

Review

Inferential social learning: cognitive foundations of human social learning and teaching

Hyowon Gweon^{1,*}

Social learning is often portrayed as a passive process of copying and trusting others. This view, however, does not fully capture what makes human social learning so powerful: social information is often ‘curated’ by helpful teachers. I argue that both learning from others (social learning) and helping others learn (teaching) can be characterized as probabilistic inferences guided by an intuitive understanding of how people think, plan, and act. Consistent with this idea, even young children draw rich inferences from evidence provided by others and generate informative evidence that helps others learn. By studying social learning and teaching through a common theoretical lens, inferential social learning provides an integrated account of how human cognition supports acquisition and communication of abstract knowledge.

What makes human social learning so powerful, smart, and distinctive?

Humans are hardly the only species that learns from others. Chimpanzees, crows, fish, and even bumblebees engage in various forms of social learning [1–5]. Yet only humans learn from others in ways that cultivate rich repertoires of knowledge, and only humans have developed cultural institutions, from parenting practices to formal education, to facilitate social learning at scale. What makes human social learning so powerful, smart, and distinctive?

Social learning – learning from others – has long been a topic of research in cognitive science and beyond [6–14]. Much of this work, understandably, has focused on ‘what learners do’ to learn from others; the underlying assumption is that studying what is special about how humans learn from others can help reveal the reasons behind the power of human social learning. This line of inquiry has been fruitful, identifying a range of behaviors that seem unique to our species. For instance, research suggests that even infants are sensitive to cues that signal others’ intent to teach (e.g., pointing, gaze, and infant-directed speech), interpreting information accompanied by these cues as generalizable knowledge (natural pedagogy [9]); children are faithful imitators, even copying actions that are causally unnecessary (overimitation [12]); children also trust what they are told, sometimes almost to a fault, unless they are given blatantly false information (e.g., calling a ball a shoe, epistemic trust [13,14]).

For all the richness of this literature, however, the learner-centric approach provides only a partial understanding of what makes human social learning truly special. First, this literature has studied how humans learn from others separately from how humans enable, facilitate, and promote others’ learning; thus, the literature on social learning has remained rather disconnected from the literature on teaching and communication. Second, much of the existing literature on social learning has also advanced separately from theories of learning broadly construed, particularly of how humans learn from observing and exploring the physical world. To explain how children know when to copy, what to generalize, and whom to trust, researchers have generally appealed

Highlights

Human knowledge is abstract, structured, and theory-like, and cannot be directly copied or transferred. Instead, learners must interpret the meaning of others’ behaviors, which are often generated by helpful, knowledgeable teachers.

Rather than studying social learning separately from teaching, inferential social learning explains these processes as two sides of the same coin: the interpretation and generation of evidence in social contexts.

Children interpret and evaluate the meaning of evidence based on how it was generated, and even generate helpful evidence by considering others’ mental states and utilities; far from passive and credulous, children are active social learners and teachers.

Inferential social learning paints an integrated picture of how human cognition gives rise to a form of learning and communication that is powerful, smart, and distinctively human.

¹Department of Psychology, Stanford University, 450 Jane Stanford Way, Stanford, CA 94305, USA

*Correspondence: gweon@stanford.edu (H. Gweon).



to associative processes [11] or specialized mechanisms evolved specifically for social learning (Box 1). Thus, compared to the image of ‘children as scientists’ who are curious and intrinsically motivated to explore the world [15,16], the image of children as social learners has been more passive, credulous, and deferential.

Here, I present a rather different picture of social learning. This picture is not only about how we learn from others but also about how we help others learn, framed within a broader perspective on how humans think, learn, and communicate in the social context. By studying both social learning and teaching through the same theoretical lens, we can begin to understand how the two processes work together towards a common, shared epistemic goal [17,18], giving rise to a form of learning and communication that is powerful, smart, and distinctively human.

Learning from others, helping others learn: towards a unified account of human learning and teaching

To understand what it means to learn from others, it is useful to first consider the nature of what we as humans know. Human knowledge is far more than a collection of adaptive behaviors or isolated facts about the world; our minds construct abstract, structured, and theory-like beliefs that reflect our understanding of the world and how to survive in it. These beliefs manifest not only as formal, scientific theories, but also as intuitive (folk, naïve) theories [19–21] that are baked into every aspect of our lives: old wisdoms, new technologies, cultural norms, social conventions, and even artistic creations. Critically, their abstract contents cannot be copied or transferred directly from one mind to another [22], presenting a serious challenge for accounts that emphasize selective copying as the main vehicle of cultural transmission. They also raise questions about

Box 1. From studying ‘whether’ to ‘how’ children imitate, trust, and generalize

A large body of literature on social learning has focused on imitation [11,75–78] as a major mechanism for cultural transmission that provides an easier, cheaper alternative to individual (‘asocial’) learning. Given the prevalence of copying and mimicry in non-human animals (often characterized as transmission biases [8,75] or social learning strategies [10,79,80]) and its apparent benefit of reducing the costs of acquiring adaptive behaviors, it may seem reasonable that human social learning also has similar evolutionary roots. Other theories also emphasize specialized mechanisms that ‘trigger’ or ‘elicit’ social learning, such as sensitivity to particular behaviors (e.g., pointing, gaze, and infant-directed speech) as a cue to trust [14] and generalize [9] information provided by others.

The idea that humans have evolved specialized mechanisms for social learning has a strong intuitive appeal. The empirical findings, however, seem rather mixed. Children indiscriminately over-imitate unnecessary actions [12] but they also selectively imitate intentional actions depending on the context [81,82] and even parse out irrelevant actions when demonstrating for others [60]; they trust anyone who points [83] or just those who were reliable in the past [14]; infants consider pedagogical cues as a signal to generalize object properties [44] but they also consider the probability of sampled evidence [37]. How do we make sense of these findings?

From the perspective of inferential social learning, these seemingly conflicting results actually reflect rational responses to observed evidence given the social context; the question at stake is not ‘whether’ children imitate, trust, and generalize (of course they do) but ‘how’ they use their intelligence to decide when to imitate, whom to trust, and how far to generalize. Rather than shifting the explanatory burden to specialized mechanisms or hypothesizing a particular kind of cognition specifically for teaching [72], inferential social learning makes explicit assumptions about what cognitive capacities might support these decisions; grounded in a computational framework, it also allows researchers to formalize the underlying representations and inferential processes to test quantitative predictions about early competence and developmental change.

Importantly, inferential social learning does not aim to replace or refute the idea that humans may have some specialized mechanisms that support early imitation and generalization. Rather, it points out their insufficiency for explaining how humans acquire, communicate, and accumulate abstract knowledge by selectively engaging in various epistemic practices as learners and as teachers. By explaining these abilities as interpretation and generation of evidence in social contexts, we can construct a more unified understanding of how cognition shapes seemingly disparate aspects of children’s behaviors – not just as learners and teachers but also as explorers, cooperators, and even competitors.

how social learning enables acquisition of abstract knowledge well before children can benefit from verbal instruction, or in cultures where direct adult-to-child teaching is rare [23].

In the past decade, collaborative efforts at the intersection of developmental and computational cognitive science have made significant progress in explaining how humans learn from observed evidence. Characterizing human intelligence as a powerful inference engine, Bayesian hierarchical models have formalized learning as probabilistic inferences that operate over abstract, structured representations of the world [24]. Consistent with this idea, developmental research has identified early-emerging signatures of these representations and inferential processes. Infants exhibit abstract, theory-like knowledge in core domains (e.g., objects [25], agents [26], number [28], and geometry [29]) that allows them to predict future events [30], explore events that violate these predictions [31], perform actions to learn from self-generated evidence [32,33], and even infer the cause of their failures when such actions fail [34]. This literature has thus emphasized the active nature of early learning: much like scientists, children use observed evidence to build, test, and revise their intuitive theories about how the world works [15,16].

