Infant statisticians: The origins of reasoning under uncertainty

Stephanie Denison* and Fei Xu†

*Department of Psychology, University of Waterloo

†Department of Psychology, University of California, Berkeley

(In press, Perspectives on Psychological Science)

Word count (main text and footnotes): 6,116

Corresponding author: Stephanie Denison (stephanie.denison@uwaterloo.ca);

200 University Ave W, Waterloo, Ontario, N2L 3G1; 519-888-4567 x 33409

Abstract

Human learners frequently make inferences about uncertain future events based on limited data. What are the developmental origins of this ability? Over the past 10 years, foundational work in this area suggests that infants and other primates can make surprisingly sophisticated inferences under uncertainty (e.g., Rakoczy et al., 2014; Téglás et al., 2007; Xu & Garcia, 2008). The current paper first asks what underlying cognitive mechanisms might allow young learners to make such sophisticated inferences under uncertainty. We outline three possibilities, the *logic*, *probabilistic*, and *heuristics* views, and assess the empirical evidence for and against each. We argue that the weight of the empirical work points in favor of the *probabilistic* view, in which early reasoning under uncertainty is grounded in proportional reasoning and inferences about the relationship between samples and populations, as opposed to being grounded in simple heuristics. Second, we discuss the apparent contradiction between this early-emerging sensitivity to probabilities when reasoning under uncertainty with the decades of literature suggesting that adults show limited use of base-rate and sampling principles in their inductive inferences (e.g., Kahneman, 2011). Third, we ask how these early inductive abilities can be harnessed for improving later mathematics education and inductive inference. We make a number of suggestions throughout for future empirical work, which should go a long way in addressing the many remaining open questions in this growing research area.

Keywords: Infant cognition; probability; learning; inductive inference

Infant statisticians: The origins of reasoning under uncertainty

Imagine you are a receptionist at an office and you have a bowl of mini candy bars at your desk. Around noon each day, various staff members come by and take one. You currently have a mix of mostly Snickers bars and just a few Kit-Kats. For a few days in a row, your coworker Jordan walks by and grabs a Snickers bar each time and your coworker Alex walks by and always grabs a Kit-Kat bar. Does Jordan really like Snickers? Does Alex really like Kit-Kats? You probably have the intuition that Alex definitely prefers Kit-Kats, but what about Jordan? Because she could have grabbed those Snickers bars at random, it is harder to know what she prefers as compared to Alex, whose choices would be unlikely given random selection. Thus, despite the fact that each person made consistent choices, the distribution of available bars made one person's choices seem more intentional, providing a stronger basis for a preference attribution in the case of Alex than Jordan. This intuition relies on an appreciation of the basic principles of probability, including recognizing that a sample containing only the majority item from a distribution could easily arise from chance but a sample containing only the minority item is suggestive of non-random selection.

While using probability (as opposed to facial expressions or explicit statements about desires) to infer another's preferences seems very advanced, research suggests that even young children can make rational inferences such as these. That is, in experimental paradigms very similar to the candy bar example, infants, toddlers, and 4-year-old children have successfully used this kind of statistical information to infer agents' preferences (Kushnir, Xu, & Wellman, 2010; Wellman, Kushnir, Xu, & Brink, 2016). These findings are part of a larger literature examining the developmental origins of inductive inference more broadly, and the basic statistical intuitions that underlie those inferences (for reviews see Xu & Kushnir, 2012, 2013). The ability to engage in inductive inference – which can be defined as generating an expectation based on incomplete, and sometimes sparse, information – is particularly challenging because the learner must use this variable input to arrive at a best guess, about which they cannot be certain. The recent studies examining preference attributions (Diesendruck, Salzer, Kushnir, & Xu, 2015; Kushnir et al., 2010; Ma & Xu, 2011; Wellman et al. 2016) represent one example from a burgeoning literature showing that infants, toddlers, and preschoolers can make inferences using probabilistic data.

With the accumulation of this recent literature showing impressive inductive reasoning in young children, researchers have been struck by an apparent contradiction. On the one hand, these findings suggest that children are sensitive to base-rates and sampling, but on the other hand, decades of research in cognitive psychology suggests that adults often fail to integrate base-rate data and random sampling in their judgments and decision-making (see Kahneman, 2011 for a review). That is, in many classic experiments, adults tend to base their judgements almost exclusively on information about a person's personality traits or other personal diagnostic information and under-value or even ignore the relevant statistical or base-rate information (Kahneman & Tversky, 1971, 1973; Tversky & Kahneman, 1974). For example, in the classic lawyer-engineer problem, study participants were given the base-rate, that 70 people in a group were lawyers and 30 were engineers. They were then given a personality description of that individual, which was highly representative of the stereotypes associated with engineers and were asked to judge how likely it was that the person in question fell into one of the categories. In these experiments, participants almost entirely ignore the statistical base rate information, and focus on the personality description when making their ratings.

These classic tasks differ in numerous and important ways from the child experiments mentioned above and from the infant literature discussed below. Nonetheless, the contrast of young children's acute sensitivity to base-rates, and adults' tendency to ignore them in a variety of contexts, has raised important questions about the developmental origins of reasoning under uncertainty. We discuss this contrast, beginning with a review and critique of the recent empirical literature examining young children's inductive inferences. Thus far the evidence suggests that adults from pre-numeric cultures, human infants, all species of great ape, and at least one species of New World monkey can make judgments about future uncertain events using base-rate information (Denison, Reed, & Xu, 2013; Denison, Trikutam, & Xu, 2014; Denison & Xu, 2010, 2014; Fontanari, Gonzalez, Vallortigara, & Girotto, 2014; Lawson, & Rakison, 2013; Rakoczy et al., 2014; Tecwyn, Denison, Messer, & Buchsbaum, 2017; Téglás, Girotto, Gonzalez, & Bonatti, 2007; Téglás, Ibanez-Lillo, Costa, & Bonatti, 2015; Téglás et al. 2011; Xu, & Garcia, 2008). This body of work raises a number of questions about the nature and development of this ability, most notably: What underlying cognitive mechanisms make these inferences possible? We then turn to the question of why adults often neglect base-rate and other statistical information when infants and young children are so adept at using it? And finally, we ask whether we can capitalize on some of these intuitive abilities to improve later mathematical and inductive reasoning.

