Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Current Opinion in
Behavioral
Sciences

CrossMark

Computational models of development, social influences

Elizabeth Bonawitz and Patrick Shafto

In the article we argue that past Bayesian approaches that model children's learning from data are missing an important element — the role of other people in generating that data. We propose that children take the origin of data into account when learning, which can be understood through ideal observer analyses of the social situation. Moreover, when observing evidence, children are not just learning *from* others, but also *about* others. We review recent literature suggesting that children can make inferences about the knowledge and goals of the individual selecting the data and use this knowledge to bolster learning from this evidence.

Address
Department of Psychology, Rutgers University–Newark, 334 Smith Hall, Newark, NJ 07102, USA

Corresponding author: Bonawitz, Elizabeth (elizabeth.bonawitz@rutgers.edu)

Conventional wisdom often points to the myriad ways in which children have unique limitations and cognitive capacities from adults — children may be lacking the knowledge, skills, and logical reasoning abilities to perform as competently as their parents [1–8,61]. However, children must be powerful learners. They enter the world with few expectations about objects, events, others and themselves, but within a year they begin express complicated concepts like 'all done', 'uh oh', or 'mine'. To explain such remarkable learning, researchers have proposed elements of the learning toolkit including statistical learning mechanisms [9–11] and the ability to integrate these statistics with inductive constraints [12•,13–16,17•,18,60].

One way to understand how even very young children can attend to the evidence in their environment and — from this data — rapidly update their beliefs, is through computational models based on ideal observer analyses. Ideal observer analyses answer the question, what is the learning problem that is being solved? The answer is formalized as a computational model that assumes ideal performance. In cognitive development, the learning problems being solved are things like learning language and learning the causal structure of the environment. The computational models formalize learning as Bayesian inference over structured representations. Bayesian inference captures ideal learning in the face of uncertainty and noise, and structured representations captures the underlying regularities of, for example, the grammar of language or the causal structure of the environment that must be inferred. The result of this marriage is 'Theory-based Bayesian' models of learning [19•].

In recent years, myriad studies have demonstrated how ideal observer approaches can be used to describe how children might learn language [20] and develop causal beliefs [21,22]. Although these models provide important insight into how we can capture the remarkably sophisticated learning in childhood, we suggest that they miss an important aspect of learning: learning from others. Evidence does not appear in a social vacuum; children often learn from observing others. When a word is uttered, it comes from the mouth of a person trying to communicate a concept. When a light switch is flicked, it is almost always because a person wants there to be light.

Computational models of learning have demonstrated the importance of considering social information in language acquisition (e.g. [23]), social referencing (e.g. [24]), and imitation (for review see, [25]). These past approaches use social information as an additional source of evidence, independent from other streams of information. Here, we focus on how social inferences shape the *interpretation* of evidence. The critical contribution of *socially generated data* — that people are part of the generative process — is that it can lead to even stronger inferences and more rapid learning. We will stress is that children are not only paying attention to the evidence (e.g. the light switch in the up position and the light being on), but that children are paying attention to the people that are demonstrating those events (e.g. that even if the person tries, but fails, to flip the light switch up, there was a reason that they tried to flip it). We will discuss recent probabilistic models that include this socially generative component. These models relate the *knowledge* and *goals* of a demonstrator to actions they choose, and formalize how these purposeful actions can influence the interpretation of the data being

demonstrated. We will present research suggesting that children can make inferences about the knowledge and goals of the individual selecting the data and use this knowledge to bolster learning from this evidence.

Intrinsic to our proposal is the idea that children are able to represent and reason about others' beliefs. These are non-trivial abilities, combining Theory of Mind and planning capabilities, and the development of these skills may extend into middle childhood (e.g. [26,27]). However, many of the types of social learning we discuss do not necessarily require competency with false beliefs; many other forms of social learning based on reasoning about others' beliefs and goals can be observed much earlier in development [28–30], and even infants may understand false-beliefs [31].

