Learning about Others and Learning from Others: Bayesian Probabilistic Models of Intuitive Psychology and Social Learning

Rongzhi Liu

Fei Xu

University of California, Berkeley

In press, In J. Lockman (Ed.), *Advances in Child Development and Behavior*, Vol. 63. Waltham, MA: Academic Press.

Word count: 11332 (text)

Please address correspondence to Rongzhi Liu, 2121 Berkeley Way, University of California, Berkeley. Berkeley, CA 94720; Email: rongzhi_liu@berkeley.edu

1. Introduction

Infants are born into a world full of other individuals. From the beginning of life, infants are interested in other people (e.g., newborns prefer looking at human faces; Johnson et al., 1991). As they observe and interact with them, infants gradually come to understand other individuals. They learn that individuals' actions are driven by unobservable and distinct mental states. They make inferences about others' mental states from their actions. To what extent are these reasoning and inferences rational?

Other people are also an important source to learn from. Humans possess powerful social learning abilities that allow the accumulation of rich knowledge and skills over generations (Boyd et al., 2011; Tomasello, 2016). What kinds of mechanisms underlie our abilities to learn from other people and teach other people? How do we decide whom to learn from? To what extent are these inferences rational?

The purpose of this chapter is to review a growing body of Bayesian probabilistic models on intuitive psychology and social learning that can shed light on the development of these abilities. We will situate our discussion of these models in the *rational constructivism* theory of cognitive development (Gopnik & Wellman, 2012; Xu, 2019; Xu & Griffiths, 2011; see Ullman & Tenenbaum, 2021, for a related perspective on learning as hierarchical Bayesian program induction). Rational constructivism characterizes infants' initial representations as a set of proto-conceptual primitives and the mature state of human conceptual system as a set of domain-specific intuitive theories. Most importantly, a key mechanism that drives learning from the initial state to the mature state is the rational statistical inference that underlies Bayesian probabilistic models. This theoretical framework has been successful in bringing together

computational modeling and empirical work to shed light on a variety of topics in cognitive development (Xu & Griffiths, 2011; Xu & Kushnir, 2012). In this chapter, we will integrate recent models on intuitive psychology and social learning with past and new empirical studies on these topics within the framework of rational constructivism. We will highlight new insights provided by the models, and propose future directions that would advance our understanding of intuitive psychology and social learning.

2. Reasoning about others' mental states and actions

As children interact with and learn about the social world, one big challenge is to learn about and understand other individuals. People differ from objects in that each individual has unobservable and distinct mental states (e.g., goals, preferences, beliefs, motivations, emotions) that could only be inferred through behaviors. In developmental psychology, there is a long tradition of studying infants' and children's abilities to reason about others' mental states and actions. Past research has identified several key phenomena in the domain of intuitive psychology.

Infants interpret human's reaching and grasping motions as goal-directed actions (Woodward, 1998). In these studies, six-month-old infants observed an agent repeatedly reach for and grasp one of two toys. Then, the two toys switched locations. Infants expected the agent to reach for the object she previously selected, which is now at a new location, instead of reaching for the new object at the original location. Thus, infants encoded the selected object as the agent's goal, and expected agents to perform actions consistent with the goal.

Infants expect agents to use the most efficient means to achieve their goals (i.e., the rationality principle), and use this principle to guide their inferences about agent's actions, goals, and situational constraints (Gergely & Csibra, 2003). Nine- and 12-month-olds saw simple

animations where a small circle (the actor) reached a large circle (the goal) by jumping over a wall separating them. When the wall was removed, infants expected the agent to take a novel but efficient, straight-line path to reach the goal, instead of the same, jumping path (Csibra et al., 1999; Gergely et al., 1995). Twelve-month-old infants can infer agents' goals even when they have not seen the goals being achieved (Csibra et al., 2003). They observed a large circle approaching a moving small circle as if it was chasing the small circle. Then, the small circle went through a small gap between two walls. The large circle cannot go through the walls, so it went around the walls. When the small circle stopped, 12-month-olds expected the large circle to approach and make contact with the small circle, rather than pass by the small circle and continue moving. In other words, they attributed the goal of catching the small circle to the large circle even though they have not seen the large circle achieve this goal. Lastly, 12-month-olds can infer unobserved constraints from agents' actions and goals (Csibra et al., 2003). A screen occluded the view between a small circle (the actor) and a large circle (the goal). Infants observed the small circle taking a jumping path to reach the large circle. When the screen was removed, infants expected that there would be a wall (a constraint) between the small circle and the large circle.

Infants and young children can also infer others' preferences from sampling behaviors (Kushnir et al., 2010; Wellman et al., 2016). For instance, 10-month-olds observed an agent sample 5 target objects from a jar that consisted of either 20% of target objects or 80% of target object. Thus, when the jar consisted of 20% of target objects, the agent violated random sampling. But when the jar consisted of 80% of target objects, the agent did not violate random sampling. Infants only inferred that the agent preferred the target objects and expected the agent to choose the target object again in the 20% condition (Wellman et al., 2016).

Lastly, infants socially evaluate agents based on their helping or hindering behaviors toward third parties (Hamlin & Wynn, 2011; Hamlin et al., 2007). For instance, 6- and 10-month-olds observed the interactions among three agents (depicted as geometric shapes with eyes). The climber attempted to climb a hill. On alternating trials, infants observed a helper pushing the climber up, and a hinderer pushing the climber down. Infants showed a preference for the helper agent over the hinderer agent – they were more likely to reach for the helper agent than the hinderer agent when they were given a choice (Hamlin et al., 2007).

What is the computational basis of these fundamental inferences about agents' mental states and actions? Can these inferences be considered rational? In recent years, there has been a surge in using Bayesian probabilistic models to capture these inferences. These models use generative models (formalized intuitive theories) to specify how agents plan their actions based on their mental states, and then use Bayesian inference over the generative models to infer unobservable mental states from observable actions. In the following sections, we review three categories of Bayesian models on early intuitive psychology: inverse planning models, inverse decision-making models, and the Naïve Utility Calculus. For each category of models, we will first lay out the technical details of a representative model from that category, and then describe how these models have been applied to capture the key intuitive psychology phenomena observed in infants and children.

2.1. Inverse planning models

Inverse planning models capture the joint inferences of a variety of mental states – goals, beliefs, desires – from agents' actions and the constraints in the environments. This approach uses a generative model to capture agents' action planning process, assuming that observers

represent other agents as rational planners. The generative model is then inverted using Bayesian inference to infer the agent's mental states based on observed actions.

A representative inverse planning model is the Bayesian theory of mind (BToM) model developed by Baker and colleagues (2017). BToM aims to reverse-engineer the elementary form of mental state inferences that emerges in infancy, including goal encoding (Woodward, 1998), teleological representation of actions (Csibra et al., 2003), false belief understanding (Onishi & Baillargeon (2005), and social evaluation (Hamlin et al., 2007). BToM formalizes these mental state inferences as Bayesian inference over a generative model of a rational agent. The generative model uses partially observable Markov decision processes (POMDPs) – an artificial-intelligence approach to rational planning and state estimation.

We will use a simple scenario in Figure 1 to illustrate the model details. A student wants to buy lunch from a food truck. There are 3 types of food trucks, Korean (K), Lebanese (L), and Mexican (M), and 2 parking spaces (the yellow squares). The blue circle represents the student, and the black dots represent the traces of the student's movements. Unshaded area in the scenes is perceptually accessible to the student; shaded area is not perceptually accessible to the student. In this example, the student started at a place where she saw the Korean food truck was parked at the bottom-left parking space, but she could not see the top-right parking space. She moved toward a place where she saw that the Lebanese food truck was parked at the top-right parking space. Then, she returned to the bottom-left parking space and bought Korean food.

Figure 1. The food truck scenario in BToM (Baker et al., 2017).

In the model, the agent's beliefs are represented as a probability distribution over all possible world states. In the food truck scenario, world states include all the possible combinations of the food trucks that came to campus and the parking locations of each food truck. The agent's beliefs can be updated given the agent's new percepts. For instance, in frame 2 (the top-right frame) of Figure 1, the agent was able to see the parking space in the top-right corner, and her beliefs were updated. The agent's desires are represented by a utility function over situations, actions, and events. In the food truck scenario, the agent's desires are her utilities for eating at each food truck. BToM starts with prior beliefs about the agent's beliefs and desires, and jointly infers the posterior probability of the agent's beliefs (B), desires (D), percepts (P) and the situations (S) given the agent's actions (A). Formally, the posterior distribution is proportional to the product of the prior distribution and three likelihood terms:

$$\Pr(B, D, P, S|A) \propto \Pr(A|B_1, D) \times \Pr(B_1|P, B_0) \times \Pr(P|S) \times \Pr(B_0, D, S)$$

$\Pr(B_0, D, S)$ correspond to the prior distribution of the agent's belief, desire and the situation. $\Pr(P|S)$ corresponds to how the agent forms her percepts given the situation; $\Pr(B_1|P, B_0)$ corresponds to how the agent updates her belief given the initial belief and new percepts; $\Pr(A|B_1, D)$ corresponds to how the agent plans her actions given her beliefs and desires. For the last likelihood term, the model assumes that the agent will achieve her desires by choosing the most efficient actions (the principle of efficiency). Since an agent's behaviors might deviate from the rational model, the model adopts a graded expectation of utility maximization: the agent is most likely to choose the highest-utility action at each step in the planning process, but she sometimes chooses a non-optimal action.