However, learning does not occur in isolation. Humans, especially early in life, are rarely left all by themselves. Although learning from exploration and social learning are often considered fundamentally different ways of learning, young children's exploration often unfolds in the presence of (or is even scaffolded by) their caregivers, siblings, and peers [35]. Even many existing developmental studies on children's causal reasoning (e.g., "blicket-detector" paradigms, [36]) can be reconstrued as studies on how children interpret evidence generated by adult experimenters.

This insight has deep implications for what it means to have a unified theory of human learning. Rather than focusing on how humans learn as individual learners (and treating 'social' learning as something different), it must explain how learning occurs in a range of contexts that vary in the degree to which others are involved, from self-guided exploration (low) to explicit instruction (high). Only then can we study learning as a whole, and the ways in which humans seamlessly transition between learning (as observers, explorers, scientists, and students) and facilitating others' learning (as onlookers, communicators, demonstrators, and teachers).

Inferential social learning: learning from evidence generated by others

Inferential social learning (Figure 1, Key figure) is a framework that characterizes social learning as inferential at its core; humans learn by drawing inferences from evidence generated by others (i.e., observations of and interactions with others) just as they learn by drawing inferences from observed physical events or self-generated evidence from exploration.

If 'social' and 'individual' learning rely on the same inferential machinery, then what differentiates the two? One key difference is the learner's understanding of how the evidence is generated, by whom, for whom, and why (i.e., the sampling process [37]; Box 2 and Figure 1A). People's goal-directed behaviors are caused by, and thus reflect, their intuitive theories about the world. Unlike naïve learners who often produce noisy data due to their lack of prior knowledge and expertise, knowledgeable agents can perform more targeted causal interventions based on their understanding of how the world works. Thus, learning from evidence generated by these agents can be more accurate, effective, and efficient than learning from self-generated evidence. Critically, simply copying the observed behaviors would be insufficient for acquiring the abstract, structured knowledge that resides in others' minds; instead, learners must use the observed evidence, such as utterances, actions, demonstrations (and even emotional expressions [38]), as information to recover the contents of others' minds that gave rise to those behaviors.

Key figure

Inferential social learning

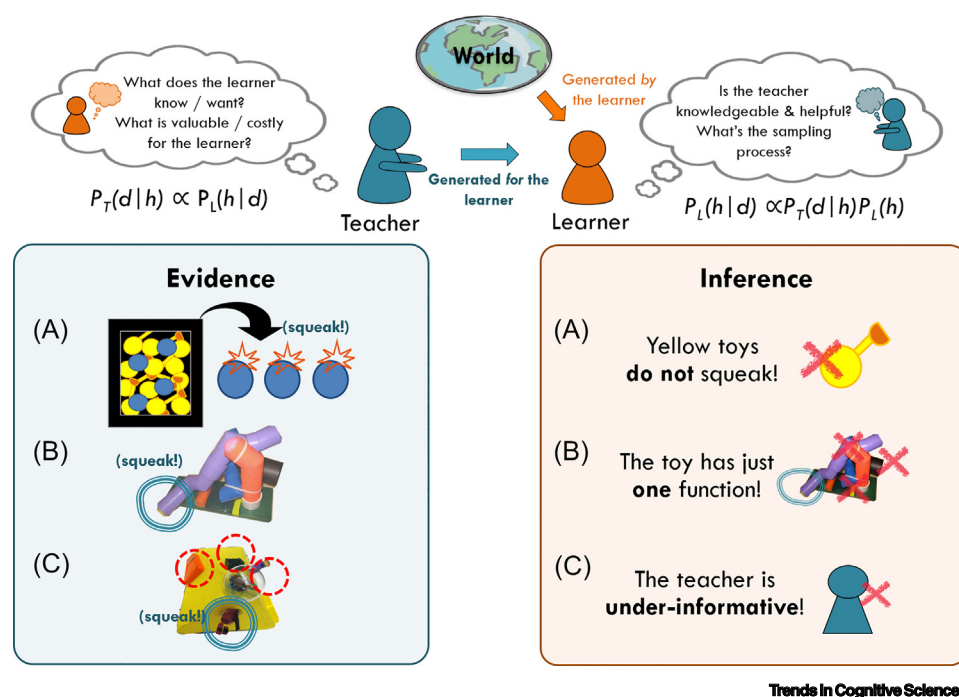


Figure 1. Human learners draw inferences from observed evidence either from their own exploration or from others' behaviors. Such learning can be particularly powerful when the evidence is generated by a knowledgeable, helpful agent (i.e., a teacher, T) whose goal is to teach the learner (L) by selecting the best set of data (d) given the target hypothesis (h); the learner and the teacher work together towards a common epistemic goal (i.e., the learner's acquisition of useful knowledge) by drawing cooperative inferences about others' mental states (e.g., goals, beliefs, and desires) and utilities (costs and rewards of goal-directed actions). The lower panels show the evidence generated by a teacher and the learner's inference. (A) Three blue toys that squeak, drawn from a box of mostly yellow toys, suggests that the teacher deliberately chose the blue ones; given such selective (strong) sampling, the learner infers that the property of the blue toys do not generalize to the yellow toys [37] (Box 2). (B) When a teacher pedagogically demonstrates one interesting function of a toy, the learner interprets the absence of additional demonstrations as evidence for the absence of other functions [48,49] (Box 2). (C) When the learner already knows about the toy's additional functions, the learner can evaluate the teacher's omission of evidence as under-informative pedagogy [50,51] (Box 3).

Drawing these inferences, however, is a nontrivial feat. To learn from others' behaviors, the learner needs a causal understanding of how and why those behaviors came to be, that is, a generative model of other minds. This model reflects our intuitive theories about how internal, unobservable states of an agent give rise to its observable behaviors. Critically, these internal states not only include the agent's subjective mental states (i.e., goals, beliefs, and desires) but also the agent's utilities (i.e., costs and rewards). While decades of developmental research have examined children's reasoning about others' mental states, often referred to as 'theory of mind', more recent work suggests that the traditional notion of belief-desire psychology is insufficient to explain how humans make sense of others' behaviors. When an agent chooses a cookie over a cracker, the 'standard' explanation is that the agent wants the cookie; however, agents might forego a desired goal because the expected costs for obtaining it exceeds the expected

Box 2. Foundations of inferential social learning: early-emerging inferential abilities

Knowing when to learn from others

Learning from exploration works insofar as learners can produce meaningful data, but not all interventions are effective. Thus, learners must decide when to explore and when to rely on others. Evidence suggests that even infants approach this decision as a fundamental inference problem; when infants imitate but fail to produce the expected outcome, they use the statistical dependencies in others' actions and outcomes (who succeeded or failed) to infer the cause of their own failures (i.e., is it me, or the toy?), and decide to explore another toy versus seek others' help depending on the likely cause [34].

Knowing what to learn from others

When someone demonstrates an interesting property of an object (e.g., squeezing the toy makes it squeak), learners face a question about how far to generalize the property: do all other toys squeak too, or just this one? Research suggests that even infants understand differences in others' sampling processes and modulate their inductive generalization accordingly [37]. For instance, when an agent draws a sample of three blue toys (each of which squeaks when pressed) from a box that contains both blue and yellow toys, infants' generalization of the squeaking property to the yellow toys depends on the proportion of yellow and blue toys in the box; for instance, if the blue toys are rare, the sample indicates that the agent engaged in 'strong' (versus weak, random) sampling to produce a biased sample (e.g., the actor chose just the blue toys because the yellow ones do not have the property) (see Figure 1A in main text). In the absence of demonstrated properties, strong sampling can indicate the agent's intrinsic preference for the exemplars [84,85].