Strong sampling in word learning

Word learning is arguably the canonical case of social learning. It is also famously wrought with uncertainty [32]: How should one infer the precise extension of a word's meaning given a few observations of a label? When a scene is labeled for a learner, the possible meanings are endless, and yet even children are likely to infer that the speaker intends something like 'rabbit', and not 'rabbits and also tomatoes but not squirrels'. Words are also learned with very few positive examples and rarely any negative examples; words can carve up the world in overlapping categories; and multiple words can be used to express similar concepts [33–37]. How might children accomplish such a feat of inductive prowess? As Quine [32] famously noted, 'Language is a social art'; the learner's reasoning depends on making inferences about the speaker's intensions. Might such intuitions leverage inductive inference?

Theory-based Bayesian models can begin to offer a solution to the riddle of induction and a precise way to characterize the contribution of social inferences in word learning [17•,18]. Imagine learning about a concept, such as 'dalmation' by observing someone randomly labeling objects from the world, and in each case stating whether the drawn object is an instance of the concept or not. Given these examples, the learner must infer which of many possible hypotheses is most likely to account for the data produced by these haphazard observations.

Of course, this is not how people use language. Intuitively, we know that, in general, labeling any given object as 'not a dalmation' provides nearly no information about what 'dalmation' actually is. Instead, we can consider language learning as a different sort of problem, one in which people almost exclusively label objects with their correct label. The examples of 'dalmation' may still be haphazard, but they are all dalmations. Thus, when a learner sees three dalmations drawn from a box and labeled with a novel word, she can reason about its extension via a sort-of counterfactual logic: 'what are the chances that I would have seen exactly these three objects if the concept I'm learning is 'dalmations', 'dogs', or 'animals'?' The narrower the concept fits to the examples, the more likely the set will have been drawn. One is most likely to see three dalmations randomly sampled from a box only containing dalmations, less likely to see such evidence if the box contains all dogs, and least likely if the box contains all animals.

Computationally, researchers have modeled these different potential selection processes. The former, called weak sampling, provides only the information present in the object-label pairing. This would lead to strong limitations on learning because the sampling process — random sampling — is unrelated to the true concept. Indeed, mathematical demonstrations formalize the idea that, if this were true, language would not be learnable [35]. In contrast, strong sampling formalizes the idea that if a person intends to communicate 'dalmation', they will choose dalmations. This does not completely solve the problem: if the person wishes to communicate 'animal' they may also choose a dalmation. However, if as a language learner we aggregate over multiple labelings, then the probability of someone choosing 3 dalmations at random from all animals is very, very small; dalmation must mean 'dalmation' [17•,18].

Learning from teaching

Another especially compelling case for social learning comes from teaching. According to an ideal observer analysis, learning from teaching ('pedagogical learning') differs from standard learning in one critical way — the data that the learner observes are chosen by a knowledgeable and helpful person, not selected at random. That is, in pedagogical sampling the learner assumes that the teacher is not merely sampling examples from within the concept, but is purposefully sampling the best possible data in order to help the learner rapidly converge on the correct concept given the learner's current beliefs. Thus, the learner not only asks herself 'what are the chances I would have observed these examples, given this category' but she asks 'what are the chances that *the teacher would have chosen these examples for me if* this was the true category.' A knowledgeable and helpful teacher would be especially unlikely to provide an overly restrictive set of examples, such as three Dalmations, if the concept to be conveyed was really something broader, such as 'dogs'. Thusly, pedagogical sampling allows the learner make even stronger inferences about the data than strong sampling would warrant [38,39•].