Baker and colleagues (2017) also developed a few alternative models. Two of them are lesioned versions of the full BToM model. One lesioned model does not represent uncertain beliefs; it assumes that agents' beliefs are always the same as the true world state. The other lesioned model assumes that agents' actions do not involve costs; therefore, the model does not incorporate the principle of efficiency. A third alternative model is a motion-based heuristic model, which makes predictions based on the learned statistical associations between motion and environmental cues and people's judgments about agents' mental states. The performances of BToM and the alternative models were compared to adults' performances in 2 types of tasks. The first type of tasks involves inferring an agent's desires and beliefs from the agent's actions and the environments. After observing scenarios similar to the one in Figure 1, adults and the models inferred the agent's preferences for the 3 food trucks and her initial beliefs about which food truck was parked at the unobservable space. In the second type of tasks, adults and models inferred the states of the world from agent's actions and desires. They observed scenarios where an agent moved in a complex environment, searching for his favorite food truck (known to the participants). They inferred the locations of the food trucks in the environment based on the agent's actions. In both types of tasks, BToM performed closer to adults, and better than the lesioned BToM models and the motion-based heuristic model. This suggests that adults' abilities to make joint inferences of beliefs, desires, and percepts in these tasks cannot be achieved by merely learning the statistical associations between motion cues and mental states. Instead, a rational planning model that uses the principle of efficiency and represents uncertain beliefs is underlying adults' mental states inferences.

BToM has only been evaluated in tasks that are more complex than the intuitive psychology phenomena observed in infants. In contrast, Shu and colleagues (2021) evaluated

models on tasks similar to the ones that infants have been tested on. In particular, they focused on 4 key phenomena: inferring goals from agents' actions (Woodward, 1998), predicting agents' actions based on goals and constraints (Gergely et al., 1995), inferring constraints from agents' actions and goals (Csibra et al., 2013), and inferring agents' preferences from the levels of costs they incurred (Liu et al., 2017). Shu and colleagues (2021) tested the performance of two models on these tasks: an inverse planning model, the Bayesian Inverse Planning and Core Knowledge (BIPaCK), and a neural network model, the Theory of Mind Neural Network (ToMnet-G). BIPaCK is based on a generative model that integrates two components – a physics simulation that depend on core knowledge of objects and physics, and an agent planning process that depend on utility computation (maximizing rewards and minimizing costs). The physics simulation extracts different types of entities from the video – the agent, the goal objects, and obstacles – and recreates an approximated physical scene. The agent planning process is simpler than that in BToM. The model only represents one type of mental state – agent's goals. The model predicts a trajectory that allows the agent to reach the goal by maximizing the agent's reward and minimizing the agent's cost. Training videos were used to calibrate parameters, and test videos were used to evaluate model performance and generalization. For each test video, the model yields a surprise rating that is defined by the expected distance between the predicted agent trajectory and the one observed in the test video. Model performances were compared to adults' surprise ratings of the test videos. BIPaCK performed closer to adults than ToMnet-G did, and achieved better performance in generalization within and across scenarios. The findings suggest that an intuitive theory that represents agents' goals and use utility computation to plan actions underlies infants' reasoning in the 4 key intuitive psychology phenomena. In addition,

core knowledge of objects and physics is an important requirement for reasoning about agents' mental states and actions.

Hamlin and colleagues (2013) used an inverse-planning model to capture infants' social evaluation of agents who help or harm third parties. They used a combination of modeling and behavioral experiments to show that infants' social evaluation is based on inferences about agents' mental states. The study aimed to rule out alternative accounts including the low-level cue-based account (e.g., infants prefer those who push things uphill versus downhill) and mid-level accounts (e.g., infants only represent the protagonist's first-order goals, and positively evaluate any individuals who completes them; infants only represent second-order goals of helping and harming, but do not understand that helping and harming require having knowledge about the protagonist's goals). In the experiment, infants observed two Lifter puppets lifting one of two doors, allowing the Protagonist puppet to reach the object behind the doors. One Lifter always allowed the Protagonist to reach the object it had repeatedly grasped before, and the other Lifter always allowed the Protagonist to reach the object it had not grasped before. The experiment varied whether the Protagonist showed preference or not (the Protagonist repeatedly chose an object from 2 options or only 1 option), whether the Lifters were knowledgeable about the Protagonist's preference (the Lifters were present or absent when the Protagonist showed preference). The Full Mental model infers the Lifter's goals (prosocial, neutral, or antisocial) based on the Lifter's beliefs and actions, as well as the protagonist's goals. Consistent with the Full Mental model, infants only positively evaluated the Lifter who allowed the Protagonist to reach the preferred object, when the protagonist showed a preference, and when the Lifters are knowledgeable about the preference. The Full Mental model provided better qualitative fit to infants' behaviors compared to a feature-based model, models that evaluate agents based on

completion of the protagonist's first-order goal, or a model that evaluates the second-order goals but not second-order beliefs. Thus, 10-month-olds' social evaluation is based on inferences about the mental states of the protagonist as well as the helper and hinderer.

These models provided new insights to our understanding of the development of intuitive psychology. Previous experiments have demonstrated that even infants can make inferences about others' goals, beliefs, and desires. However, the underlying mechanisms of these inferences were less clear. Some researchers have argued that cue-based and heuristic-based learning can account for infants' and adults' reasoning about agents (Blythe et al., 1999; Gao et al., 2009; Perner & Ruffman, 2005; Scholl & Tremoulet, 2000). For instance, Perner and Ruffman (2005) argued that infants' understanding of false belief in Onishi & Baillargeon (2005) might be achieved by learning the association between agents, objects, and location in the scene. In contrast, the inverse planning models assume that infants, children, and adults have intuitive theories of how agents' mental states relate to their actions, and can make inferences about their mental states based on these intuitive theories. By comparing Bayesian inverse planning models with alternative models, Baker et al. (2017), Shu et al. (2021), and Hamlin et al. (2013) showed that adults' and infants' abilities to reason about agents cannot be achieved by merely learning the statistical associations between motion cues and mental states. Instead, it depends on intuitive theories of agent planning. Both adults and infants have complex intuitive theories that allow joint inferences over various types of mental states (Baker et al., 2017; Hamlin et al., 2013). In addition, Shu et al. (2021) revealed that core knowledge of objects and physics is required for reasoning about agents' mental states and actions. This is consistent with the developmental trajectory of infants' abilities to reason about mental states (emerge around 6 to 12 months of age; Gergely & Csibra, 2003; Wellman et al., 2016; Woodward, 1998) and their abilities to

reason about objects (emerge around 2 to 6 months of age; Anguiar & Baillargeon, 1999; Leslie & Keeble, 1987; Spelke et al., 1992). Based on the findings of Shu et al. (2021), it is possible that mental states inferences emerge slightly later in development because these inferences depend on core knowledge of objects.

An important future direction is to use inverse planning models to examine infants' abilities to understand false belief. Given the controversy in the interpretation of the findings in Onishi & Baillargeon (2005), it would be helpful to combine modeling and behavioral approaches to reveal whether the ability to represent false belief is indeed necessary for infants' performance in this task. For instance, this could be achieved by comparing infants' performance in this task to the performance of BToM, a lesioned version of BToM that always represents agents' beliefs as the true world states, and a motion-based heuristic model.

Future works can also extend the inverse planning model of social evaluation to capture more nuanced forms of social evaluation. For instance, Hamlin and colleagues (2011) found that 8-month-olds evaluated agents who showed prosocial or antisocial behaviors based on the dispositional status of the recipient. They preferred an agent who acted prosocially toward a recipient who previously engaged in prosocial behaviors, but they preferred an agent who acted antisocially toward a recipient who previously engaged in antisocial behaviors. However, 5-month-olds did not engage in such nuanced social evaluation – they always preferred a prosocial agent, regardless of the dispositional status of the recipient. The inverse planning model can be used to reveal the generative model (the intuitive theory) underlying these nuanced forms of social evaluation, and to capture the developmental changes between 5- and 8-month-olds.

2.2. Inverse decision-making models

The next category of models on intuitive psychology is inverse decision-making models. These models capture the inferences of a particular type of mental states – preferences. An inverse decision-making model incorporates a decision-making model as the generative model, which specifies how people make choices based on their preferences. Then, the model uses Bayesian inference to invert the decision-making model, and infers agents' preferences from their choice patterns.

Lucas and colleagues (2014) adopted an econometrics model, the Mixed Multinomial Logit model (MML), to capture the developmental data on children's inferences about others' preferences and choices. The MML specifies a choice model that maps people's preferences to choices. The choice model assumes that people combine the subjective utilities of different features of each option, and choose the option that maximizes their utility. The choice rule specifies that the probability that an agent will choose an option increases exponentially with that option's utility. Then, the MML uses Bayesian inference to invert the choice model, and infer others' preferences based on their choices.

