, T., Li, T.-M., Ragan-Kelley, J., & Tenenbaum, J. (2024). Cooperative explanation as rational communication.

Chandra, K., Li, T.-M., Nigam, R., Tenenbaum, J., & Ragan-Kelley, J. (2024). Watchat: Explaining perplexing programs by debugging mental models. *arXiv preprint arXiv:2403.05334*.

Chi, M. T., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, *25*(4), 471–533.

Clark, H. H., & Marshall, C. R. (1981). Definite knowledge and mutual knowledge.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. Irvington.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153.

De Leeuw, J. R. (2015). jspsych: A javascript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*(1), 1–12.

Demszky, D., Wang, R., Geraghty, S., & Yu, C. (2024). Does feedback on talk time increase student engagement? Evidence from a randomized controlled trial on a math tutoring platform. In *Proceedings of the 14th Learning Analytics and Knowledge Conference* (pp. 632–644).

Fan, Y., Tian, F., Qin, T., Li, X.-Y., & Liu, T.-Y. (2018). Learning to teach. *arXiv preprint arXiv:1805.03643*.

FeldmanHall, O., & Shenhav, A. (2019). Resolving uncertainty in a social world. *Nature Human Behaviour*, *3*(5), 426–435.

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998–998.

Frank, M. C., & Liu, L. (2018). Modeling classroom teaching as optimal communication. *PsyArXiv*.

Fyfe, E. R., & Rittle-Johnson, B. (2016). Feedback both helps and hinders learning: The causal role of prior knowledge. *Journal of Educational Psychology*, *108*(1), 82.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. Chapman and Hall/CRC.

Gershman, S. J., Gerstenberg, T., Baker, C. L., & Cushman, F. A. (2016). Plans, habits, and theory of mind. *PLoS One*, *11*(9), e0162246.

Gershman, S. J., & Niv, Y. (2015). Novelty and inductive generalization in human reinforcement learning. *Topics in Cognitive Science*, *7*(3), 391–415.

Goodboy, A. K., & Myers, S. A. (2008). The effect of teacher confirmation on student communication and learning outcomes. *Communication Education*, *57*(2), 153–179.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818–829.

Goodman, N. D., & Stuhlmüller, A. (2014). The design and implementation of probabilistic programming languages. http://dippl.org.

Goodman, N. D., Ullman, T. D., & Tenenbaum, J. B. (2011). Learning a theory of causality. *Psychological Review*, *118*(1), 110.

Gweon, H., & Asaba, M. (2018). Order matters: Children's evaluation of underinformative teachers depends on context. *Child Development*, *89*(3), e278–e292.

Gweon, H., Pelton, H., Konopka, J. A., & Schulz, L. E. (2014). Sins of omission: Children selectively explore when teachers are under-informative. *Cognition*, *132*(3), 335–341.

Hao, X., Jhaveri, Y., & Shafto, P. (2024). Common ground in cooperative communication. *Advances in Neural Information Processing Systems*, *36*, 5211–5222.

Hauser, D. J., Moss, A. J., Rosenzweig, C., Jaffe, S. N., Robinson, J., & Litman, L. (2022). Evaluating CloudResearch's approved group as a solution for problematic data quality on MTurk. *Behavior Research Methods*, *55*(8), 3953–3964.

Hawkins, R. D., Franke, M., Frank, M. C., Goldberg, A. E., Smith, K., Griffiths, T. L., & Goodman, N. D. (2022). From partners to populations: A hierarchical Bayesian account of coordination and convention. *Psychological Review*, *130*(4), 977.

Hawkins, R. D., Gweon, H., & Goodman, N. D. (2021). The division of labor in communication: Speakers help listeners account for asymmetries in visual perspective. *Cognitive Science*, *45*(3), e12926.

Ho, M. K., Cushman, F., Littman, M. L., & Austerweil, J. L. (2021). Communication in action: Planning and interpreting communicative demonstrations. *Journal of Experimental Psychology: General*, *150*(11), 2246.

Ho, M. K., Littman, M., MacGlashan, J., Cushman, F., & Austerweil, J. L. (2016). Showing versus doing: Teaching by demonstration. *Advances in Neural Information Processing Systems*, *29*, 3027–3035.

Ho, M. K., Saxe, R., & Cushman, F. (2022). Planning with theory of mind. *Trends in Cognitive Sciences*, *26*(11), 959–971.

Hoy, A. W., Davis, H., & Pape, S. J. (2006). Teacher knowledge and beliefs.

Isaacs, E. A., & Clark, H. H. (1987). References in conversation between experts and novices. *Journal of Experimental Psychology: General*, *116*(1), 26.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, *20*(8), 589–604.