Learning from pedagogically sampled evidence

Seeing a demonstration of a toy's property might lead the learner to ask: does the toy also have other properties or only this one? Children's ability to consider the sampling process to draw flexible inductive generalization [37] also supports powerful inferences that go beyond the evidence in pedagogical contexts (see Figure 1B in main text). Given a pedagogical demonstration of a toy's property, children (as young as age 2 years) infer that it is the toy's only property and explore the toy less than when they observe an instrumental or incidental action that reveals the same property [48,49]. These results have been replicated in Yucatec Mayan culture where children rarely receive direct instruction from adults [49], suggesting that children's interpretation of pedagogically generated evidence emerges from an understanding of others' minds and sampling processes rather than prior exposure to culturally specific teaching practices.

rewards (e.g., the cookie is too high up on the shelf). A recent proposal suggests that humans assume other agents are utility-maximizers, and interpret others' actions in terms of their underlying utilities in ways that are intimately tied to their mental states and properties of the external world (naïve utility calculus [27], see also [26,39–41]). There is growing evidence that such an integrated model of other minds is already present early in life [42,43], supporting both forward-inference (i.e., using known states to infer actions) and inverse-inference (i.e., using observed actions to infer hidden states). Inverse-inference, in particular, allows learners to use others' behaviors to infer their knowledge and beliefs that gave rise to those behaviors.

In sum, inferential social learning aims to explain how these foundational cognitive capacities might support the human ability to learn from others and help others learn. Rather than appealing to built-in biases to use certain behavioral cues as an indicator of generalizable information [9,44], inferential social learning instead emphasizes the learner's capacity to flexibly generalize and interpret others' behavior, while explicitly taking into account the role that communicators and teachers play in the learning process. In what follows, I describe how these capacities manifest in how children interpret and evaluate evidence from others (as learners) and how they generate evidence for others (as teachers).

Children as learners

Interpreting the meaning of evidence generated by others

Consider an actor pulling out a tube from a complex-looking toy (Figure 1B), which in turn makes a funny, squeaking sound. A learner who observes the action–outcome association might attend more to the toy's tube (stimulus enhancement), reproduce the action (imitation), or try other ways

to generate the outcome (emulation). Yet, a learner who has a model of other minds (i.e., a generative model of how others' minds give rise to their behaviors) can consider the goals, knowledge, and beliefs that led the actor to perform that action, and draw different inferences depending on the context.

When the action is clearly incidental (e.g., the actor accidentally pulls out the tube), it suggests that the actor does not know how the toy works; the observed action–outcome relationship was produced by chance, and it does not license further inferences about the causal structure of the toy. An instrumental action (e.g., the actor intentionally pulls out the tube to make the toy squeak) does suggest that the actor has a particular goal and knows how to achieve it, but the toy's overall causal structure still remains ambiguous; it is possible that the toy has other functions, but the actor did not care to use them because they are irrelevant to the actor's current goal. By contrast, when the action is clearly instructional (e.g., the actor pedagogically demonstrates that the tube-pulling makes the toy squeak), this strongly suggests that the actor knows how the toy works and is willing to show others how it works. In such contexts, the actor is expected to engage in pedagogical sampling of evidence [45], generating the best set of evidence that helps the observer learn about the toy's causal structure. Learning from such evidence is particularly powerful when the observer (i.e., learner) recognizes the actor's goal and interprets the evidence assuming pedagogical sampling. In this case, in addition to learning that pulling the tube makes the toy squeak, the learner can go beyond the observed evidence and also infer that it is the only function of the toy (if the toy had additional functions, a knowledgeable, helpful teacher would have demonstrated them too).

These intuitions can be formalized as a set of probabilistic inferences that mutually constrain one another, where both the teacher and the learner are generating and interpreting evidence based on a generative model of the other [45–47]. Consistent with these predictions, when children observe an adult's tube-pulling action in incidental, instrumental, or instructional contexts, they draw different inferences about the toy's causal functions and modulate their exploration accordingly [48,49] (Box 2).

Evaluating the quality of evidence generated by others

The power of learning from pedagogically sampled evidence critically depends on the quality of the sampling process: did the teacher successfully select the best set of evidence for the learner? When a teacher fails to do so, the power of social learning can turn into a hazard. Consider the toy example again, where a teacher pedagogically demonstrates one function of a toy [48] (Figure 1B,C), when it actually has a few other interesting functions. Here, the teacher's demonstration is not false, but nonetheless misleading; the learner might (reasonably) go beyond the evidence to (inaccurately) infer that the demonstrated function is the toy's only function. Even though the teacher never provided false information, there was a mismatch between what the learner expected of the teacher (pedagogical sampling) and what the teacher actually did (omission of relevant evidence). This omission, or under-informativeness, constitutes a 'sin' in pedagogical contexts; the teacher could and should have shown additional evidence, yet failed to do so, leading the learner to have an inaccurate belief about the toy.

Naïve learners, unfortunately, have no choice but to be misled; they lack the relevant knowledge to detect the omission. However, learners with prior knowledge can recognize and evaluate the omission as a 'sin'. Consistent with this prediction, when a teacher demonstrates a single function of a toy to another learner who had never seen it before, preschool-aged children rate the teacher as less helpful when they know the toy has additional functions than when they know it is the toy's only function [50,51] (Figure 1C). Children also guard themselves against potential

sins of omission by modulating their inferences based on a teacher's past informativeness. Given a teacher's demonstration of a novel toy, children explore the toy more broadly in search of additional functions when the teacher had provided under-informative evidence in the past (e.g., taught just one of four functions of a familiar toy) than when the teacher was fully informative (e.g., taught one function of a familiar, single-function toy) [50].

Importantly, children's evaluations do not reflect an indiscriminate preference for teachers who provide more information; they also avoid learning from over-informative teachers who provide demonstrations that are redundant with the learner's prior knowledge [52]. A recent computational study suggests that these evaluations reflect an abstract understanding of the quality of pedagogically sampled evidence, formalized as the utility of the evidence to the learner based on the amount of sampled evidence (e.g., the number of functions shown), the epistemic value of the evidence, and the space of alternatives the teacher could have shown given their knowledge [53]. Children, like adults, show sensitivity to both the amount and value of information, but their ability to exonerate 'innocent' omissions by considering the teacher's knowledge may continue to develop throughout early childhood.

While prior work on teacher evaluation has focused on children's sensitivity to inaccurate information (e.g., calling a 'horse' a 'pen', see [13,14,54] for reviews), misinformation that children encounter in real-world contexts is often more subtle. By characterizing pedagogy as teachers' generation of evidence to be interpreted by learners, inferential social learning allows us to explain how humans evaluate various forms of misinformation (i.e., accuracy, under-informativity [50,51], and even over-informativity [52]) within the same theoretical framework. Beyond evaluating over- and under-informative teachers, recent computational work also suggests that humans make flexible and fine-grained inferences based on the process by which teachers generate evidence. One recent study asked how adults use others' advice to guide their decision-making: do they simply track the advisor's past accuracy, or do they infer the process by which the advisor generated the advice? In the critical experiment, participants received advice from advisors who were equally knowledgeable and equally accurate but differed in their strategies; one gave conservative advice and the other gave risky advice. Adults' use of advice was better explained by a computational model that jointly infers how the advisor selects advice and what information the advisor has access to, compared to a simpler model that only accepts or rejects advice based on the advisor's past accuracy [55] (see also [56]).

In sum, the way humans evaluate others as teachers goes far beyond a sensitivity to accuracy; it reflects an abstract notion of informativeness that involves reasoning about how evidence generated by one agent (e.g., a teacher) might influence the inferences drawn by another (e.g., a learner). As children develop a more sophisticated model of others' minds (e.g., goals, knowledge, beliefs, and utilities), their evaluations of others' informativeness also become more nuanced. Importantly, the scope of these evaluations extends beyond pedagogical interactions. Parallels between these results and children's evaluation of under- and over-informative speakers in linguistic communication suggest common cognitive foundations across domains that have traditionally been studied separately [57,58] (Box 3).