The origins of reasoning under uncertainty: A review of empirical research

Using a variety of methods, and from phylogenetic and ontogenetic perspectives, rational reasoning under uncertainty has been revealed surprisingly early in development. The most common methods for assessing statistical inference in pre- and non-verbal populations are violation of expectation (VOE) looking time tasks and choice tasks. Imagine that you are presented with a lottery machine on a computer screen that contains three yellow crosses and one blue cube. The objects bounce around in accord with the principles of physics, and then the contents are briefly covered as one item exits: do you think it will be a yellow cross or a blue cube? If you think it's likely to be a yellow cross, then you agree with 12-month-old infants, who look longer when the single blue item, rather than when one of the yellow items, exits the machine (Téglás et al., 2007). The logic of VOE looking time is that infants look longer at unexpected events than at expected events and it has been applied widely in research on infant perception and cognition (see Aslin, 2007 for a review of infant looking paradigms).

Recently this methodology has been used in a number of paradigms assessing reasoning under uncertainty in human infants. These experiments have revealed that infants from 6 to 12 months can make inferences about which of two event outcomes is more or less likely, given the statistical attributes of its population or source, as indexed by infants looking longer at improbable outcomes than probable outcomes (see Figure 1, panel A, B). For example, in another VOE looking paradigm, infants are shown a large box containing many red balls and a few white balls. When balls are drawn randomly from this box, infants and apes look longer at a small collection of mostly white balls (an improbable, outcome) than at a small collection of mostly red balls (a probable, outcome). Infants can also make this inference in the reverse direction, inferring that when a small collection of mostly red balls is drawn from a box, the box itself is likely to have a larger proportion of red than white balls (Eckert, Call, & Rakoczy, 2017; Xu & Garcia, 2008; see Placi et al., in press, suggesting that long-tailed macaques do not make this reverse inference). These findings suggest that infants recognize that when an item is randomly drawn from a container, the item is most likely to be of the majority type.

Recent studies have also used choice tasks with infants and non-human animals to investigate similar questions (see Figure 1, panel C, D). Thus far all of the examples of reasoning under uncertainty provided (i.e., attributing preferences to individuals and making predictions from lottery machines) seem uniquely relevant to humans. However, numerical competence is critical to survival for non-human animals, allowing them to engage in efficient foraging and providing an edge in intergroup-conflict (e.g., Addessi, Crescimbene &, Visalberghi, 2008; Wilson, Hauser, & Wrangham, 2011). Similar to humans, numerical reasoning about absolute quantities alone is insufficient for many of the inferences they must make (Rugani, Vallortigara, & Regolin, 2015). For example, in order to maximize the quantity of food an animal can access they must consider the relative relationship between available food quantities and the number of animals feeding at different locations (Harper, 1982). Also, recent work suggests that some monkey species experience inequity aversion – they are aware of the relative discrepancies between their individual effort and payoff compared to that of another individual (e.g. Brosnan & de Waal 2003; Cronin & Snowden, 2008). Therefore, similar statistical reasoning abilities are likely to be present in these non-human animal species for reasoning under uncertainty.

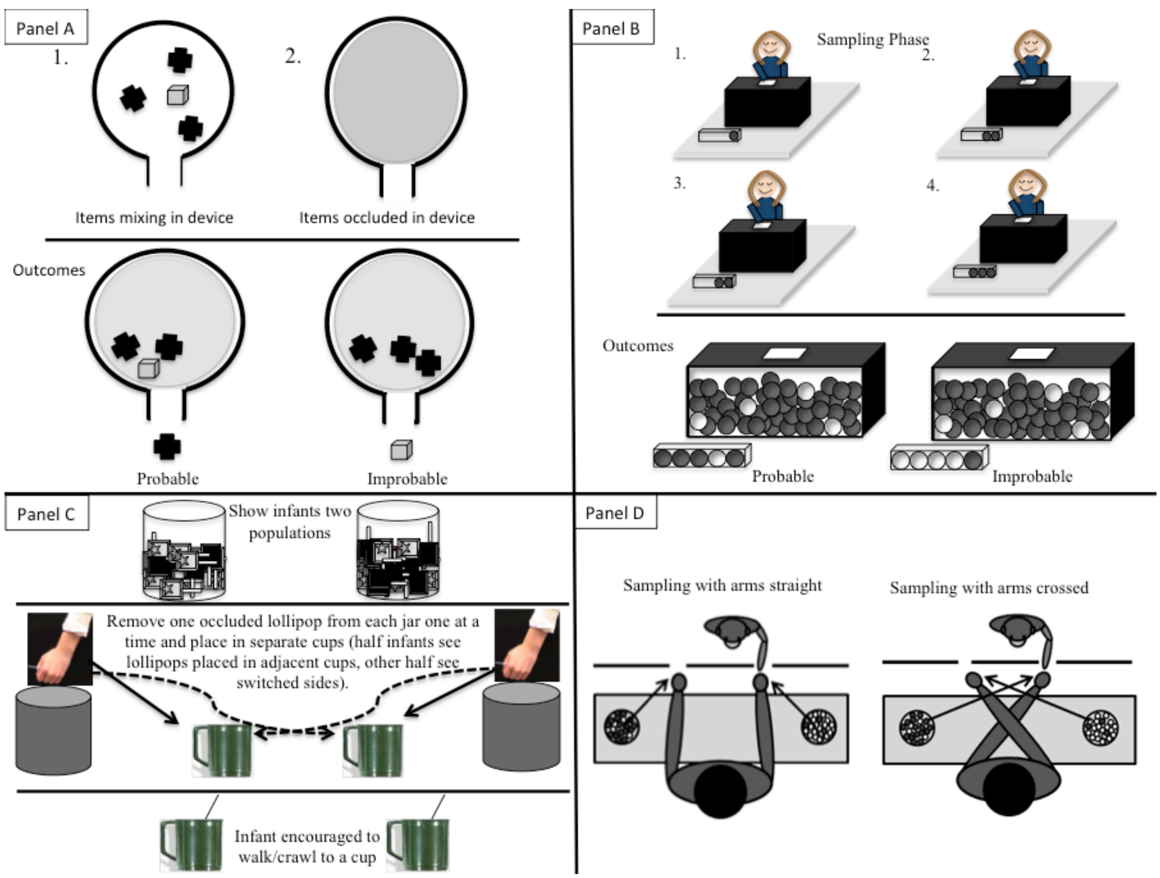
In the choice tasks depicted in Figure 1 (Panels C and D), participants (infants in some studies, non-human primates and monkeys in others) are shown two populations of items and are tasked with making a prediction about the likely outcome of a single draw based on the composition of the populations. Importantly, one population has a higher proportion of preferred to non-preferred items than the other population and so the participants are motivated to choose the sample from the container with the greater proportion of preferred items. In a number of experiments, infants, great apes, and capuchins typically choose to look for a hidden, unknown sample from the population with a higher probability of yielding a preferred item (Denison & Xu, 2010, 2014; Eckert et al., 2018a; Rakoczy et al., 2014; Tecwyn et al., 2017). The convergence between VOE and choice tasks suggests that the ability to reason under uncertainty based on statistical sampling information is robust early in human development: the representations are strong enough to support looking time differences (which may only require 'post-diction', Haith, 1998) and to guide action (which requires prediction).

