Inferential Social Learning: How humans learn from others and help others learn

Hyowon Gweon

Department of Psychology, Stanford University
Stanford, CA 94305

Keywords: Social learning, pedagogy, communication, Bayesian inference, Theory of Mind, Naïve Utility Calculus

Abstract

A widespread view of social learning is that humans, especially children, learn by observing, copying, and trusting others. This learner-centric perspective, however, fails to capture what is distinctive about human social learning: Even young children draw rich inferences from others' behaviors, and communicate information to help others learn. Recent work suggests that these early-emerging abilities as learners and as teachers have common cognitive roots: Domain-general probabilistic inferences guided by an intuitive understanding of how people think, plan, and act. Rather than studying social learning and teaching as two distinct capacities, inferential social learning explains these processes as interpretation and generation of evidence in social contexts, painting an integrated picture of how human cognition supports acquisition and communication of abstract knowledge. (120 words)

What makes human social learning so powerful 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 to promote social learning at varying scales, from parenting practices to educational systems. What makes human social learning so powerful, smart, and distinctive, and what cognitive capacities underlie this feat?

While social learning---learning from others---has long been a topic of research in cognitive science and beyond, its focus has largely remained on "what learners do": What are the behaviors that humans (especially young children) exhibit that make them successful social learners? This line of inquiry has offered several potential answers. Even infants show sensitivity to cues that signal adults' pedagogical intent (e.g., pointing, gaze, infant-directed speech), and in the presence of such cues, they interpret information from them as generalizable knowledge (natural pedagogy [6]); young children imitate others so faithfully that they even copy actions that are causally unnecessary (overimitation [7]); children readily trust what they are told, unless the speaker had been blatantly wrong in the past (epistemic trust [8]). To explain how young children know when to copy, whom to trust, and what to generalize, existing theories have often appealed to associative processes [9] or specialized mechanisms that evolved specifically for the purpose of social learning [Box 1].

Such learner-centric perspective, however, is missing what is truly special about human social learning: We learn from those who help us learn, and we also help others learn. By focusing on the learners without regard for those who enable, facilitate, and promote others' learning, existing research has often underappreciated the bidirectional nature of human social learning. Instead, social learning has often been characterized as one-way transmission of information, reduced to a collection of transmission biases, heuristics, or cue-sensitivities, and studied separately from the mechanisms that underlie how humans communicate, inform, and teach. This tradition has also shaped the way we think about how young children learn from others; in stark contrast to the image of "children as scientists" who are curious, driven to explore, and motivated to learn [10,11], the image of children as social learners has been more passive, credulous, and deferential [12].

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 [13,14], giving rise to a form of learning and communication that is powerful, smart, and distinctively human.

Human social learning: Acquisition and communication of abstract knowledge

To understand what it means to learn from others, it is useful to first consider the nature of what we 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 [15–17] 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 [18], presenting a serious challenge for accounts that emphasize selective copying as the main vehicle of cultural transmission. They also raise questions about how social learning enables acquisition of abstract knowledge well before children can benefit from verbal instruction, or in environments where adult-to-child teaching is rare [19].

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 [20]. Consistent with this idea, developmental research has identified early-emerging signatures of these representations and inferential processes; infants exhibit an abstract, theory-like knowledge in core domains---e.g., objects [21], agents [22,23], number [24], and geometry [25]---that allows them to predict future events [26], explore events that violate these predictions [27], perform actions to learn from self-generated evidence [28,29], and even infer the cause of their failures when such actions fail [30]. This literature has thus painted a picture of young children as curious, active learners; just like scientists, children build, test, and revise their (intuitive) theories of the how the world works by actively acting on the world and drawing inferences from self-generated evidence [10,11].

Learning, however, 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 as fundamentally different ways of learning, young children's "exploration" often unfolds in the presence of (or are even scaffolded by) their caregivers, siblings, and peers. While prior developmental studies on children's causal inferences have rarely acknowledged the role of experimenters who present the evidence to young participants, many of them, in fact, can be reconstrued as studies on how children learn from evidence provided by adults.

This insight has deep implications for what it means to have a comprehensive theory of human learning. Rather than focusing on how humans learn as individual learners, 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, students) and facilitating others' learning (as onlookers, communicators, demonstrators, 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 like how 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., sampling process [31], see Box 2, Figure 1A).

People’s goal-directed behaviors are caused by, and thus reflect, their abstract, structured, intuitive theories about the world. While naïve learners often produce noisy data due to their lack of prior knowledge and expertise, knowledgeable agents, who already have an intuitive understanding of how things work, can perform more targeted causal interventions that reflect their knowledge about the world. 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 learning what resides in others’ minds; to acquire the abstract, structured knowledge that led others to do what they did, learners must use the observed evidence—utterances, actions, demonstrations---as information to infer the contents of others’ mental states that gave rise to those behaviors.

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 unobservable states of an agent give rise to its observable behaviors. Critically, these internal states not only incorporate the agent’s subjective **mental states** (i.e., goals, beliefs, desires) but also its **utilities** (i.e., costs and rewards). While decades of developmental research has 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 people 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 rewards (e.g., 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, [23], see also [22,32–34]). There is growing evidence that such an integrated model of other minds is already present early in life [35,36], supporting both forward-inference (i.e., action-prediction based on known states; forward-inference) and inverse-inference (i.e., using observed actions to infer hidden states). The inverse-inference, in particular, allows learners to use others’ behaviors to infer their knowledge and beliefs that generated those behaviors.

In sum, prior work provides evidence for the foundational capacities that support inferential social learning. Compared to accounts that rely on built-in interpretive biases [6,37], inferential social learning gives more credit to the learners’ social-cognitive capacities while also explicitly taking into account the role of actors (especially as communicators or teachers) into the process of learning. 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 I: Interpreting the meaning of evidence generated by others

Consider someone pulling out a tube from a complex-looking toy (e.g., see 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---a generative model of how others’ minds give rise to their behaviors---can consider the actor's 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 doesn't know about 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), on the other hand, does suggest that the actor has a particular goal and knows how to achieve it. However, the observed evidence still does not provide further information about the toy; 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. 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 [38], 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; beyond 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 it had additional functions, the helpful teacher would've demonstrated them, too).