The model captured the developmental data on children's preference inferences. In Kushnir et al. (2010), 20-month-olds and preschoolers observed an agent picking out 5 target objects from a box of different types of objects. Target objects contained 100%, 50%, or 18% of all objects in the box (Figure 2). Children were asked to infer the agent's preferred object from the target object, the alternative object in the box, or a novel object. The posterior probability of the agent's preference for each type of object is calculated based on the prior probability of its preference, and the likelihood that the agent will make the observed choices given that preference. In the 100% condition, since the agent could only choose the target objects, his preference would not influence his choices. Thus, the posterior probability of the agent's

preference distribution is uninformative. In the 50% condition, the posterior probability would shift toward having a preference for the target object, since the likelihood of observing these choices is low (0.5^5) if the agent had no preference and was sampling randomly. In the 18% condition, the posterior probability will shift toward even stronger preference for the target object, since the likelihood of observing these choices is even lower (only 0.18^5) under random sampling. Consistent with the model predictions, children chose randomly from the 3 objects in the 100% condition, and they were increasingly more likely to choose the target object in the 50% and the 18% conditions.

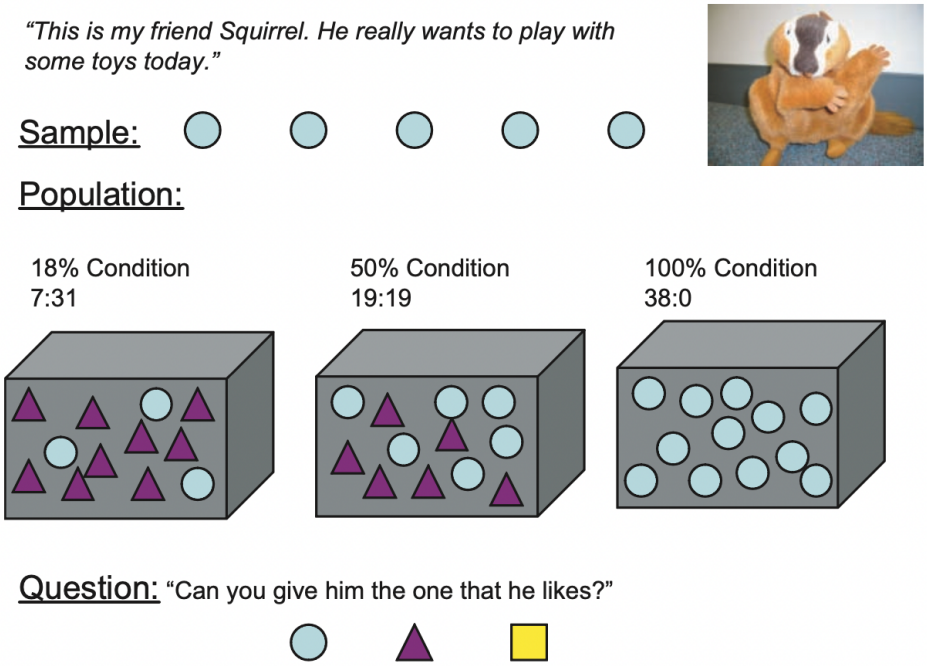

Figure 2. A schematic depiction of the experimental procedure in Kushnir et al. (2010).

The model also captured the developmental difference in preference understanding observed in Repacholi & Gopnik (1997). In the study, 14- and 18-month-olds observed an agent expressing preference either matched (liking goldfish crackers and disliking broccoli) or

unmatched (liking broccoli and disliking goldfish crackers) to their own preferences. When asked to offer the agent some food, younger children offered crackers (their own preference) regardless of condition; but older children offered what the agent preferred: crackers in the matched condition, and broccoli in the unmatched condition. Lucas et al. (2014) captured this developmental difference by proposing a shift in children's model of preference understanding. A simpler model assumes that all people have the same preferences, and the more complex model assumes that each individual has a distinct set of preferences. Lucas and colleagues (2014) ran simulations that captured how accumulation of data would lead to a shift from the simpler to the more complex models. The simulations assumed that children gradually observe more and more choices of her own, her parent and her sibling. At the beginning, their own preferences are broadly similar to their parents and siblings. The simpler and the more complex model capture the data equally well, therefore the simpler model is preferred based on Bayesian Occam's razor. As children's observations grow, the simpler model fails to account for the differences in individuals' preferences, and the more complex model is now more probable. The simulations captured the experimental results in Repacholi & Gopnik (1997): younger children should be more likely to offer crackers to an agent with preferences unmatched to their own preferences, and older children should be more likely to offer broccoli. Consistent with the simulation results, a training study showed that after observing two experimenters expressing different preferences repeatedly, even 14-month-olds can represent others' preferences that are different from theirs (Doan, Denison, Lucas, & Gopnik, 2015). In addition, the model predicts that children who are in the process of shifting from the simpler to the more complex model should be sensitive to the strength of evidence indicating that an agent has a preference different from their own. This prediction is corroborated by evidence from 16-month-olds (Ma & Xu, 2011): when an agent

chose 6 boring toys from a jar containing 13% of boring toys and 87% of interesting toys (strong evidence that the agent preferred the boring toy), 16-month-olds were more likely to offer a boring toy to the agent, compared to 16-month-olds who saw an agent chose 6 boring toys from a jar containing 100% boring toys.

Inverse decision-making models have charted the developmental trajectory of infants' intuitive theories of preference. While 14-month-olds' performance in Repacholi & Gopnik (1997) appears irrational, the modeling work showed that it is because they have a simpler model of preference understanding, which is the most reasonable model given the data that they have observed. The modeling work further showed that infants can construct a new, more complex model of preference understanding through Bayesian inductive learning. The transition from the simpler to the more complex model is predicted by both the amount and the strength of additional data that infants observe.

2.3. Naïve Utility Calculus

The last category of models, the Naïve Utility Calculus (NUC; Jara-Ettinger et al., 2016; 2020), builds on both the inverse decision-making models and the inverse planning models. The NUC extended the inverse decision-making models by allowing joint inferences over costs and rewards (instead of just rewards). It extended the inverse planning models by adding a few additional levels of analysis in agents' action planning process, distinguishing between agents' desires, goals, intentions, and actions.

The general principle behind the NUC is that we reason about agents' actions based on the assumption that agents are intuitive utility maximizers. We assume that agents would maximize their rewards and minimize their costs when they choose their goals and actions. The model consists of a generative model that produces utility-maximizing behaviors given an

agent's costs and rewards, and a mechanism that uses the generative model to infer an agent's costs and rewards from observed action sequences.

We will use a simple scenario shown in Figure 3 to illustrate the model details. In this scenario, an astronaut is on an alien planet. She starts at the middle-left location on the map, and her goal is to get to the space station at the middle-right location. She can collect two types of care packages on the way, the orange one at the top and the white one in the middle. There are two different types of terrains, the grey terrain and the blue terrain.

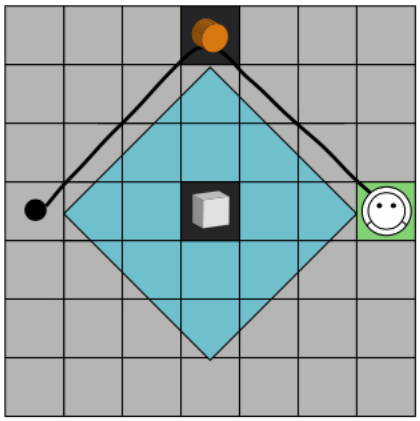

Figure 3. The astronaut scenario in the NUC (Jara-Ettinger et al., 2020).

The generative model is a hierarchical representation of how an agent's actions are produced. An agent's desires are rewards associated with having different objects or helping different agents. In the astronaut scenario, the astronaut has different rewards for the space station, the yellow package and the white package. These rewards, combined with beliefs about the location of the objects or agents determine the space of possible goals. Goals are defined as states of the world that the agent finds rewarding. In the astronaut scenario, the goals include getting to the space station, collecting the white package, and collecting the orange package. Next, the space of goals determines the space of intentions, which are ordered sequences of

goals. The model only considers intentions that satisfy certain context-specific constraints (e.g., always have a specific goal as the final goal). In the astronaut scenario, the intentions must have the space station as the final goal. Thus, the space of intentions includes: 1) go to the space station; 2) collect the orange package, and go to the space station; 3) collect the white package, and go to the space station; 4) collect the orange package, collect the white package, and go to the space station; 5) collect the white package, collect the orange package, and go to the space station.

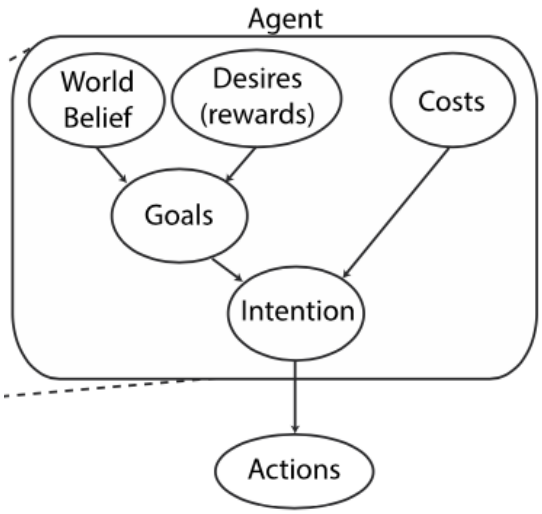

Figure 4. The generative model of NUC (Jara-Ettinger et al., 2020).