These findings have been replicated multiple times with a number of subject populations, demonstrating sophisticated quantitative and inferential abilities.[1] However, a critical open question remains unanswered in this line of work: *How* do infants and non-human primates make these judgments about uncertain events? Does the early emerging ability to make inferences under uncertainty stem from foundational knowledge about logic (i.e., comparisons of logically possible outcomes), probabilities (i.e.,

[1] While probabilistic reasoning has been shown in infants, and sophisticated applications of these abilities have been shown in young children, recent work has also demonstrated important limitations on how children apply probabilistic reasoning. For example, in mental state reasoning, where previous studies have found that children successfully infer preferences from non-random sampling, Garvin and Woodward (2015) found that statistical information alone was in fact insufficient for 3-year-olds to infer preferences. Their work suggests that children may struggle to select appropriate hypotheses to consider in the first place and verbal framing provides a context in which children can apply their probabilistic intuitions.

statistical reasoning about samples and populations), or simple heuristics (i.e., shortcuts that can appear on the surface as rational inference but also introduce systematic errors)? We outline these three possibilities, termed the *Logic*, *Probabilistic*, and *Heuristics* views, respectively, and discuss the relevant empirical evidence for and against each. Specifying the basic cognitive mechanisms that underlie these abilities will refine our understanding of how inductive inference unfolds early in development and will shape the research questions to be pursued in future work.

Figure 1. Schematics of typical paradigms. Panel A: Lottery-machine paradigm (Téglás et al., 2007). Infants watch objects moving in machine; items briefly occluded; probable and improbable outcomes revealed on alternating trials; infants look longer at improbable outcome. Panel B: Ping-Pong ball paradigm (Xu & Garcia, 2008). Experimenter samples balls one at a time from closed box then large distribution revealed; samples alternate between probable and improbable; infants look longer at improbable outcome. Panel C: Lollipop paradigm (Denison & Xu, 2010). Infants shown two large distributions with different proportions of preferred and non-preferred items; single lollipop removed from each container and placed in separate cups; infants choose cup with lollipop from more favorable distribution. Panel D: Non-human primate paradigm (Rakoczy et al., 2014). Participants shown two large distributions with different proportions of preferred to non-preferred food; experimenter samples one from each in closed hand; participants choose hand from more favorable distribution.



Underlying cognitive mechanisms: Three proposals

Consider the problems in Figure 1, posed to infants in looking time tasks and to infants and other primates in action tasks. How could one go about solving them?

One way to solve these problems is through logical inference, a view that has been proposed and explicated in several recent papers (Cesana-Arlotti, Téglás, & Bonatti, 2012; Téglás et al., 2007; 2011; 2015). The *logic view* suggests that the foundation of reasoning under uncertainty is grounded in intuitive modal logic – infants represent future events via a logical sense of possibilities. This is in contrast with reasoning under uncertainty being grounded in a statistical sense of probability or heuristics. A logic proposal based on modal reasoning suggests that infants reason about the likelihood of particular outcomes for a novel, single event by enumerating the possibilities, and then comparing the number of possible outcomes of each kind. Consider again the lottery machine example presented earlier (which is based on the scenario represented in Figure 1, Panel A). In these VOE looking time experiments, infants observe a lottery machine displaying three yellow crosses and one blue cube moving randomly in the machine. From previous familiarization with the machine, they know that just one object will eventually reach an opening, apparently at random, and exit. Twelve-month-old infants look longer when observing the unlikely event of a blue square exiting the machine than the more likely event of a yellow cross exiting the machine. Téglás, Bonatti, and their colleagues propose that these intuitions about uncertain future events result from infants' intuitive logical capacities. That is, when infants view this scene, they represent it modally, as a set of logically possible future states: three in which a yellow item exits the machine and one in which a blue item exits. The infants keep track of these potential outcomes via object tracking, or subitizing, and then compare the number of outcomes of each object type to determine the most likely outcome. In the example, there are three outcomes in which a yellow cross exits the machine and one in which the blue square exits. Infants can compare the numbers of these possible events and conclude that a yellow cross is the most likely outcome.