These intuitions can be formalized as a set of probabilistic inferences that mutually constrain one another, where both the actor (teacher) and the observer (learner) are generating and interpreting evidence based on a generative model of the other [38–40]. Consistent with their 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 [41,42], see Box 2.

Children as Learners II: 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 be a hazard. Consider the toy example again, where a teacher pedagogically demonstrates one function of a toy (Figure 1B, 1C) when it actually has a few other interesting functions. The teacher's demonstration, while true of the toy, is actually misleading; the learner might (reasonably) go beyond the evidence to (inaccurately) infer that it is the toy's only function. Note that the teacher never provided false information; however, 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.

Naïve learners, unfortunately, have no choice but to be misled; they lack the relevant knowledge to even detect the omission itself. Those with prior knowledge, however (e.g., a learner who already played with the toy and discovered other functions), can not only recognize and evaluate the omission as a sin, but also shield themselves from under-informative pedagogy. For instance, when a teacher demonstrates a single function of a toy to a naïve learner, children rate the teacher as less helpful when they already know the toy has additional functions than when they know it is the toy's only function [43,44], 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) [43].

Recent computational work has formalized such evaluation as an inference about the quality of teacher's pedagogical sampling, operationalized as the utility of evidence. This model considers the amount of evidence sampled by a teacher (e.g., how many functions did the teacher show?) and the epistemic value of the evidence (e.g., is the demonstrated function high or low in value?) given the teacher's prior knowledge. While preschool-aged children are sensitive to sins of omission, the ability to exonerate "innocent" omissions (e.g., omission due to ignorance) may continue to develop throughout early childhood [45].

Prior work on teacher evaluation (e.g., epistemic trust) has focused on whether children selectively avoid teachers who provide inaccurate information (e.g., calling a `horse' a `pen', see [8,12,46] for reviews). Recent work with adults, however, suggests that a simple strategy that tracks others' past accuracy is insufficient to explain how humans use social information; rather, people use others' advice or instruction as a way to infer the content that resides in their minds [47]. Children's sensitivity to sins of omission is consistent with this idea. Rather than using explicit cues to determine whom to approach or avoid, humans consider the process by which the teacher generated evidence [31] and the utility of sampled evidence [45] to evaluate the quality of others' teaching. These evaluations reflect an abstract notion of informativeness that incorporates others' wants, needs, and knowledge. Critically, these evaluations are not constrained to pedagogical interactions; they also manifest in linguistic communication, suggesting common cognitive foundations across domains that traditionally been studied separately (see [48,49], Box 3).

Children as Teachers I: Generating useful evidence by thinking about the learner

For human knowledge to accumulate and develop over time, learners must eventually become teachers themselves. Children's abilities as learners---especially their evaluation of others' informativeness---raises 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. Although research on early social learning and teaching have largely remained as separate literatures, 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 3 of which play music (Figure 2A). If you had to teach someone how the toy works, how many buttons would you press? While you might be tempted to simply show the three working buttons, for learners who know nothing about the toy, it is important to demonstrate all 20 buttons; given that all buttons look identical, the learner might (reasonably) generalize their function to the undemonstrated buttons and (inaccurately) infer that all buttons work play music. Such selective evidence, however, might actually suffice for learners who had already seen similar toys; they would resist generalizing the evidence and accurately infer that only those buttons play music. Consistent with this intuition, by late preschool years, children understand that what counts as “just the right amount of evidence” depends on the learner’s prior knowledge. Critically, this understanding manifests not only in their evaluation of teachers but also in how they generate evidence as teachers themselves; children take the time and effort to demonstrate all 20 buttons for a learner who had never seen the toys before, but press just 3 buttons for a learner who has already seen similar toys [50].

Children can also generate the right set of evidence based on the learner’s goal and competence [51]. In this study, 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. When they were asked to play with the toy, 4- to 6-year-old children generated lots of self-generated evidence and learned which variable controls the effect (the color of blocks determine which light turns on, regardless of the mat color). When they were asked to teach a naïve learner, however, the evidence they generated differed depending on what the learner wants (i.e., the learner’s goal) and needs (i.e., the learner’s competence). Children were more likely to generate costly (yet causally informative) evidence that fully deconfounds the two variables selectively when the learner wanted to learn how it works (but not when the learner simply wanted to observe the effect), and when the learner appeared to be incompetent and needed help (but not when the learner was an exceptionally competent learner).

In light of prior work on children’s sampling of evidence as teachers [52,53], these results are worth noting for two reasons. First, consistent with the idea that successful teaching requires Theory of Mind [54–56], children in these studies considered the learner’s goals and prior knowledge to choose the best set of evidence for the learner. Second, these results also suggest that children, as teachers, considering more than the learner’s mental states (goals, beliefs, desires). Children in these studies not only provided “enough” evidence for learners to draw accurate inference, but also resisted providing more than what is required, such as pressing just 3 buttons for learners with prior experience [50], suggesting that they were also considering the “cost” of communication (e.g., effort, time). Just as children’s evaluation of teachers reveals their expectations about others’ informativeness in cooperative communication (i.e., Gricean Maxims, see [48,49], Box 3), children’s generation of evidence as teachers reveals their willingness to be cooperative communicators themselves.

Children as Teachers II: Providing information that maximizes *others’* utilities

Children’s resistance to superfluous demonstrations or over-informing is consistent with the idea that children are sensitive to the utility of demonstrations. Nonetheless, studies mentioned above leave open a critical question: *whose* utility did children consider? Were children trying to minimize their own effort and time, or the learner’s? A recent study combined computational and developmental approaches to provide more direct evidence for the hypothesis that children, as teachers, consider (and maximize) others’ expected utilities [57].

In the studies described above, it was already clear what had to be taught; given the goal to teach how a particular toy works, children had to decide what, and how much, evidence to generate for the learner. Yet, good teaching is just as much, if not more, about deciding “what” to teach in the first place. Prior studies on children’s teaching have often sidestepped this decision by providing a clear target concept to be taught, focusing instead on whether children sample evidence appropriately [50,51], show adult-like teaching practices (e.g., [58,59], see [60] for a review) or how their concept of teaching develops [61]. Our decisions about what to teach, however, is critical to explaining the power of human social learning. Given that teaching requires time, effort, and resources [60,62], teaching can be far more effective if teachers can prioritize the most important information that must be acquired through instruction, and leave the rest for learners to discover on their own.

