Cognition 130 (2014) 335–347

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

The origins of probabilistic inference in human infants

Stephanie Denison [a,*], Fei Xu [b]

[a] University of Waterloo, Department of Psychology, Canada
[b] University of California, Berkeley, Department of Psychology, USA

ARTICLE INFO

Article history:
Received 31 January 2013
Revised 28 November 2013
Accepted 1 December 2013

Keywords:
Statistical inference
Infant learning
Probabilistic reasoning
Single-event probability
Action methodology

ABSTRACT

Reasoning under uncertainty is the bread and butter of everyday life. Many areas of psychology, from cognitive, developmental, social, to clinical, are interested in how individuals make inferences and decisions with incomplete information. The ability to reason under uncertainty necessarily involves probability computations, be they exact calculations or estimations. What are the developmental origins of probabilistic reasoning? Recent work has begun to examine whether infants and toddlers can compute probabilities; however, previous experiments have confounded quantity and probability—in most cases young human learners could have relied on simple comparisons of absolute quantities, as opposed to proportions, to succeed in these tasks. We present four experiments providing evidence that infants younger than 12 months show sensitivity to probabilities based on proportions. Furthermore, infants use this sensitivity to make predictions and fulfill their own desires, providing the first demonstration that even preverbal learners use probabilistic information to navigate the world. These results provide strong evidence for a rich quantitative and statistical reasoning system in infants.

1. Introduction

Reasoning under uncertainty pervades nearly every discipline of study, from social and natural sciences such as psychology, economics, biology and physics to law and medicine (Bell, Raiffa, & Tversky, 1988; Koehler & Harvey, 2004; Pauker & Kassirer, 1980). For example, in a volatile stock market, economists calculate the odds of making a profit by assuming rational economic laws and making educated guesses about how people's emotions may interfere with their judgments. In medicine, doctors are almost never certain of a patient's diagnosis upon initial assessment; all they have are symptoms that provide the basis of an estimate, e.g., the probability of a patient having a cold or lung cancer.

The current experiments ask where our probabilistic intuitions come from. Do untutored infants use probabilities to make predictions that guide their actions? Traditional developmental theory suggests that children do not become proficient at making inferences on even the most basic probabilistic reasoning problems until age 7 (Piaget & Inhelder, 1975). However, recent research indicates that children as young as 4 years of age are capable of engaging in rudimentary probability calculations when task demands are reduced (Acredolo, O'Connor, Banks & Horobin, 1989; Davies, 1965; Goldberg, 1966; Reyna & Brainerd, 1994; Yost, Siegel, & Andrews, 1962; Zhu & Gigerenzer, 2006). For example, in one experiment, children saw two collections of red and green marbles, one with a higher proportion of red marbles, the other with a higher proportion of green marbles and were asked which array they would prefer to draw from to obtain a red marble. With verbal demands minimized, by allowing children to point to a collection of marbles, 4-year-olds chose the collection with more red than green marbles at higher than chance levels (Yost et al., 1962). In addition, 5- to 7-year-olds can make quite sophisticated inferences about the likely outcomes of probabilistic events in a variety of contexts:

* Corresponding author. Tel.: +1 519 888 4567.
E-mail addresses: stephanie.denison@uwaterloo.ca (S. Denison), fei_xu@berkeley.edu (F. Xu).

0010-0277/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cognition.2013.12.001

Children make accurate probabilistic inferences in tasks involving complex judgments of expected values (Schlottmann & Anderson, 1994), and in tasks requiring the integration of prior probabilities with subsequent evidence (Denison, Bonawitz, Gopnik, & Griffiths, 2013; Girotto & Gonzalez, 2008; Gonzalez & Girotto, 2011).

Several recent studies have asked whether infants are capable of rudimentary probabilistic reasoning. First, in two looking-time experiments on single-event probability, 12-month-old infants were shown a computer screen displaying a lottery machine containing 3 yellow and 1 blue objects. The machine was briefly occluded and, on alternating trials, either a yellow or a blue object exited. Infants looked longer on trials when the blue object exited, suggesting that they had expected to see the more probable outcome (Teglas, Girotto, Gonzalez, & Bonatti, 2007; Teglas et al., 2011). In other experiments, 6- and 8-month-old infants were shown alternating samples of, for example, 4 red and 1 white Ping-Pong balls or 4 white and 1 red Ping-Pong balls being drawn from a large covered box. After each sampling event, the box was opened to reveal a population containing a ratio of 9 red to 1 white balls. Infants looked longer at the 4 white and 1 red ball sample (the less probable outcome) than the 1 white and 4 red ball sample (the more probable outcome) (Denison, Reed, & Xu, 2013; Xu & Garcia, 2008; see also Denison & Xu, 2010a; Xu & Denison, 2009 for evidence from 11-month-olds using variants of this method).