Another way to solve the problems in Figure 1 is to use heuristics, or mental shortcuts, as opposed to logical or probabilistic reasoning. The *heuristics view* suggests that if infants demonstrate any skills at all in these kinds of problems, then what they would really be engaging in is heuristic reasoning that only appears on the surface as logical or probabilistic inference, and this reasoning would be prone to bias. That is, many dual processing views assume that human learners start out relying heavily, or perhaps exclusively, on heuristics and they only proceed to full analytical reasoning (specifically in this case, probabilistic reasoning) with the onset of language (see Kokis et al., 2002 for a review of dual process accounts). Note that this is different from the assumption that analytical reasoning replaces heuristic reasoning as development progresses, which has been referred to as the illusion of replacement and has found limited support (Stanovich, Toplak, & West, 2011). We are instead referring to the assumption that true analytical reasoning may not be present *at all* in preverbal infants and that heuristic processing constitutes the majority of infants' reasoning.

A number of heuristics or shortcuts could be employed to solve the problems posed to infants and other primates in the experiments on reasoning under uncertainty. One such heuristic would unfold as follows: to determine which of the two item types is likely to exit the machine in Figure 1, Panel A, infants see that there are more yellow than blue items and conclude that a yellow object should be drawn based on the heuristic that more items of one type will lead to an item of that type being selected (the *more* heuristic). This view predicts that when analytical reasoning is pitted against heuristic reasoning, infants should behave in accord with the heuristic response, and not the analytical response. Thus, infants should not be capable of overriding a response based on the *more* heuristic (or any other heuristic) if the situation were to call for it.

Finally, infants could solve all the tasks discussed above via probabilistic inference, as opposed to logical or heuristic reasoning. The *probabilistic view* suggests that infants' ability to engage in non-verbal reasoning under uncertainty results from their statistical intuitions, in which they estimate proportions and consider the relationship between samples and populations (Denison, et al., 2013; Denison et al., 2014; Denison & Xu, 2014; 2010a, b; Rakoczy et al., 2014; Xu & Garcia, 2008). On this view, when posed with the problem in Figure 1, Panel A, infants begin by encoding the proportion of items. This estimate of proportions could be derived from either of infants' two quantitative systems for representing number or continuous variables, the approximate number system or the object tracking system.[2] Then, if the sample was generated randomly, they infer that the yellow item is the more likely outcome. The central predictions of this view are (1) the computations performed will be predicated on an assumption of random sampling and (2) analytic responses should, at least sometimes, override simple heuristic responses, such as a response from the *more* heuristic.

For the lottery-machine task and the Ping-Pong ball task, the correct application of logic, simple heuristics, or probability would result in the same response patterns in infants, thus each proposed mechanism could account for the data from these initial

[2] How numerosities are represented by each system differs (e.g., the ANS represents large approximate numerosities, while the object tracking system indirectly represents precise numbers up to four). Very little infant work has examined how proportions or ratios are encoded by infants. One study by McCrink and Wynn (2006) found that 6-month-old infants' discrimination of different ratios of visual-spatial arrays showed the same signatures of ANS as in infant studies with large numbers (e.g., Lipton & Spelke, 2003; Xu & Spelke, 2000).

studies. Research conducted over the past few years has begun to address which mechanism underlies the origins of reasoning under uncertainty by posing problems that should yield different patterns of behavior depending on which mechanism is at work. In particular, the probabilistic view has been contrasted with both the logic and the heuristics views. In the following section, we review the findings from this line of research, first contrasting the probabilistic and logic views and then contrasting the probabilistic and heuristics views.

A comparison of the Logic and Probabilistic views

The central predictions of the Logic and Probabilistic views converge on two points: First, both predict that infants should not predominantly rely on heuristics when reasoning under uncertainty (a prediction that we return to in the next section). Second, both predict that infants should be capable of making inferences in the absence of experiencing past frequencies. The initial research by Bonatti, Denison, Téglás, Xu, and their colleagues clearly shows that infants do not require past frequency information to make inferences about future events. In these tasks, infants are shown collections of objects and they infer the most likely outcomes from random draws without having had the opportunity to accumulate information from observing the outcomes of sampling events (see Denison & Xu, 2010; Téglás et al., 2007; Xu & Garcia, 2008).

The views diverge in that the Logic view suggests that this reasoning is grounded in the enumeration of logical possibilities, whereas the Probabilistic view suggests that this reasoning is grounded in the statistical relationship between populations and samples. Thus, the most diagnostic empirical tests of which mechanism underlies reasoning will be ones that determine whether infants are only capable of making these inferences when presented with small numbers of items: If infants are using statistical representations of the relationship between samples and populations, then there should be no such limit. If infants are representing the scene in a modal way, deriving the possible states of affairs and comparing the number of outcomes of each type, then their ability should be limited by the number of items or events infants can represent in parallel. Téglás and colleagues (2015) found support for such a limit in an experiment using the lottery machine paradigm, in which infants were unable to make predictions when the total number of items were increased to 16 (12:4, still the same 3:1 ratio). However, evidence from choice tasks with human infants, great apes, and capuchin monkeys suggest that all of these populations can make inferences about single items from large sets (see Figure 1, Panels C & D). In each of these experiments, participants have chosen a single hidden item from a population with a higher proportion of their preferred item, with numbers of items ranging from 12 to 500 (e.g., Denison & Xu, 2014; Rakoczy et al., 2014; Tecwyn et al., 2017). Any or all of the many differences between the two paradigms could be responsible for the discrepant findings, including the difference in dependent measures (i.e., looking versus choice), and possible inequities in motivation between the stimuli (i.e., looking at neutral items versus choosing preferred items), to name a few. There is a strong possibility that basic statistical intuitions and basic logical intuitions might both support reasoning under uncertainty, in different contexts. We return to this possibility later.