The key insight here is that such prioritization can be formalized as decisions that maximize the learner's expected utilities. Learners, who still have much to learn, are often ignorant of the cost-reward structure of the world, making them susceptible to ineffective exploration that are costly yet minimally rewarding. Knowledgeable individuals, however, already understand what is valuable or costly to learn; by prioritizing teaching things that increase the learner's rewards or reduce their costs of learning, 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).

The study (Figure 2B) presented children with a choice to teach one of two toys to a naïve learner. What varied across conditions was the relative difference in the toys' expected cost of discovery (i.e., difficult or easy to learn from exploration) and expected reward from activating their functions (i.e., the toy generates a high (enjoyable) or low (dull) effect). Children's responses could not be explained by a simple preference for exciting (high-reward) or complicated (high-cost) toys, or an attempt to maximize their own utilities. 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, consider others' mental states (e.g., teaching what others don't know); they suggest that children, before entering formal schooling as learners, already understand the consequence of others' ignorance on their expected utilities (e.g., ignorance about a high-cost toy leads to high exploration costs with low chances of reaping its reward) and make choices as teachers that maximize others' utilities.

Conceptualizing these decisions as a prosocial utility-maximization can help explain both the widespread presence of pedagogical practices and the cross-cultural differences in the degree to which societies rely on these practices. Cultures 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---often very high in discovery costs (e.g., reading or counting system 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 [63,64] may also reflect the distinctively human (yet culturally universal) abilities to select, compile, and curate useful knowledge in ways that balance our collective utilities.

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

The studies reviewed here provide initial empirical support for inferential social learning; humans learn from evidence generated by others, and generate useful evidence to help others learn. Both processes are inherently communicative and cooperative in nature, supported by cognitive capacities that emerge early in life: domain-general inferential abilities 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 the curious, active, scientist-like image of young children not only as independent explorers but also as social learners. Beyond the focus on well-studied sources of social information such as speech and action-based demonstrations, this perspective calls for studying how children learn from other kinds of social information, such as emotional expressions or evaluative feedback from others (e.g., praise).

Second, it explains social learning and teaching within the same integrated framework that characterizes how humans learn, communicate, and teach in social contexts. While children may generate relatively noisier and less structured evidence as learners, they can also seamlessly transition from acting-to-learn to acting-to-inform, using their knowledge to generate structured evidence that are tailored to maximize others' learning [50,51]. Of course, children's ability to hold, process, remember information, and their mastery of words and action to articulate their knowledge, continue to develop throughout childhood; thus their apparent failures—both as learners and 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, of other minds, and of the ways in which people learn, teach, and communicate.

Finally, it also urges us to revisit the idea that teaching is a capacity that is fundamentally different from learning, or the emphasis on mental-state reasoning as the key to human teaching [56,61,55,65]. Children's behaviors and decisions reflect a rich, intuitive model of other minds that critically incorporates others' utilities as well as mental states, supporting effective, flexible, and "smart" inferences as learners and as teachers. Indeed, these capacities also manifest in other domains, such as inferring the difficulty of novel tasks for others and themselves [66].

Concluding remarks

What makes human social learning so distinctively powerful and smart? Inferential social learning provides an integrated framework for understanding how we learn, teach, and communicate, and how these processes are intimately intertwined in everyday social interactions. Rather than a list of "social learning strategies" [67] or "transmission biases" [6], these interactions are a natural consequence of our inferential and representational capacities for assessing the value, meaning, and interpretability of evidence. These inferences, evaluations, and decision-making processes are hard at work regardless of whether we're on our own or interacting with others.

Humans are driven to learn, improve, and connect with others [68]; even as young children, they 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.

Box 1. Copying what others do, trusting what others say: How and *Why?*

A large body of literature on social learning has focused on imitation and selective trust as major mechanisms for cultural transmission [9,69–72]. An underlying assumption in this body of work is that copying provides an easier, cheaper alternative to individual ("asocial") learning. Given the prevalence of copying and mimicry in nonhuman animals (often characterized as transmission biases [69,73] or social learning strategies [74–76]) and its apparent benefit of reducing the costs of acquiring adaptive behaviors, it may seem reasonable that human social learning also has a similar evolutionary root; perhaps we are genetically predisposed to copy what others do and trust what others say, relying on specialized mechanisms that detect particular cues (e.g., size, age, or social status) that "trigger" or "elicit" these behaviors.

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 [7] or selectively imitate intentional actions [77]; they trust anyone who points [78] or trust those who were reliable in the past [8]; infants interpret object properties as generalizable given pedagogical cues [37] but they also consider the probability of sampled evidence [31]. How do we make sense of these findings?

From the perspective of inferential social learning, these results all reflect reasonable responses to observed evidence given the context. The question at stake is not whether children copy or trust (of course they do) but *how and why* children engage in these epistemic practices, and what cognitive capacities support their abilities as learners and teachers. Rather than shifting the explanatory burden to specialized mechanisms that trigger social learning with an added layer of late-developing cognition (e.g., language, theory of mind), or hypothesizing a particular kind of cognition specifically for teaching [65], inferential social learning makes explicit assumptions about what cognitive capacities give rise to these epistemic practices, and explains *how* by using computational models that formalize the underlying representations and inferential processes. *Why* do children imitate, trust, and generalize? It's not because they have cue-detecting mechanisms that elicit these responses, but because they interpret observed actions to recover their underlying meaning and actively decide whether to re-enact, trust, or generalize; after all, imitation is yet another way to generate useful evidence based on our observations of others. These abilities support far more than social learning and teaching, shaping children's everyday behaviors from self-guided exploration to cooperation and even deception.

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 vs. seek others' help depending on the likely cause [30].

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 [31]. For instance, when an agent draws a sample of 3 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 (vs. 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, Figure 1A, main text). In the absence of demonstrated properties, strong sampling can indicate the agent's intrinsic preference for the exemplars [79,80].

Learning from pedagogically sampled evidence: Seeing a demonstration of a toy's property also raises a question about the toy itself: does the toy have other properties, or is this it? Beyond flexible inductive generalization [31], children's sensitivity to sampling process also supports causal inferences that go beyond the evidence (Figure 1B, main text). Given a demonstration of a toy's squeaking property, children are more likely to infer it is the toy's *only* property and explore less when the action suggests pedagogical sampling (instructional) than when it does not (e.g., instrumental or incidental actions) [41,42]. These results have been replicated in Yucatec Mayan culture where children rarely receive direct instruction from adults [42], 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.