These pedagogical inferences also extend beyond word learning to other kinds of causal reasoning. For example, imagine you are a child, and a person walks into the room with a complex looking novel toy. The person says, 'This is my toy. I'm going to show you how it works' then pulls a

non-obvious yellow knob, eliciting a squeak, before handing the toy to the child. What should a child infer about the number of causal functions of the toy? Contrast this with a situation in which the person says, 'Look at this toy I just found' and, as they are handing the toy to the child, they accidentally pull the knob, eliciting a squeak, before handing the toy over to the child. What should the child infer about the number of causal functions of the toy in this other situation? It seems obvious that, in the second case, the child can infer that while there is evidence for one function, there could be more functions. However, in the first case, by reasoning about why the teacher chose to demonstrate only one function, the child can infer that there is in fact only one function: if there were more functions, a knowledgeable and helpful teacher would have chosen to demonstrate those too. Recent research using this and related paradigms have shown that, before they are arriving at formal schooling, 4- and 5-year-old children can use the knowledge and intent of the teacher to draw stronger inferences than are afforded by the data alone [40•,41]. Indeed, parallel research in education has shown similar implications of evidence selection on children's learning of the mathematical concept of equality [42–46].

Recent computational research has formalized a framework for modeling why other people's purposeful data

selection affects learning [39•]. Consider the case of learning from teaching, where the data are chosen by a knowledgeable and helpful informant ('pedagogical reasoning'; [38,47]). The model captures the deceptively simple intuition that a teacher will choose the evidence that tend to increase the learner's belief in the correct hypothesis; it makes precise the implications of that for learners who expect this sort of purposeful selection of data. Formally, we can describe the probability of a teacher selecting data given a hypothesis, $P_T(d|h)$, and the relationship of the teacher's choice to the learner's posterior beliefs after observing the data, $P_L(h|d)$. To capture the intuition that a teacher will choose data that tend to lead the learner to the correct hypothesis, we state: $P_T(d|h) \propto P_L(h|d)$. Learning from teaching is similar to standard learning based on Bayes rule, where the sampling process is modified to take into account the fact that the data are chosen by the teacher: $P_L(h|d) \propto P_T(d|h)P_L(h)$ (see Figure 1).

The model has generated a variety of testable predictions about how learning from a teacher should differ from situations in which the observed data is generated by accidental or other random selection methods. Indeed, the model was the inspiration for the novel toy experiment described above ([40•]; see also [41]), and explains why children have difficulty learning the concept of

Figure 1

$$P_T(d|h) \propto P_L(h|d)$$

$$P_L(h|d) \propto P_T(d|h)P(h)$$

Current Opinion in Behavioral Sciences
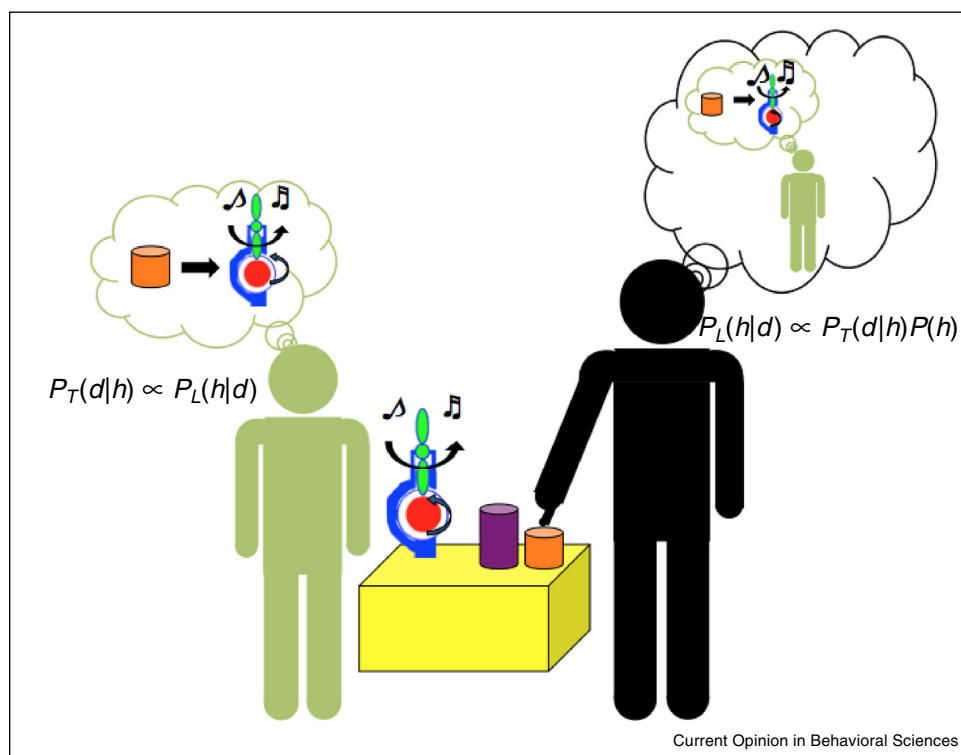
Depiction of the learner's inferences (on the left) that depend on their beliefs about how the teacher selects data (on the left). The teacher must reason about the beliefs of the learner in order to generate the most effective data.