Johns, E., Mac Aodha, O., & Brostow, G. J. (2015). Becoming the expert-interactive multi-class machine teaching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2616–2624).

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, *19*, 509–539.

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science*, *10*(3), 307–321.

Kemp, C., & Tenenbaum, J. B. (2008). The discovery of structural form. *Proceedings of the National Academy of Sciences*, *105*(31), 10687–10692.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*(2), 148.

Kline, M. A. (2015). How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behavioral and Brain Sciences*, *38*, e31.

Koedinger, K. R., Brunskill, E., Baker, R. S., McLaughlin, E. A., & Stamper, J. (2013). New potentials for data-driven intelligent tutoring system development and optimization. *AI Magazine*, *34*(3), 27–41.

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The knowledge-learning-instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science*, *36*(5), 757–798.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmertest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1–26.

Lemos, M. S. (1996). Students' and teachers' goals in the classroom. *Learning and Instruction*, *6*(2), 151–171.

Lindsey, R. V., Shroyer, J. D., Pashler, H., & Mozer, M. C. (2014). Improving students' long-term knowledge retention through personalized review. *Psychological Science*, *25*(3), 639–647.

Litman, L., Robinson, J., & Abberbock, T. (2017). Turkprime. com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, *49*(2), 433–442.

Narciss, S. (2008). Feedback strategies for interactive learning tasks. In David Jonassen, Michael J. Spector, Marcy Driscoll, M. David Merrill, Jeroen van Merrienboer, Marcy P. Driscoll. *Handbook of research on educational communications and technology* (pp. 125–143). Routledge.

Okita, S. Y., & Schwartz, D. L. (2013). Learning by teaching human pupils and teachable agents: The importance of recursive feedback. *Journal of the Learning Sciences*, *22*(3), 375–412.

Peltola, T., Çelikok, M. M., Daee, P., & Kaski, S. (2019). Machine teaching of active sequential learners. *Advances in Neural Information Processing Systems*, *32*, 11204–11215.

Popp, P. O., & Gureckis, T. M. (2020). Ask or tell: Balancing questions and instructions in intuitive teaching. In *CogSci*.

Radkani, S., Tenenbaum, J., & Saxe, R. (2022). Modeling punishment as a rational communicative social action. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 44.

Rafferty, A. N., Brunskill, E., Griffiths, T. L., & Shafto, P. (2016). Faster teaching via POMDP planning. *Cognitive Science*, *40*(6), 1290–1332.

Rafferty, A. N., Jansen, R. A., & Griffiths, T. L. (2020). Assessing mathematics misunderstandings via Bayesian inverse planning. *Cognitive Science*, *44*(10), e12900.

Redding, C. (2019). A teacher like me: A review of the effect of student–teacher racial/ethnic matching on teacher perceptions of students and student academic and behavioral outcomes. *Review of Educational Research*, *89*(4), 499–535.

Rocca, K. A. (2010). Student participation in the college classroom: An extended multidisciplinary literature review. *Communication Education*, *59*(2), 185–213.

Ross, A., & Andreas, J. (2024). Toward in-context teaching: Adapting examples to students' misconceptions. *arXiv preprint arXiv:2405.04495*.

Shafto, P., & Goodman, N. (2008). Teaching games: Statistical sampling assumptions for learning in pedagogical situations. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 1632–1637). Austin, TX: Cognitive Science Society.

Shafto, P., Goodman, N. D., & Frank, M. C. (2012). Learning from others: The consequences of psychological reasoning for human learning. *Perspectives on Psychological Science*, *7*(4), 341–351.

Shafto, P., Goodman, N. D., & Griffiths, T. L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology*, *71*, 55–89.

Slavich, G. M., & Zimbardo, P. G. (2012). Transformational teaching: Theoretical underpinnings, basic principles, and core methods. *Educational Psychology Review*, *24*(4), 569–608.

Staub, F. C., & Stern, E. (2002). The nature of teachers' pedagogical content beliefs matters for students' achievement gains: Quasi-experimental evidence from elementary mathematics. *Journal of Educational Psychology*, *94*(2), 344.

Sumers, T. R., Hawkins, R. D., Ho, M. K., Griffiths, T. L., & Hadfield-Menell, D. (2022). How to talk so AI will learn: Instructions, descriptions, and autonomy. In *Advances in Neural Information Processing Systems*.

Sumers, T. R., Ho, M. K., Hawkins, R. D., & Griffiths, T. L. (2023). Show or tell? Exploring when (and why) teaching with language outperforms demonstration. *Cognition*, *232*, 105326.