A comparison of the Heuristics and Probabilistic Views

The main predictions of the heuristics and probabilistic views naturally contrast with one another, as each predicts that when pitted against one another, the response from that "system" will prevail. Three lines of research have examined the predictions from these views.

First, researchers have begun examining whether infants and other primates use heuristics based on simple absolute quantity comparisons that children and adults often rely on in both formal mathematics and in some choice problems (e.g., Falk, Yudilevich-Assouline, & Elstein, 2012). In recent work using the choice paradigms depicted in Figure 1, Panels C and D, researchers have systematically contrasted the predictions from analytical reasoning with variations of the *more* heuristic, which often result in denominator neglect (e.g., Falk et al., 2012). For example, infants, non-human primates and capuchins have been presented with tasks in which the absolute quantity of target (i.e., preferred) items in two contrasting populations are equated, thus eliminating the ability to make choices based simply on which population has *more targets* and forcing participants to consider the proportions of items. These populations succeed at these tasks, as well as other tasks in which the absolute number of targets is lower in the more probable population (Denison & Xu, 2014; Rakoczy et al., 2014; Tecwyn et al., 2017).

While these findings suggest that infants and non-human primates can override a heuristic response in these particular paradigms, heuristics clearly continue to influence reasoning throughout the lifespan. As referenced earlier, school-age children continue to sometimes rely on versions of the *more* heuristic when making explicit judgments about which of two urns is most likely to yield a particular color ball (Falk et al., 2012). Also, in recent experiments, 3- and 4-year-old children failed to make correct inferences on a number of choice tasks that were very similar in design to the infant choice tasks (Girotto et. al, 2016). Girotto and colleagues suggest that preschoolers' difficulties might result from the preschooler tasks placing higher executive functioning demands on children than the infant tasks. Preschoolers are explicitly told they must wait for a reward based on their choice, whereas in the infant tasks, choices and subsequent rewards are produced more quickly. It is difficult to know why infants (and other primates) sometimes show competence in cases where older children do not, but this work from the choice tasks indicates that, at minimum, pre-verbal infants do not *always* rely on simple heuristics.

While variants of the *more* heuristic have received considerable attention in the cognitive development and education literatures, it is not the only heuristic that learners could use. There are numerous demonstrations of adults relying on a variety of judgment heuristics rather than applying the principles of probability in more complex inductive inference tasks. Do infants also rely on these judgment heuristics? One heuristic that adults tend to rely on is the *representativeness heuristic*, which can lead to base-rate neglect under a variety of circumstances. Use of this heuristic results in the biased judgments shown in tasks like the lawyer-engineer problem described earlier – adults rely on the personality description and ignore base rates because the description fits their representation of a typical engineer. Because this heuristic, in its simplest form, can be described as an assumption that the surface features of a sample should represent the surface features of the population from which it is drawn (Tversky & Kahneman, 1974), this could explain infants' success in problems such as the Ping-Pong ball paradigm (i.e., Xu & Garcia, 2008). To address this, a number of experiments have examined how infants behave when posed with problems that pit a response from perceptual representativeness against a response based on base-rates (Denison & Xu, 2010; Denison et al., 2014). In these looking time experiments, infants were shown that the more numerous balls in a population had a property that caused a large proportion of them to remain stuck inside the box, therefore unavailable for sampling. In these cases, infants reversed their expectations, reasoning that the sample should have a greater number of the minority-colored balls, rather than a greater number of majority-colored balls.

Finally, a major component of making correct probabilistic inferences is that the learner should consider how a sample is generated when judging its likelihood. Thus, if infants are engaging in true probabilistic inference they should be flexible in their expectations depending on whether a sample is drawn intentionally or randomly. This flexibility has been tested with 11-month-old infants using the paradigm depicted in Figure 1, Panel B (Xu & Denison, 2009; see also Gweon, Tenenbaum, & Schulz, 2010 for converging evidence from 16-month-olds, and Eckert et al., 2018b in apes using different methods and Denison & Xu, 2010, 2014; Téglás et al., 2007 for similar results in the naïve physics domain). In this experiment, when an agent expressed a goal or preference for one color of balls and then intentionally drew balls from the box (she looked into the box and deliberately chose balls), infants expected that the sample should reflect the agent's goal and not the statistical properties of the box. However, when the agent demonstrated an initial preference but then drew balls randomly from the box (she carefully demonstrated that she could not see what she was sampling by using a blindfold), infants expected that the sample should be similar in statistical properties to the larger population. This finding suggests that infants do not automatically assume that a sample should always match a distribution in statistical properties, which provides strong support for true probabilistic inference.

Looking forward: Implications and future directions

We argue that the balance of the reviewed evidence leans in favor of the probabilistic view but the logic view also has much empirical support. A simple heuristics proposal in which infants are *only* capable of relying on heuristics is becoming increasingly unlikely.

The discrepancy in findings regarding single-event probabilistic inferences with large versus small quantities should be further explored, as resolving this will be critical to teasing apart whether logic or probability underlie these inferences. To gain additional clarity on infants' abilities, some gaps in empirical evidence should be addressed: First, the Ping-Pong ball looking-time paradigm, which has been used with large populations and multi-item samples (Figure 1, Panel B) has not yet been used with large populations and single-item sampling events. If infants succeed at this type of task, it would lend support to the interpretation that differences between the stimuli in the lottery-machine paradigm versus the choice paradigm are responsible for the discrepant findings. One way to test this directly is to employ an experimental technique that manipulates the speed and number of the moving objects to reveal infants' limits on tracking, enumerating, and extracting the ratio of such objects. Multiple-object tracking experiments with adults suggest that the upper limit of four objects can be increased if the speed at which the objects move is reduced (e.g., Alvarez & Franconeri, 2007), thus infants might be more or less successful at these tasks depending on the number of objects and their rate of motion.