equality [42–46]. Even children's causal inferences are sensitive to the goals and knowledge of the person generating the data, and such inferences can be modeled with this framework.

Epistemic trust

If, as indicated by research on direct instruction, learners are going to treat people as knowledgeable and helpful, then it is necessary that one consider the alternative: some people are not knowledgeable nor helpful; they can even be deceptive. Research on epistemic trust investigates whom children choose to provide information. The literature has focused on a host of superficial cues including accuracy [48], familiarity [49], consensus [50], and accent [51] to name a few. There is robust evidence that children by at least the age of 4 (and on some tasks much earlier, e.g. [52]) children are discerning about whose information they will endorse and whom they will ask.

Building on the learning *from* others framework, probabilistic models of learning *about* others provide a framework for understanding children's remarkable proficiency, and the developmental changes in their abilities, to assess who to trust for information. Shafto *et al.* [53•,54,55] formalized a model that inferred, based on the evidence provided by an informant, whether the informant was knowledgeable or not and helpful or deceptive. The choice of who to trust is then an inference about which informant's information is most likely to be correct. Modeling three representative studies from the literature, they showed that 4-year-old's ability to track informant reliability [48], to balance familiarity and recent accuracy [49], and reason about consensus [50], was explained by reasoning about the informant's knowledge and intent. The model also revealed developmental changes in reasoning. In contrast with 4-year-olds' behavior, 3-year-olds' behavior was best explained by reasoning only about informants' knowledge, assuming that informants are always helpful (see also [56–58]).

A convergent line of research has shown that trust is about more than simply tracking informants' accuracy. Consistent with the model, Gweon *et al.* [59] presented 6- and 7-year-old with informants who either demonstrated all of the functions of a toy or only one of the functions of a toy. Prior to the demonstration, children had been allowed to explore the toy, and they therefore knew how many functions the toy had. After the demonstration children rated the demonstrator on their effectiveness. Results showed that children rated informants who *omitted* information lower than those who demonstrated all of the functions, and lower than other informants who demonstrated the same number of functions on an identical toy with only one function.

The probabilistic model of epistemic trust thus formalizes social learning as a problem of *simultaneous* learning about the world *and* about informants. Literature on epistemic trust typically focuses on cases where you know the answer [48–50] and literature on children's pedagogical reasoning tends to focuses on the case where you know about the informant [40•,41]. Probabilistic modeling of joint inferences about the world and informants provide a framework for modeling learning from and about informants as a dynamic, interactive Theory-based process [57].

Summary

Developmental science has suggested that children — far from limited, broken versions of adults — are in fact powerful learners capable of remarkable feats of inference. Ideal observer approaches, instantiated in Theory-based Bayesian models, have argued that children succeed by probabilistic inference over structured knowledge representations. We have argued that these approaches are missing an important element — other people.

Research from word learning and pedagogical reasoning shows that children take the origin of data into account when learning. Given the exact same evidence, they draw different conclusions. And, these inferences can be understood through ideal observer analyses of the *social* situation. People choose evidence purposefully — in word learning, by choosing to label instances of the concept, and in pedagogical reasoning, by tailoring the amount of information provided to the complexity of the situation — and children's behavior is consistent with models that reason about the implications of people's choices for learning.