Children as teachers

Generating useful evidence by considering the learner's mental states

For human knowledge to accumulate and develop over time, learners must eventually become teachers themselves. Although some prior work explored how children learn and teach within the same paradigm [59–61], research on early social learning and teaching has largely remained separate literatures without a unifying theoretical framework. From the perspective of inferential

Box 3. Cooperative communication via actions and words

The main idea behind inferential social learning is not specific to explicitly pedagogical interactions. Rather, it can be applied broadly to explain communication via verbal and non-verbal behaviors. Mutual, recursive reasoning about others' minds, a key assumption in computational models of pedagogical reasoning [46,47], generates the expectation that evidence provided by a teacher should be true, relevant, and sufficient for accurate inference, yet not superfluous for the learner. It is not a coincidence that such an expectation reflects Gricean Maxims, or principles of cooperative communication generally [57,86], also formalized as recursive mental-state reasoning between a speaker and a listener [Rational Speech Acts (RSA) [58]]. Together, RSA and models of pedagogical reasoning provide a productive formal framework for modeling communication as probabilistic inference, grounded in psychological reasoning as a generative process for sampling evidence, either as goal-directed actions [46,53] or as linguistic utterances [58,87]. Their parallels suggest that teacher–learner interactions and speaker–listener interactions might involve similar representations and inferential processes. Consistent with this, children show similar developmental trajectories in their inferences and evaluations in linguistic communication and pedagogical interactions.

Inference

By age 3, children expect speakers to select utterances in a way that helps the listener infer their intended, pragmatic meaning [88]; such *ad hoc* implicature (i.e., implicature based on contextual alternatives rather than lexical alternatives, as in scalar implicature) parallels children's inferences in pedagogical contexts [48,49] which also emerge from expectations about how teachers ought to select evidence (Figure 1A); both inferences require the understanding that a helpful agent would have provided more information if the alternative hypothesis were true.

Evaluation

Preschool-aged children can reject literally true but pragmatically misleading statements (e.g., 'some of the blickets have a crayon' when all blickets have crayons), but unlike older children, they succeed only when the appropriate scalar alternatives are made clear in context [89,90]. This resembles children's evaluation of under-informative pedagogy where, unlike older children, 4- and 5-year-olds penalize an under-informative teacher only after having seen a fully-informative teacher [51]. This suggests that both evaluations require an understanding of what a helpful communicator would have said or demonstrated (Figure 1B).

Language (speech acts) and action-based demonstrations involve different modalities that offer distinct advantages in communication (e.g., a few words can replace costly demonstrations, or a demonstration can be worth a thousand words) and are often studied in separate literatures. Despite their differences, the way we interpret and evaluate others' communication may rely on similar cognitive processes [91] that are modulated by the level of cooperative expectations between those who generate evidence (speakers and teachers) and those who interpret the evidence (listeners and learners).

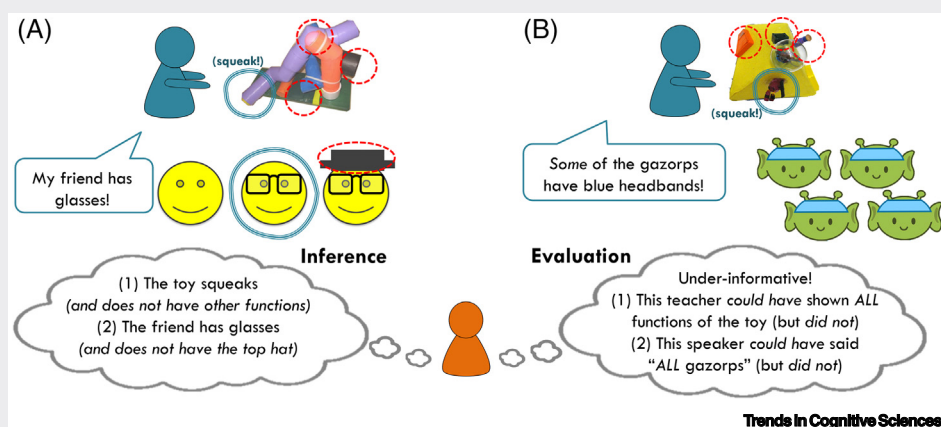


Figure 1. Parallels between pragmatic inferences and evaluation in pedagogical contexts and linguistic communication. (A) Similar to how children infer that a toy has no additional functions when a teacher demonstrates just one [48,49], given the speaker's utterance, children infer that the friend is the second agent (glasses but no hat), even though the utterance is also consistent with the third agent (glasses and hat) [88]. (B) Similar to how they evaluate teachers who omit relevant information [51], preschool-aged children can evaluate the speaker's utterance as infelicitous, but only when they have seen accurate and relevant uses of the word 'some' (e.g., used to describe three of four gazorps sharing a property).

social learning, children's abilities as learners, especially their evaluations of others' informativeness, raise an intriguing possibility: insofar as children understand what constitutes informative evidence for another learner, they also ought to be able to generate such evidence as teachers themselves. Recent work has begun to provide empirical evidence for this hypothesis, suggesting common cognitive roots for our abilities as learners and as teachers.

Consider a toy with 20 identical buttons, only three of which play music (Figure 2A). If you had to teach someone how the toy works, how many buttons would you press? For learners who know nothing about the toy, it is important to demonstrate all 20 buttons. Given that all buttons look identical, selectively demonstrating only the three working buttons is likely to lead the learner to draw a reasonable but inaccurate inference that all buttons play music. Such selective evidence, however, might suffice for learners who had already seen similar toys that also have just a few working buttons; they would resist generalizing the evidence and accurately infer that only the three demonstrated buttons play music. In fact, for those learners, demonstrating all 20 buttons (17 of which are inert) would be unnecessary, inefficient, and not worth the effort. Consistent with this intuition, by late preschool years, children understand that what counts as 'the right amount