Implications for the development of judgment and decision-making

One of the most intriguing questions raised by rejecting a heuristic account of infant probabilistic inference is how this relates to well-documented heuristic use, and base-rate neglect, in adults. One initial caveat is that the infant tasks conducted thus far have been necessarily much simpler in design than those used with adults, as illustrated in the infant experiments examining representativeness (Denison et al., 2014). Relatedly, the format of the presented base-rates is notably different than in classic tasks in that they are presented visually, whereas most adult tasks were presented verbally. Due to this visual format, it is likely that participants encode the base-rates in frequency format. As we know very well from research with adults and older children, this frequency format often facilitates the use of statistical information such as base rate (Cosmides & Tooby, 1996; Gigerenzer, 1991; Gigerenzer & Hoffrage, 1995; Hoffrage & Gigerenzer, 1998). One practical take-away from this is that adult comparison groups will be vital to any explorations of the development of heuristic use, to ensure that the format of the statistical information alone is not the driving force if greater base-rate use is in fact observed in younger children than typical adult performance. Second, much work is still needed to render a full picture of whether infants and very young children will commit other reasoning fallacies that stem from misuse or neglect of numerical information, such as falling prey to the “law of small numbers” and the gambler’s fallacy, engaging in anchoring and adjustment, failing to integrate base-rates with diagnostic information, and so on (but see Girotto & Gonzalez, 2008 for evidence that 5-year-olds can compute posterior probabilities by integrating priors and likelihoods). The accumulation of this empirical evidence will be critical to mapping the emergence of base-rate neglect and other biases.

Recent empirical work is beginning to tackle the question of whether real developmental differences exist in base-rate versus heuristics use across the lifespan. One possibility is that heuristic use (and its corresponding biases) develop later in childhood and strengthen as learners engage in more and more real-world judgments. A recent examination of the representativeness heuristic in 4- to 6-year-old children supports this idea (Gualtieri & Denison, 2018). In these experiments, children were presented with child-friendly versions of the classic lawyer-engineer problem. For example, they were told that an individual, who likes to play with trucks and trainsets (stereotypes that a separate group of same-aged children readily endorsed as being more indicative of boys than girls) was sampled from a visually presented group of 8 girls and 2 boys. Children (and a group of adults) were asked to classify the group membership of the mystery individual. By age 6, children nearly always guessed that the individual was a boy, whether they were presented with this conflicting 2:8 base-rate or an opposite base-rate of 8:2, showing base-rate neglect at levels similar to adults in this task. Interestingly, at age 4, children's aggregated responses were much closer to the base-rates, and 5-year-olds' judgments fell in between. This suggests that favoring representativeness at the expense of base-rates increases during the preschool years. But a remaining question is why children would start out favoring the seemingly more rational approach and then later settle on one that is less ideal? When considering this question, it is important to keep in mind that heuristic use often leads to accurate inferences, and therefore relying on them will produce a rational response much of the time. Recent adult work has examined the idea of "resource rational" reasoning, in which people appear to show an intuitive sense of the costs and benefits of deploying a fully analytic strategy versus a heuristic shortcut (Griffiths, Lieder, & Goodman, 2015; Lieder, Griffiths, Hsu, & Goodman, 2018a,b,). Though this possibility has not been investigated in children, it will be interesting to explore whether children are developing a similar sense and might be engaging in efficient strategy selection when relying on heuristics. In other words, older children might be employing strategies, like heuristic shortcut use, to allow them to reason more efficiently and to conserve cognitive resources.

Implications for improving inductive inference and mathematics in childhood

Finally, the developmental research on intuitive statistics and inductive reasoning has significant implications for later reasoning abilities. Future work should focus on these important connections.

Although we discussed the idea that heuristic use might be more rational than some have previously proposed, it is still important to intervene on the unequivocally errant applications of statistical concepts seen in older children and adults, such as the gambler's fallacy. Researchers have already begun examining whether the early intuitive principles that are present in the first year can be used to improve older children's general inductive reasoning based on applications of sampling principles. For example, in one study, preschool children were given training with the Ping-Pong ball paradigm in Figure 1, Panel B (Stanley & Lawson, 2014). Children were given a pre-test, assessing their ability to consider elements such as sample size and random sampling in real-world inductive inference problems (e.g., guessing which kind of cookie a child might get). They then received training with Ping-Pong balls: while drawing balls, the experimenter remarked on the random nature of the sampling process and on the correspondence in statistical properties between the larger distribution and the items drawn. Specifically the experimenter pointed out the correspondence between the randomly selected balls and the larger distribution. Post-training tests of additional inductive inference problems revealed that children in this training group were better able to consider statistical principles such as random sampling and sample size in their predictions than were children in a control group. This research is inspired by adult work, in which training on formal statistical concepts such as the law of large numbers has been shown to improve inductive inference (e.g., Fong, Krantz, & Nisbett, 1986). Importantly, in the adult work, the experimental paradigms are set up to assess how generally participants can apply the recently trained statistical concept. Similar assessments of the generalizability of trained statistical concepts should be implemented in future training paradigms with children.

We end with a few speculations on how these infant studies may inform mathematics education in older children. Proportional reasoning, probability, and statistics are difficult mathematical concepts to teach in schools (Bryant & Nunes, 2012). The main focus in the early years of mathematics curriculum is on learning about whole numbers via counting, addition, and subtraction. A number of researchers have argued that this whole-number focus may cause or exacerbate a “whole-number bias”, a tendency in young children to struggle with ratios and fractions because the well-learned principles governing whole numbers bias their reasoning (see Braithwaite & Siegler, 2018; Ni & Zhou, 2005; O’Grady & Xu, in press; Siegler, Thompson, & Schneider, 2011 Vamakoussi and Vosniadou, 2010). There is support for the idea that children’s familiarity with the rules of whole numbers negatively impacts their proportional and/or probabilistic reasoning in the developmental literature. For example, Huttenlocher and colleagues have shown that children perform worse on proportional reasoning tasks when the stimuli are discrete and countable versus continuous and uncountable, suggesting that counting leads them astray in these cases (Boyer, Levine, & Huttenlocher, 2008; Jeong, Levine, & Huttenlocher, 2007). Falk et. al (2012) also show that children apply erroneous subtraction strategies and erroneous comparisons of whole numbers in numerators when computing probabilities. These findings might be the result of an over-learning of counting principles and arithmetic, as they relate to whole numbers. These entrenched notions about how whole numbers work may also lead to difficulties in understanding that they cannot be rigidly applied to fractions and ratios.