For each goal, the model computes an action policy that probabilistically maximizes utility; the action policy determines what actions the agent should take to achieve the goal state most efficiently. The action policy is computed through individual special-purpose Markov Decision Processes (MDPs). The individual MDP uses the reward associated with the goal, and plans movement in space. In the astronaut scenario, there are 49 positions in the map. At any time, the agent can move in one of eight directions (4 cardinal directions and 4 diagonal directions). The cost of each action is determined by the type of terrain and length of the

movement. A reward is obtained if the goal state is achieved (e.g., if the astronaut reached the orange package). The MDP computes an action policy that maps the goal states to actions with the highest utility, so that when the action policy is executed, the agent reaches the goal state as efficiently as possible. The utility of each intention is the sum of the goal's rewards minus the costs that the agent would incur to complete these goals in the specific order. The model selects an intention with the highest utility. Then, the selected intention is transformed into actions by executing each goal's action policy.

Another component of the model is a mechanism that uses Bayesian inference over the generative model to infer costs and rewards given a sequence of actions. The posterior probability of the cost and the reward functions is proportional to the product of the prior probability of the cost and the reward functions and the likelihood of observing the action sequence given the cost and the reward functions:

$$p(C, R|A) \propto \; l(A|C, R) \; \times \; p(C, R)$$

Since the terrains and the objects are novel, the model uses a uniform prior distribution for the cost function of each terrain, and a uniform prior distribution for the reward function of each object. The likelihood term is the probability that the generative model would produce the observed action sequence given each intention, summed over all possible intentions considered by the model:

$$l(A|C, R) \; = \sum_{I \in Intentions} p(A|I)p(I|C, R)$$

The likelihood term can be approximate with one or few high-probability intentions, rather than a sum over all possible intentions. The model approximates this Bayesian inference through Monte Carlo likelihood weighting. Instead of computing the posterior probability of all cost and rewards functions, the model samples *n* cost and reward functions from the prior distribution,

and computes the likelihood of observing the action sequence under each sample of cost and reward functions.

The predictions of the NUC have been tested in adults with a variety of tasks. In one experiment (Jara-Ettinger et al., 2020), adults observed astronauts traveling to space stations in different environments (similar to Figure 3). Then, they were asked to judge the agents' abilities to travel through each type of terrain, and their desires to collect each care package. Participants' judgments are predicted by the NUC with quantitative accuracy, and shows more sensitivity than the alternative models, such as a heuristic model that infers cost from the amount of time an agent spends in each terrain and infers reward by the objects she collects. Other experiments (Jara-Ettinger et al., 2020) showed that the NUC also captured adults' performances when participants were asked to predict agents' future behaviors, to infer agents' knowledgeability about costs and rewards, and to reason about the agents' social goals (helping other agents).

Some of these predictions have also been tested in children and infants. One set of studies (Jara-Ettinger, Gweon, et al., 2015) found that 5- to 6-year-olds expect agents to maximize their utilities. They were able to use agents' choices under different situations to infer subjective rewards, and use agents' choices and subjective rewards to infer subjective costs. For instance, 5- to 6-year-olds learned that puppet A likes cookies over crackers, puppet B likes both equally, and that only one puppet can climb the tall box. When both puppets picked the cracker on the short box over the cookie on the tall box, children inferred that puppet A cannot climb the tall box.

Another study (Jara-Ettinger, Tenenbaum, et al., 2015) showed that even toddlers understand that different agents need to incur different costs for the same action. Seventeen- to 28-month-old toddlers observed a competent puppet made a toy play music on the first attempt and an incompetent puppet made the toy play music after many attempts. They preferred to play

with the competent puppet, but when both puppets refused to help others activate the toy, they inferred that the incompetent puppet is nicer. Thus, consistent with the expectation that agents maximize utility, toddlers understand that failure to engage in a low-cost prosocial action implies that the agent has lower reward associated with prosocial behaviors.

Infants understand that agents prefer goals that they achieved through costlier actions. In Liu et al. (2017), infants observed agents who were willing to incur different levels of cost to reach other agents. In one scenario (adapted from Gergely et al., 1995), the main character jumped over barriers of different heights to reach other agents. The main character was willing to jump over a low barrier, but not a medium barrier, to reach the blue agent; it was willing to jump over a medium barrier, but not a high barrier, to reach the yellow agent. Infants inferred that the main character preferred the yellow agent over the blue agent. Similar results were obtained when cost was defined as the width of the gaps that the agent had to jump over, or the incline angle of ramps.

Children also understand that agents maximize their *expected* utilities, instead of *actual* utilities, since agents might be ignorant about their own costs and rewards (Jara-Ettinger et al., 2017). Four- to 5-year-olds understand that knowledgeable agents are more likely to choose high utility options and their choices are more likely to be stable, compared to agents who are ignorant about their own rewards or costs. For instance, children were presented with two puppets and two novel fruits. They were told that one puppet has tasted both fruits before and the other has not. Both puppets chose the same fruit; one said “Yum” and the other said “Yuck”. Children inferred that the puppet who said “Yuck” has not tasted the fruits before.

The NUC model has been extended to capture the reasoning behind teaching decisions (Bridgers et al., 2019). Children learned about two toys that varied in discovery rewards and

discovery costs, and the experimenter (the teacher) could only teach one of the toys to a naïve learner. Children helped the experimenter decide which toy she should teach the learner, and which toy she should let the learner discover on her own. The model makes this decision by maximizing the learner's expected utility of learning from instruction and exploration. Thus, the learner's utility is the activation reward of the instructed toy and the discovery reward of the explored toy, subtracted by the activation cost of instructed toy and the discovery cost of the explored toy. Five- to 7-year-olds' behaviors were consistent with the model's predictions – they were more likely to teach a toy with higher reward than a toy with lower reward when the costs were equal; they were more likely to teach a toy with higher cost than a toy with lower cost when the rewards were equal; when one toy had higher reward and lower cost and the other toy had lower reward and higher cost, children were increasingly more likely to teach the higher cost toy as the difference in costs increased.

Lastly, Meng and Xu (2020; 2021) have shown that the rational inference behind NUC alone can "reproduce" observed disparity among unbiased people. For example, when people observe police officers stopping and searching Black people at higher rates compared to the other groups, they may infer that Black people are more likely to commit crimes because they assume that the police officers would maximize their utilities by searching groups with higher crime rates. In one study (Meng & Xu, 2021), adults observed a knowledgeable agent, a border patrol officer on a planet, using a costly method (a scanner) to check whether aliens from other planets have stolen gems. The officer knew the "theft rate" of aliens from each planet and checked aliens from a series of planets at different rates. In one condition, participants also observed the sample hit rates – the number of aliens who actually stole gems among the ones the officer checked. In the critical trials, the check rate and the hit rate conflicted (e.g., the hit rate was low when the

check rate was high). Then, participants were asked to infer the theft rate of all aliens on each planet. Adults' inferences were consistent with the predictions of the NUC models. When they were only given check rates, they inferred higher theft rates as the officer checked the aliens more often. When they were given both check rates and hit rates, participants rationally integrated the two types of information in their inferences (e.g., they inferred moderate theft rate for groups that were checked often but had low sample hit rate). Thus, observing a knowledgeable, utility-maximizing agent checking different groups at different rates leads to reproduction of disparities. Providing sample hit rates might reduce this negative consequence.

The NUC provides a unified account for reasoning about agents' actions across a variety of contexts. We assume that others choose actions that maximize their utilities – they tradeoff between costs and rewards in a precise manner. Furthermore, this assumption is consistent across development. Even infants use this assumption to reason about agents' actions in simple scenarios. The NUC also motivated many additional behavioral experiments that tested different aspects of this unified theory.

A future direction for both the NUC models and the inverse planning models is to incorporate mental costs of planning in the models. One reason that agents do not always act rationally is because planning complex sequences of actions takes time and mental resources. When people do not have enough time and resources to engage in planning, they execute non-optimal plans and inefficient actions. People might also tradeoff between the mental cost of planning and the reward of finding an optimal plan. That is, they will only engage in planning if the optimal plan leads to increase in rewards that offsets the mental costs of planning. There are two ways to extend the NUC and inverse planning models in this direction. First, the models can move towards the algorithmic level of analysis, and consider the plausible algorithms that are

underlying human reasoning given the limited amount of cognitive resources. For instance, Lieder and Griffiths (2020) proposed a new modeling paradigm called *resource-rational analysis*. *Recourse-rational analysis* assumes that people make rational use of their limited cognitive resources, maximizing the utility of the decision while minimizing the costs of the decision-making process. Second, a variable representing the mental costs of planning can be directly added to the models. The models can tradeoff between the mental costs and the rewards of finding an optimal plan to determine the amount of planning the agent should engage in.

In sum, the three categories of models reviewed above have successfully captured infants' and children's abilities to make mental state inferences across several key phenomena in the domain of intuitive psychology. These models showed that infants, children, and adults reason about agents based on intuitive theories of psychology that specifies how agents plan their actions. Utility calculation – the assumption that agents choose actions that maximize their rewards and minimizes their costs – is a critical component in inverse planning and NUC models, suggesting that it might be a unifying principle underlying our intuitive theory of psychology.