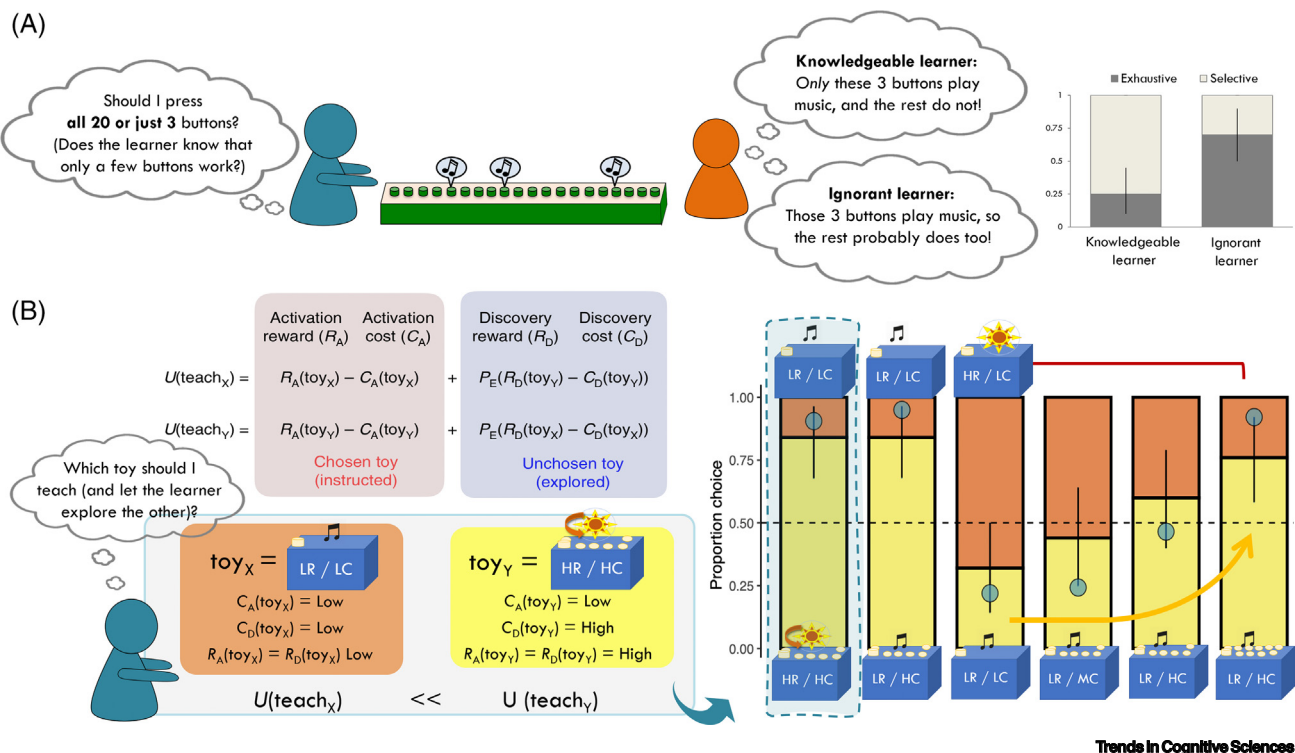


Figure 2. Children as teachers. (A) Pressing all 20 buttons (exhaustive demonstration) guarantees accurate inference, but it is costly (effortful) and unnecessary for a knowledgeable learner who already expects the toy to have just a few working buttons; pressing just the three working buttons (selective demonstration) is less costly and sufficient for a knowledgeable learner (top thought bubble) but misleading for a learner who lacks prior knowledge (bottom thought bubble). The graph shows the proportion of children who provided exhaustive or selective demonstration depending on the learner's knowledge [52]. (B) Given a choice to teach one of two toys (and let the learner explore the other one), children maximize the learner's utilities by reasoning about the costs a learner might incur from exploration (discovery costs), the rewards gained from discovery (discovery reward), as well as how teaching might minimize the discovery costs while ensuring the reward [67]. In this condition (marked by the blue broken line) the teacher must choose between a low-reward (LR, music), low-cost (LC, single button) toy and a high-reward (HR, spinning light globe), high-cost (HC, multiple buttons) toy. By choosing to teach the HR/HC toy, the teacher minimizes the high discovery cost and ensures the learner reaps the high reward. The graph shows six different conditions where children were presented with different pairs of toys, and the proportion of children who chose each toy. Green dots indicate model predictions from the 'full model' that considers the rewards and costs of being taught (chosen toy) and exploration (unchosen toy).

of evidence' depends on the learner's prior knowledge, not only in how they evaluate teachers but also in how they generate evidence as teachers [52]; while children take the time and effort to demonstrate all 20 buttons for a learner who had never seen the toys before, they resist being over-informative and press just three buttons for a learner who has already seen similar toys.

Children can also generate the appropriate set of evidence based on the learner's goal and competence. In one study [62], 4- to 6-year-old children first explored a rather complex causal system with two potential causal variables and two causal effects; either blue or yellow blocks could be placed on either black or white mats to activate either red or green lights. Importantly, one of the mats was placed further away from the rest of the toy, on the other side of the room, and was thus much costlier (harder) to use. Given a chance to play with the toy, children learned, from noisy self-generated evidence, which variable controls the effects of the toy: the color of blocks determines which light turns on, regardless of the mat color. When they were asked to teach a naïve learner, however, they did not reproduce all the evidence they had gathered first-hand. Instead, they generated evidence tailored to the learner's wants (i.e., the learner's goal) and needs (i.e., the learner's competence). Children were more likely to go through the trouble of providing causally informative evidence that fully de-confounded the two variables when the learner wanted to learn how the toy works compared to when the learner simply wanted to observe the toy's effect. Similarly, children provided more evidence when the learner appeared to be incompetent and needing of help, compared to when the learner was introduced as an exceptionally competent learner.

In light of prior work on children's sampling of evidence as teachers [63,64], these results are noteworthy for two reasons. First, in line with prior work, they are consistent with the idea that successful teaching requires theory of mind [59,65,66]. Children in these studies considered the learner's goals and prior knowledge to choose the best set of evidence for the learner. Second, they also suggest that a traditional notion of theory of mind, one that focuses on reasoning about goals, beliefs, and desires, is insufficient to fully explain these results. Children's resistance to providing more than what was necessary (e.g., pressing just three buttons for learners with prior experience [52]) suggests that they also considered the 'cost' of communication, assessing whether the learner's information gain from the additional evidence would be worth the cost of generating it. Just as children's evaluation of teachers reveals their expectations about others' informativeness in cooperative communication (i.e., Gricean Maxims) (Box 3), children's generation of evidence as teachers reveals their willingness and ability to act as cooperative communicators themselves.

Generating useful evidence by considering the learner's utilities

Children's resistance to superfluous demonstrations is consistent with the idea that they are sensitive to the utility of demonstrations. Nonetheless, the studies mentioned above leave open a critical question: whose utility did children consider? It is possible that children were only trying to minimize their own costs (i.e., the effort and time for generating the evidence) rather than considering the learner's utility. A recent study combined computational and developmental approaches to provide more direct evidence for the hypothesis that children, as teachers, can maximize others' expected utilities when deciding what to teach [67].

Learners, who still have much to learn, are often ignorant of the cost-reward structure of the world, making them susceptible to ineffective exploration (e.g., costly yet minimally rewarding). However, knowledgeable individuals (i.e., those who already understand what is valuable or costly to learn), can prioritize teaching things that increase the learner's rewards or reduce their costs of learning. In doing so, they can help learners benefit through a combination of instruction

(which ensures that learners acquire the most critical, valuable knowledge) and exploration (which allows learners to explore and discover additional things beyond what was taught). Within the framework of inferential social learning, such prioritization can be formalized as a decision to provide information that maximizes the learner's expected utilities.

In this study, 5- to 7-year-old children chose which one of two toys to teach to a naïve learner (Figure 2B). Across different conditions, the toys differed in their expected cost of discovering their causal mechanism (i.e., difficult or easy to learn from exploration) and the expected reward of the functions (i.e., whether the toy produced an exciting, high-reward effect or a dull, low-reward effect). Children's choices of what to teach could not be explained by simpler models that selected the most enjoyable (high-reward) or complicated (high-cost) toys, nor by models that only maximized the utilities of being taught or the utilities of learning from exploration. Rather, their choices were most consistent with a model that computes the learner's expected costs and rewards across both toys (i.e., being taught one toy and exploring the other) and chooses the plan that maximizes the expected utility. These results go beyond showing that children, as teachers, teach what others do not know. Even before entering formal schooling, children can infer the consequence of others' ignorance on their expected utilities (i.e., ignorance about a high-cost toy leads to high exploration costs with low chances of reaping its reward) and can make choices as teachers that maximize others' utilities.

To study what and how much evidence children generate as teachers, prior work on early development of teaching (including the work mentioned above [52,62]) has often provided participants with a preestablished goal (e.g., teaching how a particular toy works). Good teaching, however, is just as much, if not more, about deciding 'what' to teach in the first place. By conceptualizing these decisions as a prosocial utility-maximization, inferential social learning explains how human cognition supports 'curated' accumulation of cultural knowledge that prioritizes high-utility knowledge to be transmitted to the next generation.

This conceptualization of teaching also helps explain both the widespread presence of pedagogical practices and the cross-cultural differences in the degree to which societies rely on these practices. Cultures can vary in the kinds of knowledge and skills that make up their cultural repertoire. When a society faces the need to share large amounts of abstract, structured knowledge (e.g., reading or counting systems), which can be challenging or nearly impossible to learn without being taught, this can create the pressure to develop more standardized, scalable ways to educate their young. From this perspective, the diversity and variability in child-rearing and educational practices across cultures [68,69] may also reflect the distinctively human (yet culturally universal) abilities to select, compile, and curate useful knowledge in ways that balance our collective utilities.