Moreover, when observing evidence, children are not just learning *from* others, but also *about* others. Research on epistemic trust has shown that young children select informants based on the goodness of the information that they provide. And, the same models that explain children's learning from pedagogically selected data have been extended to explain children's learning about the informants selecting the data.

A consensus is emerging in which children's understanding of other people plays a key role in explaining children's powerful learning abilities. This consensus draws strength from the combination of empirical results and computational models derived from ideal observer analyses of social learning. There are many important open questions, such as whether and how social ideal observer analysis should inform our understanding of children's actions and choices and practices in learning sciences more generally. These will require new research and more elaborate models. In the interim, we can conclude that to understand power of children's learning, it is important to investigate it in a social context.

Conflict of interest statement

Nothing declared.

Acknowledgements

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Diamond A, Kirkham N: **Not quite as grown-up as we like to think parallels between cognition in childhood and adulthood**. *Psychol Sci* 2005, **16**:291-297.

2. Flavell JH: **Stage-related properties of cognitive development**. *Cogn Psychol* 1971, **2**:421-453.

3. Kuhn D: **Metacognitive development**. *Curr Dir Psychol Sci* 2000, **9**:178-181.

4. Kuhn D, Amsel E, O'Laughlin M: *The development of scientific thinking skills*. Orlando, FL: Academic Press; 1988.

5. Klahr D, Dunbar K: **Developmental differences in scientific discovery processes**. In *Complex information processing: the impact of Herbert A. Simon.* Edited by Klahr D, Kotovsky K. Hillsdale, NJ: Erlbaum; 1989:109-143.

6. Piaget J: **The stages of the intellectual development of the child**. *Educational psychology in context: readings for future teachers*. 1965:98-106.

7. Zelazo PD: **The development of conscious control in childhood**. *Trends Cogn Sci* 2004, **8**:12-17.

8. Zelazo PD, Frye D, Rapus T: **An age-related dissociation between knowing rules and using them**. *Cogn Dev* 1996, **11**:37-63.

9. Kirkham NZ, Slemmer JA, Johnson SP: **Visual statistical learning in infancy: evidence for a domain general learning mechanism**. *Cognition* 2002, **83**:B35-B42.

10. Saffran JR: **Statistical language learning mechanisms and constraints**. *Curr Dir Psychol Sci* 2003, **12**:110-114.

11. Saffran JR, Johnson EK, Aslin RN, Newport EL: **Statistical learning of tone sequences by human infants and adults**. *Cognition* 1999, **70**:27-52.

12. Gopnik A, Schulz L: **Mechanisms of theory formation in young
- children**. *Trends Cogn Sci* 2004, **8**:371-377.
A review of the causal learning literature (by two leaders in the field) presenting a summary of evidence showing that young children learn from conditional probabilities and causal interventions in a way consistent with Bayes net learning algorithms.

13. Tenenbaum JB, Kemp C, Griffiths TL, Goodman ND: *How to grow a mind: statistics, structure, and abstraction. science*. 2011, **331**:1279-1285.

14. Perfors A, Tenenbaum JB, Griffiths TL, Xu F: **A tutorial introduction to Bayesian models of cognitive development**. *Cognition* 2011, **120**:302-321.

15. Schulz LE, Bonawitz EB, Griffiths TL: **Can being scared cause tummy aches? Naive theories, ambiguous evidence, and preschoolers' causal inferences**. *Dev Psychol* 2007, **43**:1124.

16. Griffiths TL, Chater N, Kemp C, Perfors A, Tenenbaum JB: **Probabilistic models of cognition: exploring representations and inductive biases**. *Trends Cogn Sci* 2010, **14**:357-364.

17. Xu F, Tenenbaum JB: **Word learning as Bayesian inference**.
- *Psychol Rev* 2007, **114**:245.
The authors present a Bayesian framework for word learning, exploring both children and adults' predictions about word meanings at multiple levels of a taxonomy. It presents a particularly nice overview of the shortcomings of past theories of word learning, as well as a compelling set of experiments to test quantitative and qualitative predictions of their model.