Sumers, T. R., Ho, M. K., Hawkins, R. D., Narasimhan, K., & Griffiths, T. L. (2021). Learning rewards from linguistic feedback. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume *35* (pp. 6002–6010).

van Baar, J. M., Nassar, M. R., Deng, W., & FeldmanHall, O. (2022). Latent motives guide structure learning during adaptive social choice. *Nature Human Behaviour*, *6*(3), 404–414.

Van de Pol, J., Volman, M., & Beishuizen, J. (2011). Patterns of contingent teaching in teacher–student interaction. *Learning and Instruction*, *21*(1), 46–57.

van de Pol, J., Volman, M., & Beishuizen, J. (2012). Promoting teacher scaffolding in small-group work: A contingency perspective. *Teaching and Teacher Education*, *28*(2), 193–205.

VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. B. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction*, *21*(3), 209–249.

Vélez, N., Chen, A. M., Burke, T., Cushman, F. A., & Gershman, S. J. (2023). Teachers recruit mentalizing regions to represent learners' beliefs. *Proceedings of the National Academy of Sciences*, *120*(22), e2215015120.

Vygotsky, L. S. (1978). Mind in society: The development of higher mental processes (E. Rice, ed., & trans.).

Wang, R. E., Wirawarn, P., Goodman, N., & Demszky, D. (2023). Sight: A large annotated dataset on student insights gathered from higher education transcripts.

Wang, R. E., Wu, M., & Goodman, N. (2022). Know thy student: Interactive learning with Gaussian processes. *arXiv preprint arXiv:2204.12072*.

Wang, R. E., Zhang, Q., Robinson, C., Loeb, S., & Demszky, D. (2024). Bridging the novice-expert gap via models of decision-making: A case study on remediating math mistakes.

Whitehill, J., & Movellan, J. (2017). Approximately optimal teaching of approximately optimal learners. *IEEE Transactions on Learning Technologies*, *11*(2), 152–164.

Wijngaards-de Meij, L., & Merx, S. (2018). Improving curriculum alignment and achieving learning goals by making the curriculum visible. *International Journal for Academic Development*, *23*(3), 219–231.

Wu, C. M., Vélez, N., & Cushman, F. A. (2022). Representational exchange in human social learning. In *The drive for knowledge: The science of human information seeking*, 169.

Yang, S. C.-H., Vong, W. K., Sojitra, R. B., Folke, T., & Shafto, P. (2021). Mitigating belief projection in explainable artificial intelligence via Bayesian teaching. *Scientific Reports*, *11*(1), 1–17.

Yang, S. C.-H., Vong, W. K., Yu, Y., & Shafto, P. (2019). A unifying computational framework for teaching and active learning. *Topics in Cognitive Science*, *11*(2), 316–337.

Yoon, E. J., Tessler, M. H., Goodman, N. D., & Frank, M. C. (2020). Polite speech emerges from competing social goals. *Open Mind*, *4*, 71–87.

Yuan, L., Zhou, D., Shen, J., Gao, J., Chen, J. L., Gu, Q., Wu, Y. N., & Zhu, S.-C. (2021). Iterative teacher-aware learning. *Advances in Neural Information Processing Systems*, *34*, 29231–29245.

Yudelson, M. V., Koedinger, K. R., & Gordon, G. J. (2013). Individualized Bayesian knowledge tracing models. In *Artificial Intelligence in Education: 16th International Conference, AIED 2013, Memphis, TN, USA, July 9-13, 2013. Proceedings 16* (pp. 171–180). Springer.

Zhu, X. (2015). Machine teaching: An inverse problem to machine learning and an approach toward optimal education. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 29.

Zhu, X., Singla, A., Zilles, S., & Rafferty, A. N. (2018). An overview of machine teaching. *arXiv preprint arXiv:1801.05927*.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. A1. Simulation results for the "non-sequential" condition in Experiment 1, under a range of values for the teacher's softmax temperature parameter α_T, and the weight on the learner's prior (i.e., multiplier on total number of turtles seen, while keeping the proportion constant). Fig. A2. Simulation results for the "sequential" conditions in Experiment 1 under a range of parameter values. Fig. A3. Accuracy of first guesses in Experiment 2 in response to examples from the teacher. Fig. A4. Simulation results for Experiment 2 (Fig. 3) under a range of values for the cost weight c and teachers' and learners' softmax temperature parameters α_T and α_L. Fig. A5. Simulation results for Experiment 2 (Fig. 3) under a range of values for the cost weight c and teachers' and learners' softmax temperature parameters α_T and α_L.