Concluding remarks

The studies reviewed here provide initial empirical support for inferential social learning. Humans, including young children, learn from evidence generated by others and generate useful evidence to help others learn. Both processes are inherently communicative and cooperative in nature and are supported by cognitive capacities that emerge early in life: domain-general inferential abilities [24] that operate over abstract (often domain-specific) representations of how the world works and how other people think, plan, and act. This view goes beyond existing theoretical perspectives in a number of ways.

First, it explains how humans learn from both self-generated evidence and evidence provided by others, painting a curious, active, scientist-like image of young children as both independent

Outstanding questions

Inferential social learning has been applied to explain how humans, especially young children, learn and communicate about the external world (e.g., objects, causal mechanisms, and other people). Can it also explain how humans learn and communicate about the inner world (i.e., the self)? How do children learn about the self by drawing inferences from others' feedback (e.g., praise and criticisms), and how do they generate evidence about the self to revise or correct others' beliefs about the self?

Although existing work on social learning has focused on learning from verbal instruction or action-based demonstrations, recent work suggests that children also learn from others' emotional expressions. How do humans integrate social signals from different modalities (e.g., speech, language, and emotional expressions) to learn from others, and how do they use different modalities as communicators to convey their message effectively?

Humans not only engage in cooperative interactions towards their epistemic goals, but also leverage social information to maximize their gains (as learners) or control the flow of information to outcompete others (as communicators). How does inferential social learning incorporate the role of motivation? Can these competitive social decisions and interactions be explained as modulations in how individuals consider their own versus others' expected utilities?

Given what we know about neural mechanisms that support social perception, mental-state reasoning, and utility-based decision-making, how can we integrate computational, developmental, and neuroimaging methods to better understand how inferential social learning is implemented in the brain?

Most machine learning and artificial intelligence (AI) algorithms for social learning rely on simple copying or imitation. How can we use insights from children's abilities as learners and as teachers to develop machines that can learn more effectively from humans and communicate more naturally with humans?

explorers and social learners. Beyond the focus on well-studied sources of social information such as speech and action-based demonstrations, this perspective calls for more research on how humans (especially children) learn by integrating multiple sources of information, including emotional expressions [38] or evaluative feedback from others [70].

Second, rather than treating social learning and teaching as fundamentally different capacities, it explains both within the same integrated framework as the interpretation and generation of evidence in social contexts. Regardless of how they learned (i.e., from self-generated evidence or evidence provided by others), humans can seamlessly transition from learners to teachers who can generate structured evidence tailored to maximize others' learning [52,62]. This perspective also urges us to reevaluate the assumption that a traditional notion of theory of mind (i.e., belief-desire psychology) is sufficient to explain how humans teach [65,66,71,72]; children's behaviors and decisions reflect a rich, intuitive model of other minds that critically incorporates others' utilities as well as their mental states.

Finally, it generates exciting directions for future work that pushes the field towards a unified account of learning and communication (see [Outstanding questions](#)). It calls for work that extends inferential social learning to explain a broader set of phenomena, including interdisciplinary research that explores the cognitive and neural bases of how individuals capitalize on social information to learn and maximize their gain [73]. It also demands more rigorous explanations of developmental change in children's abilities as learners and teachers in ways that reflect the development of their underlying cognitive capacities. Of course, children's ability to hold, process, and remember information, along with their mastery of words and actions to articulate their knowledge, also continues to develop throughout childhood; thus, their apparent failures, both as learners and

Box 4. Pedagogy and cumulative culture: why distinctively human?

Many non-human species show basic inferential capacities [92,93,106], and primates in particular have a surprisingly rich understanding of others' mental states [94,95]. Rather than relying on reflexes that are insensitive to the context (e.g., teaching-like behaviors in meerkats [96]), they imitate human actions selectively [97,98], coordinate with their conspecifics towards a common goal [99], and even engage in tool-sharing behaviors [2]. What, then, makes humans so different from these species?

One possibility is that non-human primates are still limited in their representational capacities [100,107] both about the world and of other minds, and are thus unable to generate or acquire abstract, structured, symbolic contents characteristic of human knowledge. Furthermore, evaluation and generation of evidence (unlike interpretation) require recursive reasoning about how one's communicative behaviors impacts another's epistemic state given their mutual expectations about the context. Such recursive reasoning may be limited in non-human primates, limiting their ability to grasp (as learners) or utilize (as teachers) on the intentional structure of pedagogy.

Another answer may lie in the motivations that drive us to learn from others and help others learn. Beyond epistemic motivations, the findings described in this paper wouldn't exist if children didn't care to connect, affiliate, or cooperate with others [74] (see [78,101,102] for discussions of socially motivated imitation), which stand in stark contrast even to our closest primate relatives [103]. While inferential social learning does not make explicit commitments about the motivational origins of learning and teaching, motivation does have a role in the picture, not only as a general force that drives our behaviors, but also as a product of cognitive capacities that can generate context-specific desire to learn, help, and teach: to want to remedy others' false belief, we must first represent it as such.

To explain what makes humans special, existing proposals have emphasized either the individuals' cognitive abilities (i.e., "big brains") as drivers of individual discoveries [76,104] or the inherited cultural knowledge [11,75,77] that obviates the need for individual intelligence. In inferential social learning, both are necessary: culture provides the priors that empower the inferences we make, and powerful inferences enrich our cultural knowledge. When humans, as individuals and in groups, make decisions about how best to acquire, revise, and communicate useful information, these decisions can collectively drive accumulation of abstract cultural knowledge that evolves over time. After all, the most remarkable aspect of human intelligence might lie not in its power alone but in our willingness to use it on behalf of others [105].

as teachers, may arise from any and all of these components. Behind their noisy behaviors, however, is a sophisticated, nuanced, yet continuously developing understanding of the world and of other minds that allows them to learn from others and help others learn.

So, what makes human learning so powerful, smart, and distinctive? Rather than a collection of disparate mechanisms for mimicry, imitation, and emulation that tells the learner when to copy, what to generalize, or whom to trust (Box 1), these interactions may be better understood as a natural consequence of our integrative cognitive capacities for assessing the value, meaning, and interpretability of evidence, fueled by our drive to learn, communicate, and connect with others [74] (Box 4). Even as young children, humans capitalize on social learning as a vehicle for intellectual growth and turn this motivation outwards to help others learn. Just as children have much to learn from us, we, as scientists, have much to learn from how children think, learn, and communicate.

Acknowledgments

Many thanks to Carol Dweck and Rebecca Saxe for their support and helpful comments on various versions of this paper; Mike Frank, Laura Schulz, Natalia Véléz, Mika Asaba, Ben Deen, Sophie Bridgers, as well as three anonymous reviewers have also provided helpful feedback on earlier versions of this paper. This work has been supported by a James S. McDonnell Foundation Scholars Award for Understanding Human Cognition, a Jacobs Foundation Early Career Research Fellowship, and an NSF Award BCS-2019567.

Declaration of interests

No interests to declare.