Box 3. Cooperative communication via actions and words.

Mutual, recursive reasoning about others' minds—key assumption in computational models of pedagogical reasoning [39,40]---leads to the expectation that evidence provided by a teacher should be true, relevant, sufficient for accurate inference, yet not superfluous given the learner's goal and knowledge. It is not a coincidence that these are the key properties of Gricean Maxims and cooperative communication more generally [48,81], also formalized as recursive mental-state reasoning between a speaker and a listener (Rational Speech Acts, [49]).

Both models—RSA and pedagogy---extend probabilistic models of learning by incorporating psychological reasoning as a generative process for sampling of evidence, one as goal-directed actions [39] and the other as linguistic utterances [49]. Together they provide a productive framework for formalizing learning and communication across a range of contexts and modalities, suggesting that similar inferential processes underlie interpretation and generation of evidence in teacher-learner or speaker-listener interactions [39,49,47,82]) as well as evaluation of teachers' or speakers' over- or under-informativeness [45,50,83]. These theoretical parallels also manifest in empirical results, where children show similar developmental trajectories in their inferences and evaluations in linguistic communication and pedagogical interactions.

Inference: By age 3, children draw ad-hoc implicature (i.e., implicature based on a scale constructed in context [84], expecting speakers to generate utterances in a way that helps the listener infer their intended, pragmatic meaning (e.g., "my friend has glasses" means the friend is not wearing other things, Figure 3); this parallels children's inferences in pedagogical contexts (Figure 1B, "the toy does not have other functions"), both of which emerges from an expectation about helpful communication [41,42].

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, Figure 3), but unlike older children, they succeed only when the appropriate scalar alternatives are made clear in context [85,86]. 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 [44], both of which may emerge from understanding what a helpful communicator would have said or demonstrated.

Language (speech acts) and action-based demonstrations have their distinct advantages in communication; a few words (e.g., "all buttons work") can replace costly demonstrations, but sometimes seeing a demonstration is worth a thousand words. For all their differences, however, the way we interpret their meanings may involve similar representations and inferential processes [87]. From this perspective, what distinguishes teaching (pedagogy) from other communicative behaviors is not whether it involves actions vs. words, but the strength of mutual expectations about the process of evidence-generation.

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

Many nonhuman species show basic inferential capacities [88,89], and primates in particular have a surprisingly rich understanding of others' mental states [90,91]; rather than relying on reflexes that result in teaching-like behaviors (e.g., meerkats, [92]), they imitate human actions selectively [93,94], coordinate with their conspecifics towards a common goal [95], and even engage in basic tool-sharing behaviors [2]. What, then, makes humans so distinctive in their behaviors as learners and teachers?

The ability to draw inferences from others' behaviors---the key driver of inferential social learning---may be present in other species. However, non-human species may be limited in their underlying representational and inferential capacities [96], which in turn limits their ability to represent and reason about abstract, structured contents of cultural knowledge, or capitalize on the intentional structure of pedagogy. Evaluation and generation of evidence, in particular, require recursive reasoning about how one's communicative behaviors impacts another's epistemic state given their mutual expectations about the context.

Another answer may lie in our motivation to learn from others and help others learn. The findings described in this paper wouldn't exist without children's desire to connect, communicate, and cooperate with others [68], yet even simple helping behaviors are difficult to find in nonhuman species {Citation}. While inferential social learning does not make explicit commitments about the origins of human prosociality or altruism, it invites motivation as a part of the picture by explaining how cognitive capacities can generate context-specific motivations to help, communicate, and teach (i.e., to *want* to remedy others' false belief, we must first represent it as such). After all, the most remarkable aspect of human intelligence is not its power alone but our willingness to use it on behalf of others [31].

More generally, to explain what makes humans special, existing proposals have emphasized either the individuals' cognitive abilities (i.e., "big brains") as drivers of individual discoveries [70,72,98]) or the inherited cultural knowledge [9,69,71] that obviates the need for individual intelligence. In inferential social learning, both are necessary: "Big brain" supports powerful inferences from observed evidence, and "culture" is what resides in people's minds to generate evidence they provide for others. Through each individual making their own decisions about how best to acquire, revise, and communicate what they know, we, as a group, can drive accumulation of abstract cultural knowledge that evolves over time. By understanding what makes these decisions and inferences so powerful and distinctive, we can also better understand the distinctively human nature of our cultural knowledge.