18. Xu F, Tenenbaum JB: **Sensitivity to sampling in Bayesian word learning**. *Dev Sci* 2007, **10**:288-297.

19. Tenenbaum JB, Griffiths TL, Kemp C: **Theory-based Bayesian
- models of inductive learning and reasoning**. *Trends Cogn Sci* 2006, **10**:309-318.
An important overview by a number of the leaders in the theory-based Bayesian modeling approach, arguing that both strong constraints from structured knowledge and powerful statistical learning can be combined to understand human inductive inference.

20. Kemp C, Perfors A, Tenenbaum JB: **Learning over-hypotheses with hierarchical Bayesian models**. *Dev Sci* 2007, **10**:307-321.

21. Gopnik A, Wellman HM: **Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory**. *Psychol Bull* 2012, **138**:1085.

22. Kemp C, Tenenbaum JB: **The discovery of structural form**. *Proc Natl Acad Sci* 2008, **105**:10687-10692.

23. Yu C, Ballard DH: **A unified model of early word learning: integrating statistical and social cues**. *Neurocomputing* 2007, **70**:2149-2165.

24. Thomaz AL, Berlin M, Breazeal C: **An embodied computational model of social referencing**. *IEEE international workshop on robot and human interactive communication, 2005. ROMAN 2005*. IEEE; 2005 August:591-598.

25. Breazeal C, Scassellati B: **Robots that imitate humans**. *Trends Cogn Sci* 2002, **6**:481-487.

26. Talwar V, Lee K: **Social and cognitive correlates of children's lying behavior**. *Child Dev* 2008, **79**:866-881.

27. Wellman HM, Liu D: **Scaling of theory-of-mind tasks**. *Child Dev* 2004, **75**:523-541.

28. Gergely G, Csibra G: **Teleological reasoning in infancy: the naïve theory of rational action**. *Trends Cogn Sci* 2003, **7**:287-292.

29. Repacholi BM, Gopnik A: **Early reasoning about desires: evidence from 14-and 18-month-olds**. *Dev Psychol* 1997, **33**:12.

30. Woodward AL: **The infant origins of intentional understanding**. *Adv Child Dev Behav* 2005, **33**:229-262.

31. Baillargeon R, Scott RM, He Z: **False-belief understanding in infants**. *Trends Cogn Sci* 2010, **14**:110-118.

32. Quine WV: *Word and object*. Cambridge, MA: MIT Press; 1960:. 1969.

33. Bloom L: **The intentionality model of word learning: how to learn a word, any word**. *Becoming a word learner: a debate on lexical acquisition*. 2000:19-50.

34. Carey S: **The child as a word learner**. In *Linguistic theory and psychological reality.* Edited by Halle M, Bresnan J, Miller G. Cambridge, MA: MIT Press; 1978.

35. Gold EM: **Language identification in the limit**. *Inf Control* 1967, **16**:447-474.

36. Markman EM: **Constraints children place on word meanings**. *Cogn Sci* 1990, **14**:57-77.

37. Regier T: *The human semantic potential: spatial language and constrained connectionism*. MIT Press; 1996.

38. Shafto P, Goodman N: **Teaching games: statistical sampling assumptions for learning in pedagogical situations**. In *Proceedings of the thirtieth annual conference of the cognitive science society*. Austin, TX: Cognitive Science Society; 2008, 1632-1637.

39. Shafto P, Goodman ND, Frank MC: **Learning from others the
- consequences of psychological reasoning for human learning**. *Perspect Psychol Sci* 2012, **7**:341-351.
Introduces the 'learning from others' computational framework. The paper proposes that learners reason about other people's choices. The implications are illustrated by contrasting three levels of a taxonomy of social learning: accidental actions by not knowledgeble actors, purposeful actions by knowledgeable actors, and communicative actions by knowledgeable actors. Simulations show that the exact same evidence leads to very different inferences across these contexts.