3. Pedagogical reasoning and epistemic trust

The social world also provides children with the opportunity to learn from others. Teaching allows quick transmission of important knowledge. However, choosing the best examples to teach and learning the correct hypotheses based on the examples are nontrivial inferential problems. What kinds of mechanisms underlie our abilities to learn from other people and teach other people? This question is addressed by the pedagogical model. In addition, not everyone is the best teacher to learn from. How do we decide whom we should trust and learn from? This question is addressed by the epistemic trust model.

3.1. Pedagogical model

When children learn about the world, they need to infer the correct hypotheses based on the enormous amount of data that they observe in the world. Pedagogical situations have an important impact on this process, because in these situations, data are not generated randomly; instead, the teacher is choosing data to transmit to the learner for the purpose of teaching a particular hypothesis (Csibra & Gergely, 2009). How can a teacher optimally teach a hypothesis to a learner? Imagine you want to teach a child the word "dog". You can use the word "dog" to refer to three golden retrievers, or use the word "dog" to refer to a golden retriever, a dalmatian, and a border collie. Intuitively, the second set of examples is better than the first one. The child is more likely to infer the correct hypothesis, that "dog" refers to the basic level category *dog*, given the second set of examples. Why is that the case? When you choose three golden retrievers as examples, the child would be more likely to believe in the hypothesis that "dog" refers to the subordinate level category *golden retriever*, rather than the hypothesis that it refers to the basic level category *dog*. If the latter hypothesis were true, you would have chosen examples of other kinds of dogs as well. However, the child will only engage in this kind of reasoning if she assumes that you are intentionally choosing the examples to help her learn. In contrast, if the child overheard you say "dog" three times while pointing to three golden retrievers, the child would be less certain that "dog" refers to *golden retriever*. Because the examples are generated randomly, both hypotheses ("dog" refers to *golden retriever*, and "dog" refers to *dog*) are consistent with the data she observed. Therefore, in pedagogical communications, the learner needs to reason about the teacher's intentions, and distinguish between when the examples are provided for the purpose of helping her learn and when the examples are generated randomly.

The teacher also needs to reason about what the learner would infer given different examples, and choose the examples that are most helpful for the learner.

The natural pedagogy perspective (Csibra & Gergely, 2009) proposes that such pedagogical communications are achieved by a specifically adapted human communication system called "natural pedagogy". "Natural pedagogy" takes place when communications are accompanied by ostensive cues such as eye contact, pointing, and child-directed speech. Most importantly, this perspective proposes that the knowledge being transmitted in such contexts are kind-relevant and generalizable. Indeed, evidence suggests that both infants and preschoolers expect to learn generalizable knowledge when communications are accompanied by ostensive cues (e.g., Butler & Markman, 2012; Egyed et al., 2013).

In contrast, the pedagogical model proposed by Shafto, Goodman and Griffiths (2014) does not make any assumptions about the kinds of knowledge being transmitted in pedagogical communications. Instead, it focuses on the learner's and the teacher's ability to reason about the mental states of each other. In particular, the learner reasons about the process by which the teacher chooses the data as she updates her beliefs; the teacher reasons about the learner's belief updating process and chooses data that will be most helpful to the learner.

How should the teacher sample data that are helpful to the learner? The literature on concept learning has distinguished between weak sampling and strong sampling. In weak sampling, data are randomly selected from all possible examples, and are labeled as to whether they are true of the target hypothesis (e.g., Hsu & Griffiths, 2009). In strong sampling, data are randomly selected from the set of examples that are true of the hypothesis (e.g., Xu & Tenenbaum, 2007a, 2007b). However, neither of these two types of sampling could capture the sampling process that underlies teaching. A key assumption of the pedagogical model is that the

teacher is engaging in *pedagogical sampling* – choosing data that will maximize the learner's belief in the correct hypothesis, that is, the posterior probability of the correct hypothesis. The model further assumes that the learner knows that data are sampled by a helpful teacher, and rationally updates her belief.

Based on these two assumptions, the pedagogical model formalizes a system of equations, specifying the distribution from which the teacher generates data (the sampling distribution), and the process by which the learner updates her belief given data. Critically, these equations depend on each other – the teacher's sampling distribution depends on how the learner will update her belief given data; the learner's belief updating process depends on the teacher's sampling distribution. The model solves this system of equations using a mathematical method called fixed-point iteration. This method is analogous to a process of recursive mental state reasoning (although the actual psychological mechanisms used to solve this problem do not necessarily involve explicit recursion). When the learner updates her belief given data, she needs an estimation of the likelihood that the teacher generated the data given the true hypothesis; to do this, she has to make an assumption about the teacher's sampling distribution (e.g., weak sampling, strong sampling, or pedagogical sampling); if pedagogical sampling is assumed, this process depends on the teacher's assumption about how the learner will update her belief; if the teacher assumes the learner will rationally update her belief, this assumption, in turn, depends on the learner's assumption about the teacher's sampling distribution, and so on. This recursive reasoning will eventually converge, at which point we have the solution to the system of equations. The model achieves this solution by first specifying an initial distribution from which the teacher generates data – the initial sampling distribution assumes unbiased random sampling. Depending on whether negative evidence is possible, it will be either weak sampling (random

sampling from all possible examples) or strong sampling (random sampling from all examples true of the hypothesis). Then, the fixed-point iteration process described above will transform this initial distribution into a solution satisfying all model assumptions.

For instance, in one task in Shafto et al. (2014), a teacher teaches a rule-based concept (a rectangle on a board) to a learner. In one condition, the teacher can provide 2 positive examples (i.e., 2 points inside the rectangle) to help a learner infer the correct rectangle. The model assumes that the teacher's initial sampling distribution is strong sampling – the teacher is equally likely to choose any points inside the rectangle. Then, the model makes predictions about the hypotheses that the learner will infer given the assumption that the teacher samples from this distribution. The predictions would show that some examples are more likely than others to allow the learner to infer the correct hypothesis (e.g., points near the corners of the rectangle will be more likely to lead to the correct hypothesis compared to other examples). Then, the model updates the teacher's sampling distribution by increasing the likelihood of examples that are more likely to lead to the correct hypothesis (e.g., given the new distribution, the teacher will be more likely to choose points closer to the corners). Then, the model repeats the processes of predicting the hypotheses that the learner will infer, and updating the teacher's sampling distribution based on the predictions. Eventually, the teacher's sampling distribution will converge to examples that are most likely to lead to the correct hypothesis (e.g., pairs of points at two opposite corners). Indeed, adults who played the role of the teacher in the experiments chose these examples to teach the learner. Other participants played the role of the learner. When learners were told that the examples were generated by a helpful teacher, the rectangles that they inferred were more likely to have the positive examples at the corners. That is, they understood the process by which the teacher chose the examples, and are were more likely to infer the

correct rectangle. When learners thought the examples were generated by the computer, the rectangles they inferred did not show this specific pattern, and they were less likely to infer the correct rectangle. Shafto et al. (2014) examined the model predictions in two other tasks on prototype concepts and causally structured concepts, and showed that adults' teaching and learning in these tasks were well captured by the pedagogical model. Furthermore, this model can be applied to teaching and learning any concept as long as the hypothesis space can be specified.

The general prediction of the pedagogical model is that a teacher should provide examples that would maximize the learner's belief in the correct hypothesis. A few specific predictions follow from this general prediction: first, the examples that a teacher provided should be exhaustive, and the learner can infer that any hypotheses for which the teacher did not provide evidence are not true; second, a teacher would only provide necessary examples, and the true hypothesis should be consistent with all examples, rather than a portion of the examples; third, when there are infinite numbers of possible examples for a hypothesis, such as points in rectangle in the task in Shafto et al. (2014), the most diverse set of examples should be preferred (e.g., points on the opposite corners of the rectangle). Studies with young children have supported these model predictions. First, Bonawitz and colleagues (2011) found that 4- to 5-year-olds expected a teacher's examples to be exhaustive. When a knowledgeable teacher demonstrated one function on a toy, they explored the toy less later, compared to when the function was demonstrated by an ignorant teacher. Children inferred that a knowledgeable teacher would demonstrate all functions on the toy; since she only demonstrated one, children did not expect to find additional functions through exploration. Second, Buchsbaum and colleagues (2011) found that 3- to 5-year-olds expected a teacher to only provide necessary

examples. In this study, children learned about a causally structured concept – the sequence of actions that leads to an effect. The teacher was either knowledgeable about the toy and taught the child how it worked by demonstrating several three-action sequences (pedagogical condition), or the teacher was ignorant about the toy and demonstrated the same three-actions sequences while she tried to figure out how it works (non-pedagogical condition). In some trials, the statistical evidence indicated that only two of the actions in the sequences were the overlapping cause of the effect. In accordance with the statistical evidence, children in the non-pedagogical condition produced the two-action sequence. However, children in the pedagogical condition were more likely to overimitate and produce the three-action sequence. That is, children inferred that all three actions were necessary, because they assumed that the teacher would not demonstrate superfluous actions if she was helping them learn. Last, Rhodes and colleagues (2010) found that 6-year-olds prefer to teach with a diverse set of examples. Participants taught another child a novel property that was true of a subset of animals (e.g., dogs). They found that children were more likely to choose a diverse set of examples (e.g., a golden retriever, a dalmatian, and a collie) than a non-diverse set of examples (e.g., three dalmatians).