The idea that children should first master whole numbers before other types of numbers is sensible, as one needs to start somewhere, and classic developmental literature suggests that children had no intuitions about proportions and probability until well into middle childhood. However, given the recent research with infants and non-human primates reviewed in this paper, it may be worth exploring the idea that in addition to positive integers, other types of numbers, including ratios, can be introduced earlier in education. Children appear to have intuitions about proportions and probability much earlier in development. The tasks used with infants and non-human primates, which present items visually in varying proportions, might be particularly good tools for introducing mathematical concepts such as proportions and probability. Further, the lottery-machine stimuli, which can be presented easily on computers, could be implemented as a game in which children can make predictions about future outcomes, as the numbers of different items change across trials. By introducing ratios, proportions, fractions, and decimals earlier in education, perhaps first in an intuitive manner and then more formally, children might become more flexible in their numerical reasoning. This practice could potentially help students harness the intuitive understanding they already have as babies, enrich their numerical reasoning abilities, and build a better foundation for learning high level mathematics later on.

References

Addessi, E., Crescimbene, L., & Visalberghi, E. (2008). Food and token quantity discrimination in capuchin monkeys (*Cebus apella*). *Animal Cognition 11*, 275–82.

Alvarez, G., & Franconeri, S. (2007). How many objects can you attentively track?: Evidence for a resource- limited tracking mechanism. *Journal of Vision*, *7(13)*,1–10.

Aslin, R. N. (2007). What's in a look? *Developmental Science, 10(1)*, 48–53.

Boyer, T.W., Levine, S. C., & Huttenlocher, J. (2008). Development of proportional reasoning: Where young children go wrong. *Developmental Psychology, 44*(5), 1478-1490.

Braithwaite D.W., Siegler R.S.. (2018). Developmental changes in the whole number bias. *Developmental Science*.

Brosnan S.F., & de Waal F.B.M. (2003). Monkeys reject unequal pay. *Nature 425,* 297-299.

Bryant, P., & Nunes, T. (2012). Children's understanding of probability: A literature review (full report). *A report to the Nuffield Foundation.* Retrieved on 1/1/2013 from:http://www.nuffieldfoundation.org/sites/default/files/files/Nuffield_CuP_FULL_REPORTv_FINAL.pdf

Cesana-Arlotti, N., Téglás, E., & Bonatti, L. (2012). The probable and the possible at 12 months: intuitive reasoning about the uncertain future. In F. Xu & T. Kushnir (Eds.), *Rational constructivism in cognitive development.* Advances in child development and behavior, Vol. 43. Waltham, MA: Academic Press.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition, 58(1),* 1 – 73.

Cronin K.A., & Snowdon C.T. (2008). The effects of unequal reward distributions on cooperative problem solving by cotton-top tamarins (*Saguinus oedipus*). *Animal Behavior, 75*, 245-257.

Denison, S., Reed, C., & Xu, F. (2013). The emergence of probabilistic reasoning in very young infants: Evidence from 4.5- and 6-month-olds. *Developmental Psychology, 49(2),* 243-249.

Denison, S., Trikutam, P., & Xu, F. (2014). Probability versus representativeness in infancy: Can infants use naïve physics to adjust population base rates in probabilistic inference? *Developmental Psychology, 50(8),* 2009 - 2019.

Denison, S., & Xu, F. (2014). The origins of probabilistic inference in human infants. *Cognition, 130(3)*, 335–347.

Denison, S., & Xu, F. (2010). Integrating physical constraints and statistical inference by 11-month-old infants. *Cognitive Science, 34(5)*, 885-908.

Denison, S., & Xu, F. (2010). Twelve- to fourteen-month-old infants can predict single-event probability with large set sizes. *Developmental Science, 13(5),* 798-803.

Diesendruck, G., Salzer, S., Kushnir, T. & Xu, F. (2015). When choices aren't personal: The effect of statistical and social cues on children's inferences about the scope of preferences. *Journal of Cognition and Development, 16(2),* 370 – 380.

Eckert, J., Call, J., Hermes, J., Herrmann, J. & Rakoczy, H. (2018a). Intuitive statistical inferences in chimpanzees and humans follow Weber's Law. *Cognition, (180)*, 99-107.

Eckert, J., Call, J. & Rakoczy, H. (2017). Are great apes able to reason from multi-item samples to populations of food items. *American Journal of Primatology*.

Eckert, J., Rakoczy, H., Call, J., Herrmann, E. & Hanus, D. (2018b). Chimpanzees consider humans´ psychological states when drawing statistical inferences. *Current Biology, 28(12),* 1959-1963.

Falk, R., Yudilevich-Assouline, P., & Elstein, A. (2012). Children's concept of probability as inferred from their binary choices – revisited. *Educational Studies in Mathematics, 81,* 207-233.

Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology (18),* 253-292.

Fontanari, L., Gonzalez, M., Vallortigara, G., & Girotto, V. (2014). Probabilistic cognition in two indigenous Mayan groups. *Proceedings of the National Academy of Sciences, 111(48),* 17075–17080.

Gigerenzer, G. (1991). How to Make Cognitive Illusions Disappear: Beyond "Heuristics and Biases", *European Review of Social Psychology, 2(1)*, 83-115.

Gigerenzer G., Hoffrage U. (1995). How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review, 102,* 684–704.