40. Bonawitz E, Shafto P, Gweon H, Goodman ND, Spelke E, Schulz L:
   • **The double-edged sword of pedagogy: instruction limits spontaneous exploration and discovery**. *Cognition* 2011, **120**:322-330.
   This paper provides a computational analysis showing how teaching effects exploration and discovery, and empirical support for the prediction that children tend to limit their exploratory play to a target function highlighted during direct instruction.

41. Buchsbaum D, Gopnik A, Griffiths TL, Shafto P: **Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence**. *Cognition* 2011, **120**: 331-340.

42. McNeil NM: **Limitations to teaching children 2+2=4: typical arithmetic problems can hinder learning of mathematical equivalence**. *Child Dev* 2008, **79**:1524-1537.

43. McNeil NM, Alibali MW: **Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations**. *Child Dev* 2005, **76**:883-899.

44. McNeil NM, Alibali MW: **You'll see what you mean: students encode equations based on their knowledge of arithmetic**. *Cogn Sci* 2004, **28**:451-466.

45. Behr M: **How children view the equals sign**. *Math Teach* 1980, **92**:13-15.

46. Seo KH, Ginsburg HP: *'You've got to carefully read the math sentence…'*: classroom context and children's interpretations of the equals sign. 2003.

47. Shafto P, Goodman ND, Griffiths TL: **A rational account of pedagogical reasoning: teaching by, and learning from, examples**. *Cognit Psychol* 2014, **71**:55-89.

48. Pasquini ES, Corriveau KH, Koenig M, Harris PL: **Preschoolers monitor the relative accuracy of informants**. *Dev Psychol* 2007, **43**:1216.

49. Corriveau K, Harris PL: **Choosing your informant: weighing familiarity and recent accuracy**. *Dev Sci* 2009, **12**:426-437.

50. Corriveau KH, Fusaro M, Harris PL: **Going with the flow preschoolers prefer nondissenters as informants**. *Psychol Sci* 2009, **20**:372-377.

51. Kinzler KD, Corriveau KH, Harris PL: **Children's selective trust in native-accented speakers**. *Dev Sci* 2011, **14**:106-111.

52. Tummeltshammer KS, Wu R, Sobel DM, Kirkham NZ: **Infants track the reliability of potential informants**. *Psychol Sci* 2014, **25**:1730-1738.

53. Shafto P, Eaves B, Navarro DJ, Perfors A: **Epistemic trust:**
   • **modeling children's reasoning about others' knowledge and intent**. *Dev Sci* 2012, **15**:436-447.
   Introduces a computational model of epistemic trust, formalized as a causal Bayes Net relating beliefs, intent, and actions. The model generalizes previous accounts of learning from others to provide an account of learning about others. They model three representative studies from the epistemic trust literature. Results show that behavioral differences between 3 and 4-year-olds are explained by an increased ability to reason about informants' helpfulness.

54. Bovens L, Hartmann S: **Bayesian epistemology**. *OUP Catalogue*. 2004.

55. Butterfield J, Jenkins OC, Sobel DM, Schwertfeger J: **Modeling aspects of theory of mind with Markov random fields**. *Int J Soc Robot* 2009, **1**:41-51.

56. Eaves BS Jr, Shafto P: **Unifying pedagogical reasoning and epistemic trust rational constructivism in cognitive development**. *Adv Child Dev Behav* 2012, **43**:295-320.

57. Landrum AR, Eaves BS, Shafto P: **Learning to trust and trusting to learn: a theoretical framework**. *Trends Cogn Sci* 2015, **19**:109-111.

58. Mascaro O, Sperber D: **The moral, epistemic, and mindreading components of children's vigilance towards deception**. *Cognition* 2009, **112**:367-380.

59. Gweon H, Pelton H, Konopka JA, Schulz LE: **Sins of omission: children selectively explore when teachers are under-informative**. *Cognition* 2014, **132**:335-341.

60. Bonawitz EB, Lombrozo T: **Occam's rattle: children's use of simplicity and probability to constrain inference**. *Dev Psychol* 2012, **48**:1156.

61. Piaget J, Inhelder B: *The growth of logical thinking from childhood to adolescence*. New York: Basic Books; 1958.