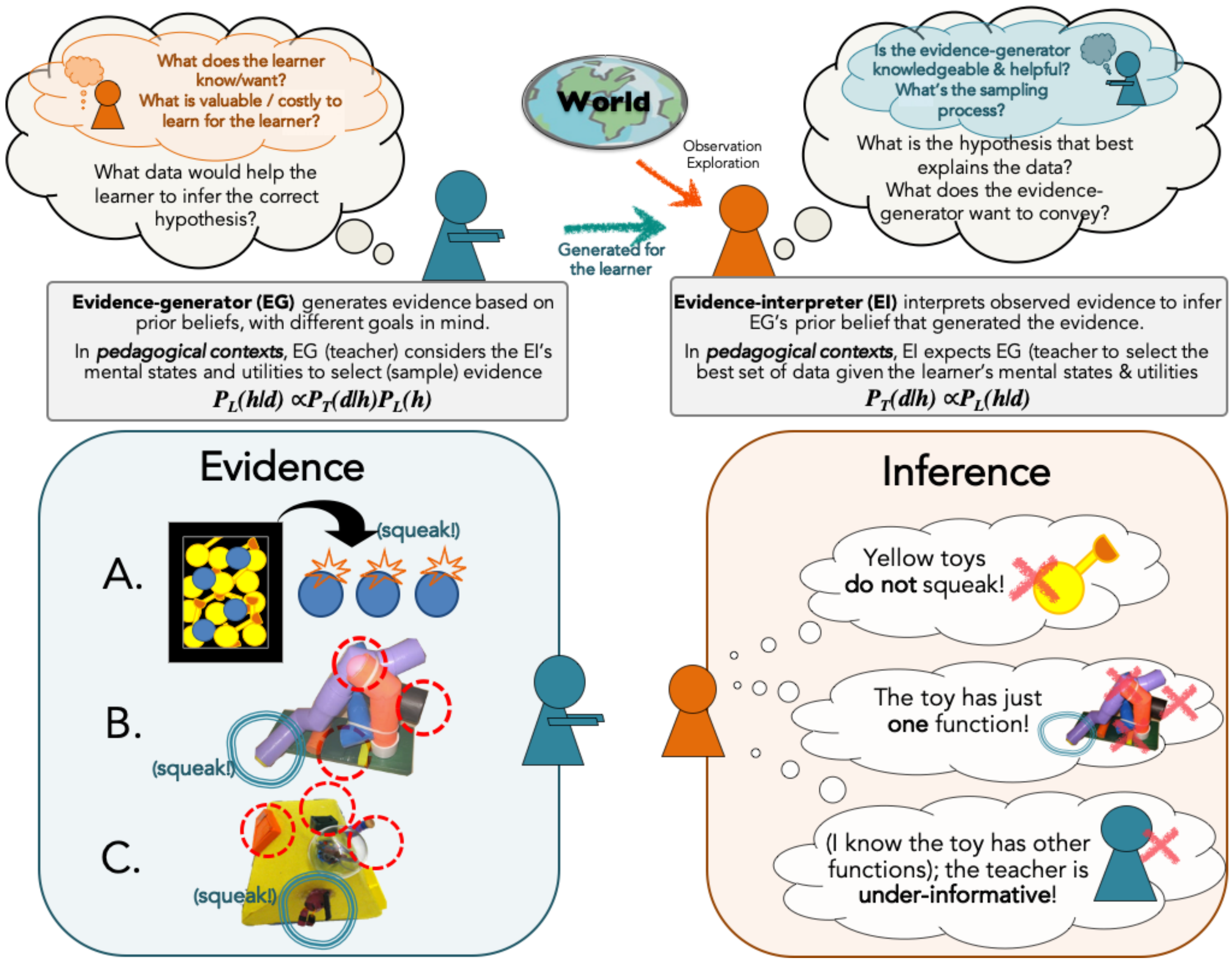

Figure 1, Key Figure. Human learners (EI, orange) draw inferences from observed evidence either from their own exploration or from others' behaviors. Such learning can be particularly powerful when the evidence is generated by knowledgeable, helpful agents (EG) whose goal is to communicate their knowledge (i.e., teacher, green). In inferential social learning, the learner and the teacher work together towards a common epistemic goal (learner's acquisition of desired knowledge) by drawing cooperative inferences about others' mental states (e.g., goals, beliefs, desires) and utilities (costs and rewards of goal-directed actions). The lower panel shows learner's inferences given demonstrated evidence. A: Given three blue toys that squeak, drawn from a box with mostly yellow toys, suggests strong sampling (indicating that the evidence-generator knew that only blue toys squeak); even infants are sensitive to the sampling process and decide whether the squeaking property generalizes to yellow toys based on the probability of the sample [31]. B: A tube-pulling action on a complex-looking novel toy in a pedagogical (instructional) context suggests the toy does not have additional functions. Toddlers and preschoolers consider the context in which the action was performed (incidental, instrumental, or instructional) and explore toys differently depending on whether the observed evidence indicates the presence of other functions [41,42]. C: When the learner already knows about the toy's additional functions, the learner can recognize the omission of evidence as "a sin" and evaluate the quality of evidence accordingly. Children as young as four evaluate a teacher as less helpful when they stop short of demonstrating all functions of a toy [43,44].

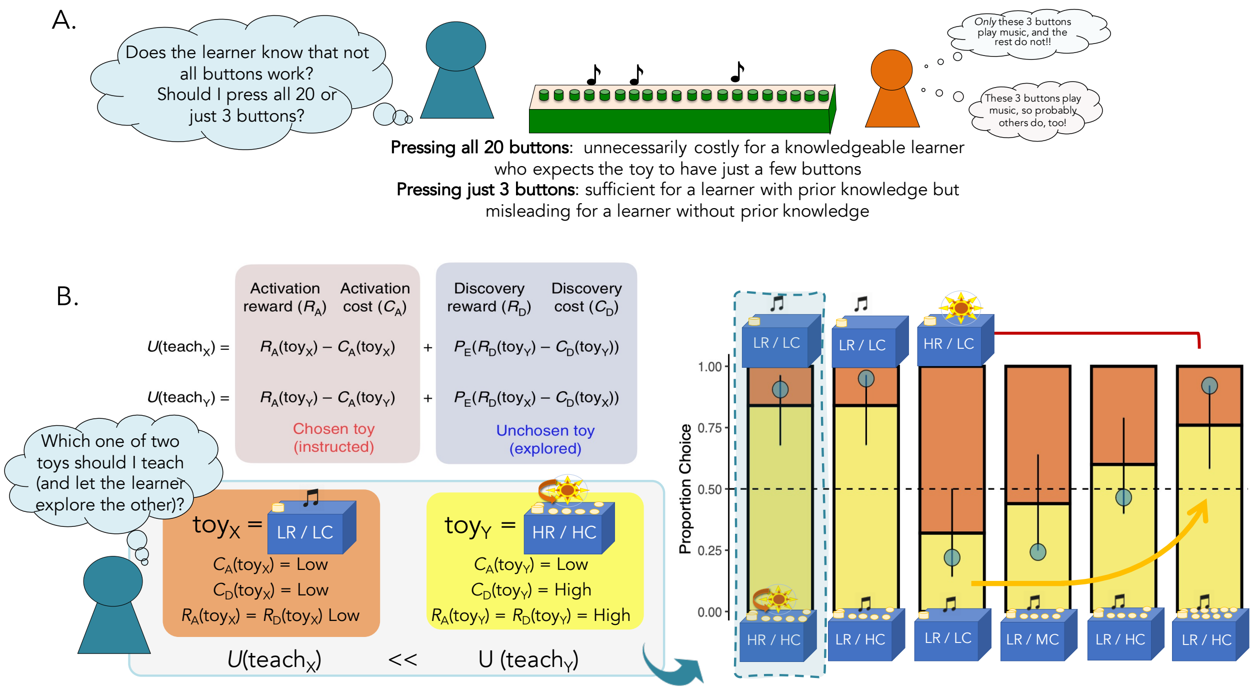

Figure 2. A: While pressing all 20 buttons ensures the learner has exhaustive evidence, but it is costly and effortful; pressing just the 3 working buttons is efficient and effective when the learner has prior knowledge (top thought bubble) but misleading for learners who lack prior knowledge (bottom thought bubble). Children consider what the learner knows to evaluate teachers who press 20 or 3 buttons, and even press 20 or 3 buttons as teachers themselves [50]. B: Given a choice to teach one of two toys (and let the learner explore the other one), a teacher can 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. In this example [57], the teacher is choosing between a low-reward (LR, music) low-cost (LC, single button) toy and a high-reward (HR, 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 6 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).