References

- Davis, S.J. *et al.* (2016) Foundations of cumulative culture in apes: improved foraging efficiency through relinquishing and combining witnessed behaviors in chimpanzees (*Pan troglodytes*). *Sci. Rep.* 6, 35953
- Musgrave, S. *et al.* (2016) Tool transfers are a form of teaching among chimpanzees. *Sci. Rep.* 6, 34783
- Cornell, H.N. *et al.* (2011) Social learning spreads knowledge about dangerous humans among American crows. *Proc. R. Soc. B Biol. Sci.* 279, 499–508
- Duffy, G.A. *et al.* (2009) Size-dependent directed social learning in nine-spined sticklebacks. *Anim. Behav.* 78, 371–375
- Loukola, O.J. *et al.* (2017) Bumblebees show cognitive flexibility by improving on an observed complex behavior. *Science* 355, 833–836
- Bandura, A. (1977) *Social Learning Theory*, Prentice-hall, Englewood Cliffs, NJ
- Tomasello, M. *et al.* (1993) Cultural learning. *Behav. Brain Sci.* 16, 495–552
- Boyd, R. and Richerson, P.J. (2005) *The origin and evolution of cultures*, Oxford University Press
- Csibra, G. and Gergely, G. (2009) Natural pedagogy. *Trends Cogn. Sci.* 13, 148–153
- Hopppitt, W. and Laland, K.N. (2013) *Social Learning: An Introduction to Mechanisms, Methods, and Models*, Princeton University Press
- Heyes, C. (2018) *Cognitive Gadgets: The Cultural Evolution of Thinking*, Harvard University Press
- Lyons, D.E. *et al.* (2007) The hidden structure of overimitation. *Proc. Natl. Acad. Sci.* 104, 19751–19756
- Harris, P.L. (2012) *Trusting What You're Told: How Children Learn from Others*, Harvard University Press
- Harris, P.L. *et al.* (2018) Cognitive foundations of learning from testimony. *Annu. Rev. Psychol.* 69, 251–273
- Gopnik, A. *et al.* (1999) *The Scientist in the Crib: Minds, Brains, and How Children Learn*, William Morrow & Co.
- Schulz, L. (2012) The origins of inquiry: inductive inference and exploration in early childhood. *Trends Cogn. Sci.* 16, 382–389
- Bratman, M.E. (2013) *Shared Agency: A Planning Theory of Acting Together*, Oxford University Press
- Tomasello, M. (2019) *Becoming Human: A Theory of Ontogeny*, Belknap Press
- Carey, S. (2011) *The Origin of Concepts*, Oxford University Press
- Kail, F.C. (1989) *Concepts, Kinds, and Conceptual Development*, MIT Press
- Wellman, H.M. and Gelman, S.A. (1992) Cognitive development: foundational theories of core domains. *Annu. Rev. Psychol.* 43, 337–375
- Sperber, D. (1996) *Explaining Culture: A Naturalistic Approach* (1st edn), Blackwell Publishers, Cambridge, MA
- Rogoff, B. *et al.* (2003) Firsthand learning through intent participation. *Annu. Rev. Psychol.* 54, 175–203
- Tenenbaum, J.B. *et al.* (2011) How to grow a mind: statistics, structure, and abstraction. *Science* 331, 1279–1285
- Spelke, E. *et al.* (1992) Origins of knowledge. *Psychol. Rev.* 99, 605–632
- Gergely, G. and Csibra, G. (2003) Teleological reasoning in infancy: the naïve theory of rational action. *Trends Cogn. Sci.* 7, 287–292
- Jara-Ettinger, J. *et al.* (2016) The naïve utility calculus: computational principles underlying commonsense psychology. *Trends Cogn. Sci.* 20, 589–604
- Feigenson, L. *et al.* (2004) Core systems of number. *Trends Cogn. Sci.* 8, 307–314
- Dillon, M.R. *et al.* (2013) Core foundations of abstract geometry. *Proc. Natl. Acad. Sci. U. S. A.* 110, 14191–14195
- Téglás, E. *et al.* (2011) Pure reasoning in 12-month-old infants as probabilistic inference. *Science* 332, 1054–1059
- Stahl, A.E. and Feigenson, L. (2015) Observing the unexpected enhances infants' learning and exploration. *Science* 348, 91–94
- Cook, C. *et al.* (2011) Where science starts: spontaneous experiments in preschoolers' exploratory play. *Cognition* 120, 341–349
- Sim, Z.L. and Xu, F. (2017) Learning higher-order generalizations through free play: evidence from 2- and 3-year-old children. *Dev. Psychol.* 53, 642–651