The development of the pedagogical model has also led to new behavioral investigations. For instance, given the important role that mental state reasoning plays in the pedagogical model, Bass and colleagues (2019) examined the link between preschoolers' ability to select evidence to correct others' false beliefs and their theory of mind abilities. Three- to 4-year-olds first learned the kind of blocks that activate a toy (e.g., red blocks activate the toy). Then, they provided evidence in the presence of a confederate with a false belief (e.g., the confederate believed square blocks activate the toy). Children with better evidence selection ability were more likely to select evidence contradicting the confederate's false belief (e.g., showing that a red, circle

block activated the toy). They found that there is a correlation between children's evidence selection ability and ToM ability, above and beyond effects of age and other cognitive abilities. Furthermore, a 6-week training of pedagogical evidence selection improved children's ToM ability. In another study, Gweon and colleagues (2018) investigated the concept of informativeness in teaching in 5- to 6-year-olds. They found that children preferred teachers who provided the sufficient number of demonstrations based on the learners' knowledge levels, and modulated their own teaching in similar fashions. The pedagogical model assumes that the teacher should maximize the probability that the learner believes in the correct hypothesis, that is, they should always generate more data until no additional data could benefit the learner. However, even children understand that teachers should not be overinformative. Gweon et al. (2018) extended the pedagogical sampling assumption by incorporating the cost of information transmission – a rational teacher should maximize the utility of information, instead of the reward of information.

Next, we will identify two sets of behavioral studies in cognitive development that could be integrated with the pedagogical model. The first set of studies supported the natural pedagogy assumption, that knowledge transmitted through pedagogical communication is generalizable (e.g., Butler & Markman, 2012; Egyed et al., 2013). As Shafto et al. (2014) also discussed, their model could be extended to incorporate the natural pedagogy assumption. For instance, the model could specify that in pedagogical situations, the learner has stronger prior beliefs for hypotheses about generalizable concepts than hypotheses about nongeneralizable concepts. The teacher would assume that the learner has these prior beliefs when she engages in pedagogical sampling.

The second set of studies examined children's normative inferences about intentional actions. Schmidt and colleagues (2016) found that 3-year-olds are "promiscuous normativitists", that is, they inferred social norms from single, spontaneous human actions, in the absence of any linguistic or behavioral cues indicating that the actions might be generic or normative. Furthermore, Butler and colleagues (2015) found that 3-year-olds' normative inferences were even stronger when the actions were demonstrated pedagogically, compared to when they were demonstrated intentionally. These findings can be analyzed in the pedagogical model framework: Children assign stronger prior probabilities to normative hypotheses compared to non-normative hypotheses about any intentional actions – they believe that intentionally demonstrated actions are more likely to be normative. Pedagogical demonstration of an action further increases the probability that the normative hypothesis is correct. Since the learner assumes that the teacher knows their prior beliefs (e.g., that an intentionally demonstrated action is more likely to be normative), the fact that the teacher chose to demonstrate this particular action increases the learner's posterior belief that this action should be normative.

Lastly, we will point out one limitation of the pedagogical model. The model assumes the teacher always has accurate representations about the learner's hypothesis space, which might not be true in naturalistic teaching situations. For instance, Aboody and colleagues (2018) designed a task in which a participant taught another participant the activation rule for a toy. Given the evidence that the "teachers" provided, only about half of the "learners" inferred the correct rule. Aboody and colleagues entered the evidence provided by the "teacher" participants into a computational model similar to the pedagogical model and examined the model's inference of the activation rule. The accuracy of the model inference was much higher (75%) when the model learned under a simple hypothesis space (e.g., only allowing rules containing

single blocks or two blocks), than under a complex hypothesis space (25%). Furthermore, constraining participants' hypothesis space led to more accurate learning given the same evidence provided by the "teachers".

In sum, the pedagogical model describes how teachers choose data and how learners update their beliefs in pedagogical situations. Past studies on adults' and children's learning and teaching are consistent with the model predictions. The model has led to interesting behavioral investigations about the relationship between ToM and pedagogical evidence selection, as well as children's understanding of informativeness in teaching. Future work could integrate the natural pedagogy perspective with the pedagogical model, and use the pedagogical model to capture children's normative inferences about intentional actions. The model can also be extended to capture teaching and learning when the teacher does not have accurate representation of the learner's hypothesis space.

3.2. Epistemic trust model

While the pedagogical model only focuses on learning in situations where the teacher is always knowledgeable and helpful, the epistemic trust model (Shafto et al., 2012) captures the phenomenon of learning from others in a broad range of situations. When children do not know who is knowledgeable and helpful, how do they decide whom to trust and simultaneously learn about the world? The epistemic trust model developed by Shafto and colleagues (2012) focuses on a particular task that has been studied extensively in the epistemic trust literature – learning a new word label for a novel object from informants.

The epistemic trust model includes a generative model about how an informant chooses a label to provide; a learner observes the provided label, and simultaneously infers the true state of the world (i.e., the actual label for the object) and the informant's knowledgeability and

helpfulness. The generative model specifies that the informant chooses a label based on her belief and helpfulness. Her belief depends on her knowledgeability and the true state of the world. The model assumes that the knowledgeability variable is binary: A knowledgeable informant always believes in the actual label, and a non-knowledgeable informant is equally likely to believe in any label in a set of possible labels. The helpfulness variable is also assumed to be binary: A helpful informant selects the label that maximizes the probability that the learner forms the same belief as the informant, whereas an unhelpful informant minimizes this probability.

Shafto and colleagues (2012) compared the fits of two models with children's performance in past studies. The Knowledge & Intent model includes parameters (free parameters that were fitted to data) that reflect children's prior beliefs about informants' knowledgeability and helpfulness on average, and the variability of knowledgeability and helpfulness across informants. The Knowledge-only model also included the knowledgeability parameters (fitted to data), and the helpfulness parameters were fixed to reflect the assumption that informants are always helpful.

Past studies have demonstrated 4-year-olds' competence in various epistemic trust tasks. However, 3-year-olds' performance in some of these tasks differed from 4-year-olds. Different studies suggest slightly different interpretations for this developmental change. In Pasquini et al. (2007), children were asked to endorse novel labels provided by 2 informants with different previous accuracies on labeling familiar objects (100% vs. 0%; 100% vs. 25%; 75% vs. 0%; 75% vs. 25%). While 3-year-olds trusted the more accurate informants only when one of the informants were always accurate, 4-year-olds trusted the more accurate informants in all conditions. These results suggest that 3-year-olds might be less sensitive to the relative frequency

of errors compared to 4-year-olds. In Corriveau & Harris (2009), children were presented with familiar and unfamiliar informants, and the informants provided correct or incorrect labels for familiar objects in the experiments. Although both 3- and 4-year-olds initially trusted familiar over unfamiliar informants, when the familiar informant provided incorrect labels for objects and the unfamiliar informant provided correct labels, only 4-year-olds trusted the unfamiliar informant more. These results suggest that younger children might have trouble integrating new accuracy information into judgments about familiar informants. In Mascaro & Sperber (2009), 4-year-olds decided not to trust an informant whom they were explicitly told to be a big liar, whereas 3-year-olds continue to trust that informant. This result suggests that younger children might ignore information about informants' intent. Lastly, in Corriveau et al. (2009), both 3- and 4-year-olds were more likely to trust an informant who previously referred to a novel object with a label that agreed with the majority of the group, compared to an informant who previously disagreed with the majority of the group.

Across all 4 studies, 4-year-olds' behaviors were best captured by the Knowledge & Intent model. The best-fitting parameters revealed that 4-year-olds believe that informants are knowledgeable and helpful on average, but different informants vary in their knowledgeability and helpfulness. However, across all 4 studies, 3-year-olds' behaviors were captured equally well by the Knowledge & Intent model and the Knowledge-only model. Moreover, the best-fitting parameters of the Knowledge & Intent model for 3-year-olds were consistent with an assumption that people are uniformly helpful. In other words, 3-year-olds believe that all informants are helpful, but some informants are more knowledgeable than others. Thus, the modeling results provided a unifying alternative explanation for the developmental changes observed in 3- and 4-year-olds' performance in various tasks – that 3- and 4-year-olds differ in

their assumptions about the helpfulness of informants. Future work could examine whether this developmental shift can be captured by Bayesian inductive learning, that is, whether the shift to the new assumption reflects an integration of children's prior beliefs (i.e., all informants are helpful) and the new data they observe (e.g., some informants are not helpful).

So far, the epistemic trust model has only been used to capture children's behaviors in a particular task: when children were given information about the reliability of informants, they selectively learn novel labels from reliable informants. However, children's epistemic reasoning abilities have been demonstrated in various tasks. Next, we identify three lines of behavioral studies that could be integrated with the epistemic trust model.