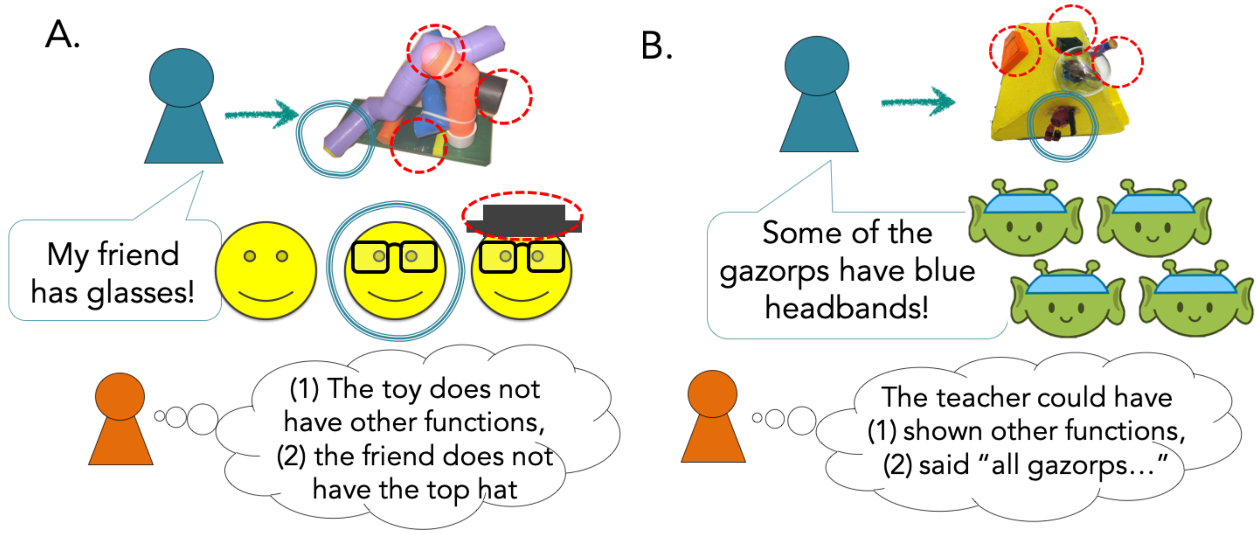

Figure 3 (to appear in Box 3). Parallels between pragmatic inferences and evaluation in pedagogical contexts and linguistic communication. In A, when a speaker says "my friend has glasses" given three friends (one with neither glasses nor hat, one with glasses only, and one with hat and glasses), children infer that the friend does not have a hat, and chooses the second one even though the third one also has glasses [84]. Just like inferring that a toy does not have additional functions [41,42], this inference requires the understanding that a helpful agent would have shown additional functions or mentioned the hat to identify the friend. In B, even preschool-aged children reject the utterance "Some of the gazorps have blue headbands" as infelicitous but only when they have seen accurate and relevant uses of the word "some" (e.g., when the property applies to 3 of 4 gazorps). Just like evaluating sins of omission, this requires the ability to understand that the teacher/speaker could have shown additional functions (but did not) and could have used the word "some" (but did not).

Highlights

- Human knowledge is abstract, structured, theory-like, and reflects our intuitive theories of how the world works.
- To understand how humans acquire and communicate abstract knowledge, we must study social learning (how we learn from others) and teaching (how we help others learn) through the same theoretical lens.
- Human social learning and teaching are best described as rich, abstract inferences grounded in basic cognition, empowered by representations of others' minds (e.g., mental states and utilities).
- Learners recover the abstract contents of others' minds by inverting an intuitive, causal model of how people think, plan, and act; these inferences are particularly powerful when the evidence is generated those who intend to teach; the learner interprets and evaluates evidence based on a model of the teacher's mind, and the teacher generates evidence based on a model of the learner's mind.
- Young children's abilities as learners and teachers suggest that children are far more than passive recipients of social information; they actively seek, interpret, and evaluate information from others, and generate useful evidence to help others learn.

Outstanding Questions

- Inferential social learning has been applied to explain how humans—especially young children---learn about the external world (e.g., objects, causal mechanisms, other people); can it also explain how humans learn about the *inner* world (i.e., the self)? How do children learn from others' feedback about the self (e.g., praise, criticisms) and how do they communicate to others about the self?
- Existing literature has focused primarily on the role of language and action-based demonstrations as information for social learning. Do humans (especially children) use others' emotional (e.g., facial, vocal, or even symbolic (emoji)) expressions as information for learning? Do they generate emotional expressions to help others learn (e.g., emotionese)?
- Humans not only engage in cooperative interactions towards their epistemic goals, but also leverage social information to maximize their gains (as learners),or withhold information to outcompete others (as communicators). How can we extend the inferential social learning framework to explain social decisions and interactions in competitive contexts?
- Inferential social learning suggests that a combination of existing cognitive capacities can explain how we learn, communicate, and teach. In the past couple of decades, we have learned a lot about the neural mechanisms that support social perception, mental-state reasoning, and utility-based decision-making; how can we integrate advances in cognitive neuroscience within the framework of inferential social learning to better understand how the brain supports cooperative communication as learners and as teachers?
- Inferential social learning suggests that copying and imitation does not successfully solve the problem of acquiring abstract knowledge. Yet, most machine learning/AI algorithms for social learning relies on these strategies. 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?