34. Gweon, H. and Schulz, L. (2011) 16-month-olds rationally infer causes of failed actions. *Science* 332, 1524
35. Weisberg, D.S. *et al.* (2016) Guided play: principles and practices. *Curr. Dir. Psychol. Sci.* 25, 177–182
36. Gopnik, A. and Wellman, H.M. (2012) Reconstructing constructivism: causal models, Bayesian learning mechanisms, and the theory. *Psychol. Bull.* 138, 1085–1108
37. Gweon, H. *et al.* (2010) Infants consider both the sample and the sampling process in inductive generalization. *Proc. Natl. Acad. Sci.* 107, 9066–9071
38. Wu, Y. and Gweon, H. (2021) Preschool-aged children jointly consider others' emotional expressions and prior knowledge to decide when to explore. *Child Dev.* 92, 862–870
39. Lucas, C.G. *et al.* (2014) The child as econometrician: a rational model of preference understanding in children. *PLoS ONE* 9, e92160
40. Baker, C.L. *et al.* (2017) *Rational quantitative attribution of beliefs, desires and percepts in human mentalizing*, 1. Nature Publishing Group, p. 0064–10
41. Jern, A. *et al.* (2017) People learn other people's preferences through inverse decision-making. *Cognition* 168, 46–64
42. Liu, S. *et al.* (2017) Ten-month-old infants infer the value of goals from the costs of actions. *Science* 358, 1038–1041
43. Jara-Ettinger, J. *et al.* (2015) Children's understanding of the costs and rewards underlying rational action. *Cognition* 140, 14–23
44. Csibra, G. and Shamsudheen, R. (2015) Nonverbal generics: human infants interpret objects as symbols of object kinds. *Annu. Rev. Psychol.* 66, 689–710
45. Bonawitz, E. and Shafto, P. (2016) Computational models of development, social influences. *Curr. Opin. Behav. Sci.* 7, 95–100
46. Shafto, P. *et al.* (2014) A rational account of pedagogical reasoning: teaching by, and learning from, examples. *Cogn. Psychol.* 71, 55–89
47. Shafto, P. *et al.* (2012) Learning from others: the consequences of psychological reasoning for human learning. *Perspect. Psychol. Sci.* 7, 341–351
48. Bonawitz, E. *et al.* (2011) The double-edged sword of pedagogy: instruction limits spontaneous exploration and discovery. *Cognition* 120, 322–330
49. Shneidman, L. *et al.* (2016) Learning from others and spontaneous exploration: a cross-cultural investigation. *Child Dev.* 87, 723–735
50. Gweon, H. *et al.* (2014) Sins of omission: children selectively explore when teachers are under-informative. *Cognition* 132, 335–341
51. Gweon, H. and Asaba, M. (2018) Order matters: children's evaluation of under-informative teachers depends on context. *Child Dev.* 89, e278–e292
52. Gweon, H. *et al.* (2018) Development of children's sensitivity to overinformativeness in learning and teaching. *Dev. Psychol.* 54, 2113–2125
53. Bass, I. *et al.* (2021) The effects of information utility and teachers' knowledge on evaluations of under-informative pedagogy across development. *PsyArXiv* Published online March 8, 2021. <http://doi.org/10.31234/osf.io/q6jxe>
54. Sobel, D.M. and Kushnir, T. (2013) Knowledge matters: how children evaluate the reliability of testimony as a process of rational inference. *Psychol. Rev.* 120, 779–797
55. Vélez, N. and Gweon, H. (2018) Integrating Incomplete Information with imperfect advice. *Top. Cogn. Sci.* 329, 1081–1117
56. Hawthorne, D. and Goodman, N.D. (2017) Reasoning about social sources to learn from actions and outcomes. *Decision* 6, 17–60
57. Grice, H.P. (1975) *Logic and Conversation*, Syntax and Semantics vol. 3. Academic Press
58. Goodman, N.D. and Frank, M.C. (2016) Pragmatic language interpretation as probabilistic inference. *Trends Cogn. Sci.* 20, 818–829
59. Ashley, J. and Tomasello, M. (1998) Cooperative problem-solving and teaching in preschoolers. *Soc. Dev.* 7, 143–163
60. Flynn, E. (2008) Investigating children as cultural magnets: do young children transmit redundant information along diffusion chains? *Philos. Trans. R. Soc. B Biol. Sci.* 363, 3541–3551
61. Ronfard, S. *et al.* (2016) Children teach methods they could not discover for themselves. *J. Exp. Child Psychol.* 142, 107–117
62. Gweon, H. and Schulz, L. (2019) From exploration to instruction: children learn from exploration and tailor their demonstrations to observers' goals and competence. *Child Dev.* 90, e148–e164
63. Rhodes, M. *et al.* (2010) Children's attention to sample composition in learning, teaching and discovery. *Dev. Sci.* 13, 421–429
64. Rhodes, M. *et al.* (2015) Controlling the message: preschoolers' use of information to teach and deceive others. *Front. Psychol.* 6, 867
65. Strauss, S. *et al.* (2002) Teaching as a natural cognition and its relations to preschoolers' developing theory of mind. *Cogn. Dev.* 17, 1473–1487
66. Bass, I. *et al.* (2019) Children's developing theory of mind and pedagogical evidence selection. *Dev. Psychol.* 55, 286–302
67. Bridgers, S. *et al.* (2020) Young children consider the expected utility of others' learning to decide what to teach. *Nat. Hum. Behav.* 4, 144–152
68. Shneidman, L. *et al.* (2015) Child-directed teaching and social learning at 18 months of age: evidence from Yucatec Mayan and US infants. *Dev. Sci.* 19, 372–381
69. Boyette, A.H. and Hewlett, B.S. (2017) Teaching in hunter-gatherers. *Rev. Philos. Psychol.* 9, 771–797
70. Asaba, M. and Gweon, H. (2020) Learning about others to learn about the self: early reasoning about the informativeness of others' praise. In *Psychological Perspectives on Praise* (1st edn) (Brummelman, E., ed.), pp. 67–74, Routledge
71. Sobel, D.M. and Letourneau, S.M. (2016) Children's developing knowledge of and reflection about teaching. *J. Exp. Child Psychol.* 143, 111–122
72. Strauss, S. *et al.* (2014) Teaching, naturally. *Trends Neurosci. Educ.* 3, 38–43
73. Vélez, N. and Gweon, H. (2021) Learning from other minds: an optimistic critique of reinforcement learning models of social learning. *Curr. Opin. Behav. Sci.* 38, 110–115
74. Dweck, C.S. (2017) From needs to goals and representations: foundations for a unified theory of motivation, personality, and development. *Psychol. Rev.* 124, 689–719
75. Henrich, J. (2015) *The Secret of Our Success: How Culture is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*, Princeton University Press
76. Pinker, S. (2010) Colloquium paper: the cognitive niche: coevolution of intelligence, sociality, and language. *Proc. Natl. Acad. Sci.* 107, 8993–8999
77. Boyd, R. *et al.* (2011) The cultural niche: why social learning is essential for human adaptation. *Proc. Natl. Acad. Sci.* 108, 10918–10925
78. Legare, C.H. and Nielsen, M. (2015) Imitation and innovation: the dual engines of cultural learning. *Trends Cogn. Sci.* 19, 688–699
79. Rendell, L. *et al.* (2011) Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends Cogn. Sci.* 15, 68–76
80. Kendal, R.L. *et al.* (2018) Social learning strategies: bridge-building between fields. *Trends Cogn. Sci.* 22, 651–665
81. Gergely, G. *et al.* (2002) Rational imitation in preverbal infants. *Nature* 415, 755
82. Yu, Y. and Kushnir, T. (2014) Social context effects in 2- and 4-year-olds' selective versus faithful imitation. *Dev. Psychol.* 50, 922–933
83. Palmquist, C.M. and Jaswal, V.K. (2012) Preschoolers expect pointers (even ignorant ones) to be knowledgeable. *Psychol. Sci.* 23, 230–231
84. Kushnir, T. *et al.* (2010) Young children use statistical sampling to infer the preferences of other people. *Psychol. Sci.* 21, 1134
85. Ma, L. and Xu, F. (2011) Young children's use of statistical sampling evidence to infer the subjectivity of preferences. *Cognition* 120, 403–411
86. Clark, H.H. (1996) *Using language*, Cambridge University Press
87. Degen, J. *et al.* (2020) When redundancy is useful: a Bayesian approach to 'overinformative' referring expressions. *Psychol. Rev.* 127, 591–621

88. Stiller, A.J. *et al.* (2014) *Ad-hoc* implicature in preschool children. *Lang. Learn. Dev.* 11, 176–190
89. Barner, D. *et al.* (2011) Accessing the unsaid: the role of scalar alternatives in children's pragmatic inference. *Cognition* 118, 84–93
90. Skordos, D. and Papafragou, A. (2016) Children's derivation of scalar implicatures: alternatives and relevance. *Cognition* 153, 6–18
91. Baldwin, D. *et al.* (2008) Pragmatics of human action. In *Understanding Events: From Perception to Action* (Shipley, T. F. and Zacks, J.M., eds), pp. 96–128
92. Jelbert, S.A. *et al.* (2019) New Caledonian crows infer the weight of objects from observing their movements in a breeze. *Proc. R. Soc. B* 286, 20182332 Chicago
93. Eckert, J. *et al.* (2018) Intuitive statistical inferences in chimpanzees and humans follow Weber's law. *Cognition* 180, 99–107
94. Krupenye, C. *et al.* (2016) Great apes anticipate that other individuals will act according to false beliefs. *Science* 354, 110–114
95. Drayton, L.A. and Santos, L.R. (2017) What do monkeys know about others' knowledge? *Cognition* 170, 201–208
96. Thornton, A. (2006) Teaching in wild meerkats. *Science* 313, 227–229
97. Tomasello, M. and Carpenter, M. (2005) Intention reading and imitative learning. In *Perspectives on Imitation: From Neuroscience to Social Science* (Hurley, S. and Chater, N., eds), pp. 133–148, The MIT Press
98. Buttelmann, D. *et al.* (2007) Enculturated chimpanzees imitate rationally. *Dev. Sci.* 10, F31–F38
99. Melis, A.P. and Tomasello, M. (2019) Chimpanzees (*Pan troglodytes*) coordinate by communicating in a collaborative problem-solving task. *Proc. Biol. Sci.* 286, 20190408
100. Phillips, J. *et al.* (2020) Knowledge before belief. *Behav. Brain Sci.* Published online September 08, 2020. <https://doi.org/10.1017/S0140525X20000618>
101. Over, H. and Carpenter, M. (2013) The social side of imitation. *Child Dev. Perspect.* 7, 6–11
102. Powell, L.J. (2021) Adopted utility calculus: origins of a concept of social affiliation. *PsyArXiv* Published online April 06, 2021. <http://dx.doi.org/10.31234/osf.io/kuwgf>
103. Warneken, F. and Tomasello, M. (2009) Varieties of altruism in children and chimpanzees. *Trends Cogn. Sci.* 13, 397–402
104. Cosmides, L. *et al.* (2010) Colloquium paper: adaptive specializations, social exchange, and the evolution of human intelligence. *Proc. Natl. Acad. Sci.* 107, 9007–9014
105. Bridgers, S. and Gweon, H. (2018) Means-inference as a source of variability in early helping. *Front. Psychol.* 9, 17071–17077
106. Blaisdell, A.P. *et al.* (2006) Causal reasoning in rats. *Science* 311, 1020–1022
107. Martin, A. and Santos, L.R. (2016) What cognitive representations support primate theory of mind? *Trends Cogn. Sci.* 20, 375–382