First, Schütte and colleagues (2020) have demonstrated 5-year-olds' ability to use informants' reliability to make epistemic trust judgments retrospectively. In their study, children first observed conflicting testimonies provided by two unfamiliar informants – they used the same novel label to refer to different novel objects. Then, they received new information about the informants' reliability: one informant consistently labeled familiar objects accurately, and the other consistently labeled them inaccurately. Children retrospectively inferred that the testimony (i.e., the referent of the novel label) provided by the reliable informant was more likely to be correct. The epistemic trust model could be used to capture these results. At first, children had no information about the knowledgeability and helpfulness of the two informants, and they might infer that the two conflicting testimonies were equally likely to be true. When they received new information about the informant's knowledgeability, their original inferences were updated – now they inferred that the knowledgeable informant was more likely to have provided the correct testimony.

Another study by Liberman and Shaw (2020) found that children understand that people can be biased in their testimony about friends or enemies. When 3- to 13-year-olds heard a negative testimony about a person (e.g., "she is bad at soccer"), they judged that the target person was worse at soccer if the testimony was provided by a friend, compared to an enemy. When they heard a positive testimony about a person, they made more positive judgments about the target person's ability if the testimony was provided by an enemy, as opposed to a friend. These results could be captured by the epistemic trust model with a slight modification. The helpfulness variable in the original model could be changed to a variable about general intention. The nature of the social relationship between the informant and the target of the testimony affects the informant's intention (e.g., the informant could be positively or negatively biased). Observers can infer the true state of the world (e.g., the target's true ability) based on the testimony and the social relationship between the informant and the target (i.e., they can correct for the informant's bias).

Lastly, studies by Butler and colleagues (2018, 2020) showed that children reason about the process by which claims are made, above and beyond reasoning about informants' knowledgeability and helpfulness. Children as young as 3 years of age judged verified claims (e.g., when a person looked inside a container and made a claim about what is inside the container) to be more acceptable than claims that have not been verified (e.g., when the person chose not to look inside the container and made the same claim) (Butler et al., 2018). In addition, 6- and 7-year-olds, but not younger children, treated verification as more important than an informant's past history of accuracy when they judged whether a claim should be trusted (Butler et al., 2020). They were more likely to trust a verified claim provided by an informant who previously provided inaccurate word labels, compared to an unverified claim provided by an

informant who previously provided accurate word labels. These results could help extend the existing epistemic trust model. For instance, an additional variable, the method that informants use to form their beliefs, could be incorporated into the model. This variable would determine whether or not the informant's belief would match the true state of the world. For instance, if a verification process was used, the informant's belief is more likely to match the true world state, and therefore the informant's testimony is more likely to be accurate. Modeling these results could also shed light on the developmental change in children's ability to integrate verification and past accuracy information in their epistemic trust judgments.

In sum, the epistemic trust model describes how children reason about the testimony provided by informants to learn about the world and to infer the reliability of informants. Past studies on children's ability to learn novel labels from informants are well captured by this model. Recent studies have demonstrated children's abilities to retrospectively update their beliefs about testimonies, to reason about biased testimony, and to reason about the process by which testimonies are made. These new findings are consistent with the predictions of the epistemic trust model, and suggest ways that the model could be extended to reflect children's sophisticated reasoning in epistemic judgments.

4. Conclusion

In the first half of the chapter, we reviewed a growing body of Bayesian probabilistic models on reasoning about agents' mental states and actions. Comparisons of these models with human performances have shown that infants, children, and adults reason about agents based on intuitive theories of psychology that specifies how agents plan their actions. Examining the transition of infants' theories of intuitive preference have shown that infants can construct new

concepts and theories through Bayesian inductive learning, consistent with the rational constructivism framework.

In the second half of the chapter, we reviewed Bayesian probabilistic models on pedagogical reasoning and epistemic trust. The pedagogical model has shown that teachers and learners engage in optimal teaching and learning by reasoning about the mental states of each other. The teacher provides evidence that would maximize the learner's belief in the correct hypothesis. The epistemic trust model has shown that children selectively trust informants by inferring the knowledgeability and helpfulness of informants. The model also revealed a shift between 3- and 4-year-olds' assumptions of the helpfulness of informants, and whether this shift can be captured through Bayesian inductive learning within the rational constructivism framework remains to be explored.

Taking these two bodies of work together, we found that the abilities to learn about others – inferring others' mental states – emerge during infancy and guide our reasoning throughout childhood and adulthood, but the abilities to learn from others – pedagogical reasoning and epistemic reasoning – emerge later during preschool. How should we reconcile the early competence in mental states reasoning and the later developing abilities in pedagogical and epistemic reasoning? As a concrete example, infants as young as 10-month-olds are capable of socially evaluating prosocial and antisocial agents based on mental state inferences, but 3-year-olds assume that informants are uniformly helpful when they learn new words from them. Future work should further integrate these two lines of research. For instance, components of the intuitive psychology models could be incoperated into the pedagogical and epistmic trust models to understand how mental state reasoning relate to and potentially contribute to the development of pedagogical and epistemic reasoning.

Given the scope of this chapter, we could not elaborate on many other social learning models that might be of interests to researchers in child development. But before we close, we would like to highlight one other group of models that focus on learning and reasoning about social groups, namely the social-structure learning models (Gershman & Cikara, 2020; Gershman et al., 2017; Lau et al., 2018; Martinez et al., 2021). These models sort individuals into latent groups based on the assumption that individuals in the same group tend to behave similarly. Social-structure learning models provide a promising approach to investigate many issues on children's reasoning about social groups, such as intergroup bias, stereotyping, essentialist beliefs, and so on.

Bayesian probabilistic models have provided formal accounts of many aspects of cognitive, language, and social-cognitive development, and inspired new empirical investigations. We hope that this review will inspire fruitful future research to tackle the open questions in both child development and computational modeling.

References

Aboody, R., Velez-Ginorio, J., Santos, L., & Jara-Ettinger, J. (2018). When teaching breaks down: Teachers rationally select what information to share, but misrepresent learners' hypothesis spaces. In T.T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), Proceedings of the 40th Annual Conference of the Cognitive Science Society (pp. 70–75). Austin, TX: Cognitive Science Society.

Aguiar, A., & Baillargeon, R. (1999). 2.5-Month-Old Infants' Reasoning about When Objects Should and Should Not Be Occluded. *Cognitive Psychology*, *39*(2), 116–157. https://doi.org/10.1006/cogp.1999.0717

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*(4), 0064. https://doi.org/10.1038/s41562-017-0064

Bass, I., Gopnik, A., Hanson, M., Ramarajan, D., Shafto, P., Wellman, H., & Bonawitz, E. (2019). Children's developing theory of mind and pedagogical evidence selection. *Developmental Psychology*, *55*(2), 286–302. http://dx.doi.org/10.1037/dev0000642

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. https://doi.org/10.1016/j.cognition.2010.10.001

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, *108*(Supplement_2), 10918–10925. https://doi.org/10.1073/pnas.1100290108

Bridgers, S., Jara-Ettinger, J., & Gweon, H. (2019). Young children consider the expected utility of others' learning to decide what to teach. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-019-0748-6

Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2011). Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. *Cognition*, *120*(3), 331–340. https://doi.org/10.1016/j.cognition.2010.12.001

Butler, L. P., Gibbs, H. M., & Tavassolie, N. S. (2020). Children's developing understanding that even reliable sources need to verify their claims. *Cognitive Development*, *54*, 100871. https://doi.org/10.1016/j.cogdev.2020.100871

Butler, L. P., & Markman, E. M. (2012). Preschoolers Use Intentional and Pedagogical Cues to Guide Inductive Inferences and Exploration: Pedagogical Cues and Exploration. *Child Development*, *83*(4), 1416–1428. https://doi.org/10.1111/j.1467-8624.2012.01775.x

Butler, L. P., Schmidt, M. F. H., Bürgel, J., & Tomasello, M. (2015). Young children use pedagogical cues to modulate the strength of normative inferences. *British Journal of Developmental Psychology*, *33*(4), 476–488. https://doi.org/10.1111/bjdp.12108

Butler, L. P., Schmidt, M. F. H., Tavassolie, N. S., & Gibbs, H. M. (2018). Children's evaluation of verified and unverified claims. *Journal of Experimental Child Psychology*, *176*, 73–83. https://doi.org/10.1016/j.jecp.2018.07.007

Blythe, P. W., Todd, P. M., & Miller, G. F. (1999). How motion reveals intention: Categorizing social interactions. In G. Gigerenzer, P. M. Todd, & The ABC Research Group, *Simple heuristics that make us smart* (pp. 257–285). Oxford University Press.