References

1 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). *Scientific reports* DOI: 10.1038/srep35953

2 Musgrave, S. *et al.* (2016) Tool transfers are a form of teaching among chimpanzees. *Sci Rep* 6, 34783

3 Cornell, H.N. *et al.* (2011) Social learning spreads knowledge about dangerous humans among American crows. *Proceedings of the Royal Society B: Biological Sciences* 279, 499–508

4 Duffy, G.A. *et al.* (2009) Size-dependent directed social learning in nine-spined sticklebacks. *Animal Behaviour* 78, 371–375

5 Loukola, O.J. *et al.* (2017) Bumblebees show cognitive flexibility by improving on an observed complex behavior. *Science* 355, 833–836

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

7 Lyons, D.E. *et al.* (2007) The hidden structure of overimitation. *Proceedings of the National Academy of Sciences* 104, 19751

8 Harris, P.L. *et al.* (2018) Cognitive Foundations of Learning from Testimony. *Annu. Rev. Psychol.* 69, 251–273

9 Heyes, C. (2018) *Cognitive gadgets: the cultural evolution of thinking*, Harvard University Press.

10 Gopnik, A. *et al.* (1999) *The scientist in the crib: Minds, brains, and how children learn.*, William Morrow & Co.

11 Schulz, L. (2012) The origins of inquiry: inductive inference and exploration in early childhood. *Trends in Cognitive Sciences* 16, 382–389

12 Harris, P.L. (2012) *Trusting what you're told: How children learn from others*, Harvard University Press.

13 Bratman, M.E. (2013) *Shared agency: A planning theory of acting together*, Oxford University Press.

14 Tomasello, M. (2019) *Becoming human: A theory of ontogeny*, Belknap Press.

15 Carey, S. (2011) *The origin of concepts*, Oxford University Press.

16 Keil, F.C. (1989) Concepts, kinds, and conceptual development.

17 Wellman, H.M. and Gelman, S.A. (1992) Cognitive development: foundational theories of core domains. *Annual Review of Psychology* 43, 337–375

18 Sperber, D. (1996) Explaining culture: A naturalistic approach. *Cambridge, MA: Cambridge*

19 Rogoff, B. *et al.* (2003) Firsthand Learning Through Intent Participation. *Annual Review of Psychology* 54, 175–203

20 Tenenbaum, J.B. *et al.* (2011) How to Grow a Mind: Statistics, Structure, and Abstraction. *Science* 331, 1279–1285

21 Spelke, E. *et al.* (1992) Origins of knowledge. *Psychol Rev* 99, 605–632

22 Gergely, G. and Csibra, G. (2003) Teleological reasoning in infancy: the naive theory of rational action. *Trends in Cognitive Sciences* 7, 287–292

23 Jara-Ettinger, J. *et al.* (2016) The Naïve Utility Calculus: Computational Principles Underlying Commonsense Psychology. *Trends in Cognitive Sciences* 20, 589–604

24 Feigenson, L. *et al.* (2004) Core systems of number. *Trends in Cognitive Sciences* 8, 307–314

25 Dillon, M.R. *et al.* (2013) Core foundations of abstract geometry. *Proceedings of the National Academy of Sciences of the United States of America* 110, 14191–14195

26 Téglás, E. *et al.* (2011) Pure reasoning in 12-month-old infants as probabilistic inference. *Science* 332, 1054–1059
27 Stahl, A.E. and Feigenson, L. (2015) Observing the unexpected enhances infants' learning and exploration. *Science* 348, 91–94
28 Cook, C. *et al.* (2011) Where science starts: spontaneous experiments in preschoolers' exploratory play. *Cognition* 120, 341–349
29 Sim, Z.L. and Xu, F. (2017) Learning higher-order generalizations through free play: Evidence from 2- and 3-year-old children. *Developmental Psychology* 53, 642–651
30 Gweon, H. and Schulz, L. (2011) 16-Month-Olds Rationally Infer Causes of Failed Actions. *Science* 332, 1524–1524
31 Gweon, H. *et al.* (2010) Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences* 107, 9066–9071
32 Lucas, C.G. *et al.* (2014) The Child as Econometrician: A Rational Model of Preference Understanding in Children. *PLoS ONE* 9, e92160
33 Baker, C.L. *et al.* (2017) Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Publishing Group* 1, 1–10
34 Jern, A. *et al.* (2017) People learn other people's preferences through inverse decision-making. *Cognition* 168, 46–64
35 Liu, S. *et al.* (2017) Ten-month-old infants infer the value of goals from the costs of actions. *Science* 358, 1038–1041
36 Jara-Ettinger, J. *et al.* (2015) Children's understanding of the costs and rewards underlying rational action. *Cognition* 140, 14–23
37 Csibra, G. and Shamsudheen, R. (2015) Nonverbal Generics: Human Infants Interpret Objects as Symbols of Object Kinds. *Annual Review of Psychology* 66, 689–710
38 Bonawitz, E. and Shafto, P. (2016) ScienceDirect Computational models of development, social influences. *Current Opinion in Behavioral Sciences* 7, 95–100
39 Shafto, P. *et al.* (2014) A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology* 71, 55–89
40 Shafto, P. *et al.* (2012) Learning From Others: The Consequences of Psychological Reasoning for Human Learning. *Perspect Psychol Sci* 7, 341–351
41 Bonawitz, E. *et al.* (2011) The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition* 120, 322–330
42 Shneidman, L. *et al.* (2016) Learning From Others and Spontaneous Exploration: A Cross-Cultural Investigation. *Child Development* 87, 723–735
43 Gweon, H. *et al.* (2014) Sins of omission: Children selectively explore when teachers are under-informative. *Cognition* 132, 335–341
44 Gweon, H. and Asaba, M. (2018) Order Matters: Children's Evaluation of Underinformative Teachers Depends on Context. *Child Dev* 89, e278–e292
45 Bass, I. *et al.* (under review) , The effects of information utility and teachers' knowledge on evaluations of under-informative pedagogy across development.
46 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
47 Vélez, N. and Gweon, H. (2018) Integrating Incomplete Information With Imperfect Advice. *Top Cogn Sci* 329, 1081–17
48 Grice, H.P. (1975) Logic and Conversation. pp. 1–11
49 Goodman, N.D. and Frank, M.C. (2016) Pragmatic Language Interpretation as Probabilistic Inference. *Trends in Cognitive Sciences* 20, 818–829
50 Gweon, H. *et al.* (2018) Development of children's sensitivity to overinformativeness in learning and teaching. *Developmental Psychology* 54, 2113–2125