Garvin, L., & Woodward, A.L. (2015). Verbal framing of statistical evidence drives children's preference inferences. *Cognition, 138,* 35 – 48.

Girotto, V., Fontanari, L., Gonzalez, M., Vallortigara, G., & Blaye, A. (2016). Young children do not succeed in choice tasks that imply evaluating chances. *Cognition, 152,* 32 – 39.

Girotto, V., & González, M. (2008). Children's understanding of posterior probability. *Cognition, 106*, 325-344.

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science, 7,* 217-229.

Gualtieri, S. & Denison, S. (2018). The development of the representativeness heuristic in young children. *Journal of Experimental Child Psychology, 174,* 60 – 76.

Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences, 107(20),* 9066–9071.

Haith, M. M. (1998). Who put the cog in infant cognition: Is rich interpretation too costly. *Infant Behavior and Development, 21,* 167–179.

Harper D.G.C. (1982). Competitive foraging in mallards: ''Ideal free'' ducks. *Animal Behavior 30*, 575–584.

Hoffrage, U., & Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Academic Medicine, 73,* 538 – 540.

Jeong, Y., Levine, S.C., & Huttenlocher, J. (2007). The development of proportional reasoning: Effect of continuous vs. discrete quantities. *Journal of Cognition and Development, 8(2),* 237-256.

Kahneman, D. (2011). *Thinking, Fast and Slow.* Farrar, Straus, & Giroux. New York, NY.

Kahneman, D., & Tversky, A. (1971). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*, 430-454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237-251.

Kokis, J., Macpherson, R., Toplak, M., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology, 83*, 26−52.

Kushnir, T., Xu, F. & Wellman, H. (2010). Young children use statistical sampling to infer the preferences of others. *Psychological Science, 21*, 1134-1140.

Lawson, C.A., & Rakison, D.H. (2013). Expectations about single event probabilities in the first year of life: The influence of perceptual and statistical information. *Infancy, 18,* 961-982.

Lieder, F., Griffiths, T. L., Huys, Q. J. M., & Goodman, N. D. (2018a). The anchoring bias reflects rational use of cognitive resources. *Psychonomic Bulletin & Review 25(1)* 322-349.

Lieder, F., Griffiths, T. L., Huys, Q. J. M., & Goodman, N. D. (2018b). Empirical evidence for resource-rational anchoring and adjustment. *Psychonomic Bulletin & Review 25(2)* 775-784.

Lipton, J.S., & Spelke, E.S. (2003). Origins of number sense. Large-number discrimination in human infants. *Psychological Science, 14(5),* 386 – 401.

Ma, L. & Xu, F. (2011). Young children's use of statistical sampling evidence to infer the subjectivity of preferences. *Cognition 120(3),* 403-411.

McCrink, K., & Wynn, K. (2007). Ratio Abstraction by 6-Month-Old Infants. *Psychological Science, 18,* 740-746.

Ni, Y., & Zhou, Y-D. (2005). Teaching and Learning Fraction and Rational Numbers: The Origins and Implications of Whole Number Bias. *Educational Psychologist, 40(1)*, 27-52,

Placi, S., Eckert, J., Rakoczy, H. & Fischer, J. (in press). Long-tailed macaques (Macaca fascicularis) can use simple heuristics but fail at drawing statistical inferences from populations to samples. *Royal Society Open Science.*

Rakoczy, H., Clüver, A., Saucke, L., Stoffregen, N., Gräbener, A., Migura, J., & Call, J. (2014). Apes are intuitive statisticians. *Cognition, 131(1),* 60-68.

Rugani, R., Vallortigara, G., & Regolin, L. (2015). The use of proportion by young domestic chicks (*Gallus gallus*). *Animal Cognition, 18*, 605–16.

Siegler, R.S., Thompson, C.A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology, 62(4),* 273 – 296.

Stanley, S.L., & Lawson, C.A. (2014). The Impact of Statistical Training on Children's Inductive Reasoning. *Proceedings of the 36th annual conference of the Cognitive Science Society* (pp. 1521-1526). Austin, TX.

Stanovich, K.E., West, R.F., & Toplak, M.E. (2011). The complexity of developmental predictions from dual process models. *Developmental Review, 31(2),* 103 – 118.

Tecwyn, E., Denison, S., Messer, E., & Buchsbaum, D. (2017). Intuitive probabilistic inference in capuchin monkeys, *Animal Cognition, 20(2),* 243 - 256.

Téglás, E., Girotto, V., Gonzalez, M., & Bonatti, L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proceedings of the*

National Academy of Sciences of the United States of America, 104, 19156–19159.

Téglás, E., Ibanez-Lillo, A., Costa, A., Bonatti, L. (2015). Numerical representations and intuitions of probabilities at 12 months. *Developmental Science*, *2,* 183 – 193.

Téglás, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science, 332(6033),* 1054-1058.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185,* 1124-1131.

Vamvakoussi, X., & Vosniadou, S. (2010). Understanding the structure of the set of rational numbers: A conceptual change approach. *Learning and Instruction, 14*, 453 – 467.

Wellman, H.M., Kushnir, T., Xu, F., & Brink, K.A. (2016). Infants use statistical sampling to understand the Psychological world. *Infancy, 21(5),* 668 – 676.

Wilson, M. L., Hauser, M. D., & Wrangham, R. W. (2001). Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Animal Behavior 6:*1203–1216.

Xu, F., & Denison, S. (2009). Statistical inference and sensitivity to sampling in 11-month-old infants. *Cognition, 112*, 97–104.

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 5012–5015.

Xu, F. & Kushnir, T. (Eds.) (2012). Rational Constructivism in Cognitive Development.

Advances in Child Development and Behavior, Vol. 43. Waltham, MA: Academic Press.

Xu, F. & Kushnir, T. (2013) Infants are rational constructive learners. *Current Directions in Psychological Science*, *22*, 28-32.

Xu, F. & Spelke, E. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74(1),* B1 – B11.