Corriveau, K. H., Fusaro, M., & Harris, P. L. (2009). Going with the flow: Preschoolers prefer nondissenters as informants. *Psychological Science*, *20*(3), 372–377. https://doi.org/10.1111/j.1467-9280.2009.02291.x

Corriveau, K., & Harris, P. L. (2009). Choosing your informant: Weighing familiarity and recent accuracy. *Developmental Science*, *12*(3), 426–437. https://doi.org/10.1111/j.1467-7687.2008.00792.x

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*(1), 111–133. https://doi.org/10.1016/S0364-0213(02)00112-X

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Csibra, G., Gergely, G., Bíró, S., Koós, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of 'pure reason' in infancy. *Cognition*, *72*(3), 237–267. https://doi.org/10.1016/S0010-0277(99)00039-6

Egyed, K., Király, I., & Gergely, G. (2013). Communicating shared knowledge in infancy. *Psychological Science*, *24*(7), 1348–1353. https://doi.org/10.1177/0956797612471952

Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive Psychology*, *59*(2), 154–179. https://doi.org/10.1016/j.cogpsych.2009.03.001

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287–292. https://doi.org/10.1016/S1364-6613(03)00128-1

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165–193. https://doi.org/10.1016/0010-0277(95)00661-H

Gershman, S. J., & Cikara, M. (2020). Social-structure learning. *Current Directions in Psychological Science*, *29*(5), 460–466. https://doi.org/10.1177/0963721420924481

Gershman, S. J., Pouncy, H. T., & Gweon, H. (2017). Learning the structure of social influence. *Cognitive Science*, *41*, 545–575. https://doi.org/10.1111/cogs.12480

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin, 138*(6), 1085–1108. https://doi.org/10.1037/a0028044

Gweon, H., Shafto, P., & Schulz, L. (2018). Development of children's sensitivity to overinformativeness in learning and teaching. *Developmental Psychology*, *54*(11), 2113–2125. https://doi.org/10.1037/dev0000580

Hamlin, J. K., Ullman, T., Tenenbaum, J., Goodman, N., & Baker, C. (2013). The mentalistic basis of core social cognition: Experiments in preverbal infants and a computational model. *Developmental Science*, *16*(2), 209–226. https://doi.org/10.1111/desc.12017

Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development*, *26*(1), 30–39. https://doi.org/10.1016/j.cogdev.2010.09.001

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557–559. https://doi.org/10.1038/nature06288

Hamlin, J. K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Sciences*, *108*(50), 19931–19936. https://doi.org/10.1073/pnas.1110306108

Hsu, A., & Griffiths, T. (2009). Differential use of implicit negative evidence in generative and discriminative language learning. In Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems* (Vol. 22). Curran Associates, Inc. https://proceedings.neurips.cc/paper/2009/file/0f96613235062963ccde717b18f97592-Paper.pdf

Jara-Ettinger, J., Floyd, S., Tenenbaum, J. B., & Schulz, L. E. (2017). Children understand that agents maximize expected utilities. *Journal of Experimental Psychology: General*, *146*(11), 1574–1585. https://doi.org/10.1037/xge0000345

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. https://doi.org/10.1016/j.tics.2016.05.011

Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2015). Children's understanding of the costs and rewards underlying rational action. *Cognition*, *140*, 14–23. https://doi.org/10.1016/j.cognition.2015.03.006

Jara-Ettinger, J., Schulz, L. E., & Tenenbaum, J. B. (2020). The Naïve Utility Calculus as a unified, quantitative framework for action understanding. *Cognitive Psychology*, *123*, 101334. https://doi.org/10.1016/j.cogpsych.2020.101334

Jara-Ettinger, J., Tenenbaum, J. B., & Schulz, L. E. (2015). Not so innocent: Toddlers' inferences about costs and culpability. *Psychological Science*, *26*(5), 633–640. https://doi.org/10.1177/0956797615572806

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1), 1–19. https://doi.org/10.1016/0010-0277(91)90045-6

Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science*, *21*(8), 1134–1140. https://doi.org/10.1177/0956797610376652

Lau, T., Pouncy, H. T., Gershman, S. J., & Cikara, M. (2018). Discovering social groups via latent structure learning. *Journal of Experimental Psychology: General*, *147*(12), 1881–1891. https://doi.org/10.1037/xge0000470

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, *25*(3), 265–288. https://doi.org/10.1016/S0010-0277(87)80006-9

Liberman, Z., & Shaw, A. (2020). Even his friend said he's bad: Children think personal alliances bias gossip. *Cognition*, *204*, 104376. https://doi.org/10.1016/j.cognition.2020.104376

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, E1. https://doi.org/10.1017/S0140525X1900061X

Liu, S., Ullman, T. D., Tenenbaum, J. B., & Spelke, E. S. (2017). Ten-month-old infants infer the value of goals from the costs of actions. *Science*, *358*(6366), 1038–1041. https://doi.org/10.1126/science.aag2132

Lucas, C. G., Griffiths, T. L., Xu, F., Fawcett, C., Gopnik, A., Kushnir, T., Markson, L., & Hu, J. (2014). The child as econometrician: A rational model of preference understanding in children. *PLoS ONE*, *9*(3), e92160. https://doi.org/10.1371/journal.pone.0092160

Ma, L., & Xu, F. (2011). Young children's use of statistical sampling evidence to infer the subjectivity of preferences. *Cognition*, *120*(3), 403–411. https://doi.org/10.1016/j.cognition.2011.02.003

Martinez, J. E., Feldman, L. A., Feldman, M. J., & Cikara, M. (2021). Narratives shape cognitive representations of immigrants and immigration-policy preferences. *Psychological Science*, *32*(2), 135–152. https://doi.org/10.1177/0956797620963610

Mascaro, O., & Sperber, D. (2009). The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition*, *112*(3), 367–380. https://doi.org/10.1016/j.cognition.2009.05.012

Meng, Y., & Xu, F. (2020). How do disparities reproduce themselves? Ground truth inference from utility-maximizing agent's sampling behavior. In S. Denison., M. Mack, Y. Xu, & B.C. Armstrong (Eds.), Proceedings of the 42nd Annual Conference of the Cognitive Science Society (pp. 903–908). Cognitive Science Society.

Meng, Y., & Xu, F. (2021). Naive utility calculus underlies the reproduction of disparities in social groups. In T. Fitch, C. Lamm, H. Leder, & K. Teßmar-Raible (Eds.), Proceedings of the 43rd Annual Conference of the Cognitive Science Society (pp. 1416–1422). Cognitive Science Society.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 255–258. https://doi.org/10.1126/science.1107621

Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology*, *43*(5), 1216–1226. https://doi.org/10.1037/0012-1649.43.5.1216

Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science*, *308*(5719), 214–216. https://doi.org/10.1126/science.1111656

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology*, *33*(1), 12–21. https://doi.org/10.1037/0012-1649.33.1.12

Rhodes, M., Gelman, S. A., & Brickman, D. (2010). Children's attention to sample composition in learning, teaching and discovery: Sample composition and learning. *Developmental Science*, *13*(3), 421–429. https://doi.org/10.1111/j.1467-7687.2009.00896.x

Schmidt, M. F. H., Butler, L. P., Heinz, J., & Tomasello, M. (2016). Young children see a single action and infer a social norm: Promiscuous normativity in 3-year-olds. *Psychological Science*, *27*(10), 1360–1370. https://doi.org/10.1177/0956797616661182

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, *4*(8), 299–309. https://doi.org/10.1016/S1364-6613(00)01506-0

Schütte, F., Mani, N., & Behne, T. (2020). Retrospective inferences in selective trust. *Royal Society Open Science*, *7*(2), 191451. https://doi.org/10.1098/rsos.191451

Shafto, P., Eaves, B., Navarro, D. J., & Perfors, A. (2012). Epistemic trust: Modeling children's reasoning about others' knowledge and intent. *Developmental Science*, *15*(3), 436–447. https://doi.org/10.1111/j.1467-7687.2012.01135.x

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. https://doi.org/10.1016/j.cogpsych.2013.12.004

Shu, T., Bhandwaldar, A., Gan, C., Smith, K. A., Liu, S., Gutfreund, D., Spelke, E., Tenenbaum, J. B., & Ullman, T. D. (2021). AGENT: A benchmark for core psychological reasoning. *ArXiv:2102.12321 [Cs]*. http://arxiv.org/abs/2102.12321

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99*(4), 605–632. https://doi.org/10.1037/0033-295X.99.4.605

Tomasello, M. (2016). Cultural Learning Redux. *Child Development*, *87*(3), 643–653. https://doi.org/10.1111/cdev.12499

Ullman, T. D., & Tenenbaum, J. B. (2020). Bayesian models of conceptual development: Learning as building models of the world. *Annual Review of Developmental Psychology*, *2*(1), 533–558. https://doi.org/10.1146/annurev-devpsych-121318-084833

Wellman, H. M., Kushnir, T., Xu, F., & Brink, K. A. (2016). Infants use statistical sampling to understand the psychological world. *Infancy*, *21*(5), 668–676. https://doi.org/10.1111/infa.12131

Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34. https://doi.org/10.1016/S0010-0277(98)00058-4

Xu, F. (2019). Towards a rational constructivist theory of cognitive development. *Psychological Review*, *126*(6), 841–864. https://doi.org/10.1037/rev0000153

Xu, F., & Griffiths, T. L. (2011). Probabilistic models of cognitive development: Towards a rational constructivist approach to the study of learning and development. *Cognition*, *120*(3), 299–301. https://doi.org/10.1016/j.cognition.2011.06.008

Xu, F. & Kushnir, T. (Eds). (2012). *Advances in Child Development and Behavior Volume 43: Rational Constructivism in Cognitive Development.* Waltham, MA: Academic Press.

Xu, F., & Tenenbaum, J. B. (2007a). Word learning as Bayesian inference. *Psychological Review*, *114*(2), 245–272. https://doi.org/10.1037/0033-295X.114.2.245

Xu, F., & Tenenbaum, J. B. (2007b). Sensitivity to sampling in Bayesian word learning. *Developmental Science*, *10*(3), 288–297. https://doi.org/10.1111/j.1467-7687.2007.00590.x