51 Gweon, H. and Schulz, L. (2019) From Exploration to Instruction: Children Learn From Exploration and Tailor Their Demonstrations to Observers' Goals and Competence. *Child Development* 90, e148–e164
52 Rhodes, M. *et al.* (2010) Children's attention to sample composition in learning, teaching and discovery. *Developmental Science* 13, 421–429
53 Rhodes, M. *et al.* (2015) Controlling the message: preschoolers' use of information to teach and deceive others. *Front Psychol* 6, 867
54 Ashley, J. and Tomasello, M. (1998) Cooperative problem-solving and teaching in preschoolers. *Social Development* 7, 143–163
55 Strauss, S. *et al.* (2002) Teaching as a natural cognition and its relations to preschoolers' developing theory of mind. *Cognitive Development* 17, 1473–1487
56 Bass, I. *et al.* (2019) Children's developing theory of mind and pedagogical evidence selection. *Developmental Psychology* 55, 286–302
57 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
58 Calero, C.I. *et al.* (2015) Young children are natural pedagogues. *Cognitive Development* 35, 65–78
59 Ronfard, S. and Corriveau, K.H. (2016) Teaching and preschoolers' ability to infer knowledge from mistakes. *Journal of Experimental Child Psychology* 150, 87–98
60 Kline, M.A. (2015) How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behav Brain Sci* 38, e31
61 Sobel, D.M. and Letourneau, S.M. (2016) Children's developing knowledge of and reflection about teaching. *Journal of Experimental Child Psychology* 143, 111–122
62 Caro, T.M. and Hauser, M.D. (1992) Teaching in nonhuman animals. *Quarterly Review of Biology* 67, 74
63 Shneidman, L. *et al.* (2015) Child-directed teaching and social learning at 18 months of age: evidence from Yucatec Mayan and US infants. *Developmental Science* 19, 372–381
64 Boyette, A.H. and Hewlett, B.S. (2017) Teaching in Hunter-Gatherers. *Rev.Phil.Psych.* 9, 771–797
65 Strauss, S. *et al.* (2014) Teaching, naturally. *Trends in Neuroscience and Education* 3, 38–43
66 Gweon, H. *et al.* (2017) , Reverse-engineering the process: Adults and preschoolers? ability to infer the difficulty of novel tasks.
67 Laland, K.N. (2004) Social learning strategies. *Learn Behav* 32, 4–14
68 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
69 Henrich, J. (2015) *The secret of our success: how culture is driving human evolution, domesticating our species, and making us smarter*, Princeton University Press.
70 Pinker, S. (2010) Colloquium paper: the cognitive niche: coevolution of intelligence, sociality, and language. *Proceedings of the National Academy of Sciences* 107 Suppl 2, 8993–8999
71 Boyd, R. *et al.* (2011) The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences* 108, 10918–10925
72 Legare, C.H. and Nielsen, M. (2015) Imitation and Innovation: The Dual Engines of Cultural Learning. *Trends in Cognitive Sciences* 19, 688–699
73 Boyd, R. and Richerson, P.J. (2005) *The origin and evolution of cultures*, Oxford University Press.
74 Hoppitt, W. and Laland, K.N. (2013) *Social learning: an introduction to mechanisms, methods, and models*, Princeton University Press.
75 Rendell, L. *et al.* (2011) Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences* 15, 68–76

76 Kendal, R.L. *et al.* (2018) Social Learning Strategies: Bridge-Building between Fields. *Trends in Cognitive Sciences* 22, 651–665

77 Gergely, G. *et al.* (2002) Rational imitation in preverbal infants. *Nature* 415, 755

78 Palmquist, C.M. and Jaswal, V.K. (2012) Preschoolers Expect Pointers (Even Ignorant Ones) to Be Knowledgeable. *Psychological Science* 23, 230–231

79 Kushnir, T. *et al.* (2010) Young Children Use Statistical Sampling to Infer the Preferences of Other People. *Psychological Science* 21, 1134

80 Ma, L. and Xu, F. (2011) Young children's use of statistical sampling evidence to infer the subjectivity of preferences. *Cognition* 120, 403–411

81 Clark, H.H. (1996) *Using language*, Cambridge university press.

82 Hawthorne, D. and Goodman, N.D. (2017) Reasoning about Social Sources to Learn from Actions and Outcomes. *Decision*

83 Degen, J. *et al.* (2020) When redundancy is useful: A Bayesian approach to "overinformative" referring expressions. *Psychological Review* DOI: 10.1037/rev0000186

84 Stiller, A.J. *et al.* (2014) Ad-hoc Implicature in Preschool Children. *Language Learning and Development* DOI: 10.1080/15475441.2014.927328

85 Barner, D. *et al.* (2011) Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition* 118, 84–93

86 Skordos, D. and Papafragou, A. (2016) Children's derivation of scalar implicatures: Alternatives and relevance. *Cognition* 153, 6–18

87 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

88 Taylor, A.H. *et al.* (2012) New Caledonian crows reason about hidden causal agents. *Proceedings of the National Academy of Sciences* DOI: 10.1073/pnas.1208724109

89 Eckert, J. *et al.* (2018) Intuitive statistical inferences in chimpanzees and humans follow Weber's law. *Cognition* 180, 99–107

90 Krupenye, C. *et al.* (2016) Great apes anticipate that other individuals will act according to false beliefs. *Science* 354, 110–114

91 Drayton, L.A. and Santos, L.R. (2017) What do monkeys know about others' knowledge? *Cognition* 170, 201–208

92 Thornton, A. (2006) Teaching in Wild Meerkats. *Science* 313, 227–229

93 Tomasello, M. and Carpenter, M. (2005) Intention reading and imitative learning. *Perspectives on imitation: From neuroscience to social science* 2, 133–148

94 Buttelmann, D. *et al.* (2007) Enculturated chimpanzees imitate rationally. *Developmental Sci* 10, F31–F38

95 Melis, A.P. and Tomasello, M. (2019) Chimpanzees (Pan troglodytes) coordinate by communicating in a collaborative problem-solving task. *Proc. Biol. Sci.* 286, 20190408–10

96 Phillips, J. *et al.* (2020) Knowledge before Belief. *Behav Brain Sci* DOI: 10.1017/S0140525X20000618

97 Bridgers, S. and Gweon, H. (2018) Means-Inference as a Source of Variability in Early Helping. *Front Psychol* 9, 17071–7

98 Cosmides, L. *et al.* (2010) Colloquium Paper: Adaptive specializations, social exchange, and the evolution of human intelligence. *Proceedings of the National Academy of Sciences* 107, 9007–9014