Unfortunately, all of these experiments have confounded probability and quantity, leaving unknown whether infants solve these problems using either proportional reasoning or a shortcut based on comparisons of quantities.[1] For example, in the lottery machine experiments, infants may have used a heuristic such as "if there are more yellow than blue objects, then it is more likely that a sample will consist of a yellow object than a blue object". A control experiment was conducted in which a barrier was placed in the lottery machine and the 3 yellow objects were above the barrier and the blue object was below, preventing the yellow objects from exiting the lottery machine. In this experiment, infants' looking times were reversed; they expected the blue object, rather than a yellow object to exit. This design rules out the concern that infants' looking behavior was driven by a simple perceptual preference for tracking and attending to the one blue object; however, it does not rule out the use of a quantity heuristic in the experimental condition. In the Ping-Pong ball experiments, infants may have thought, "If there are more red than white balls in the box, then a small sample should consist of more red than white balls". In other words, in all of these experiments, infants could have assumed that *more numerous equals more probable*, and in cases where only one population is present, this shortcut provides the correct answer.

In addition to these infant experiments, two studies using other methodologies – choice and property generalization – have tested slightly older toddlers. Denison and Xu (2010b), for example, tested 12- to 14-month-old infants' abilities to compute single-event probabilities in a choice task. They found that infants could predict which of two populations was most likely to yield a desirable object on a random draw. One population consisted of a distribution of 40 desirable to 10 undesirable objects (4:1) and the other contained the opposite distribution (1:4). Infants searched in the location that contained an object drawn from the 4 desirable to 1 undesirable distribution. This design also confounds absolute quantity of desirable objects with proportions; infants could have made predictions based on a simple comparison of 40 desirable objects in one population versus 10 desirable objects in the other population, or the relative quantities of desirable to undesirable objects within each population (4:1 versus 1:4). In a series of experiments investigating property generalization, 15-month-old infants demonstrated the impressive ability to use the probabilities of samples (e.g. 1 versus 3 yellow balls from a box with mostly blue balls and a few yellow ones) as the basis for generalizing a non-obvious internal property (Gweon, Tenenbaum, & Schulz, 2010). However, because there were many more blue balls than yellow balls in the box, these toddlers may have used the quantity heuristic to decide that drawing out 1 or 3 yellow balls was a low-probability event.

It is important to tease apart whether infants compute probabilities based on proportions, or use more straightforward comparisons of quantities for a variety of reasons. First, it is desirable to have a more precise account of how infants compute probabilities in previous experiments that claim to test probabilistic reasoning. Mathematically the concept of probability is instantiated by proportions, not simple comparisons of quantities (see Bryant & Nunes, 2012). In the probability literature with older children, researchers are careful to use methods that differentiate a full probability concept (based on proportional reasoning) from heuristics, which only yield the correct inferences some of the time (Falk, Yudilevich-Assouline, & Elstein, 2012). Second, young children succeed at a number of tasks that are not solvable with simple quantity comparisons, suggesting that there might be some foundation for probabilistic reasoning already in place. For example, in a variety of causal learning experiments, preschoolers are required to track the probability of objects or people causing particular events, and not just the frequency or absolute number of times that those objects or people are associated with certain events (e.g., Kushnir & Gopnik, 2007; Waismeyer, Meltzoff, & Gopnik, 2013). Additionally, outside of the lab, children are often faced with decisions that are best made using probability judgments and not straightforward quantity comparisons. For example, a child might want to track and compare the proportion of times that two caregivers agree to a request for ice cream, rather than simply tracking the number of times that each person has agreed, in order to efficiently decide which parent to approach in such situations. Third, evidence is accumulating to suggest that children make rational inferences in a number of cognitive domains that are consistent with the

[1] Although all of these tasks have confounded probability and quantity, other experiments investigating statistical learning in infancy have addressed frequency confounds in the auditory domain (Aslin, Saffran, & Newport, 1998). The primary aim of the current research is to investigate the origins of reasoning and decision-making under uncertainty, rather than statistical learning. The computations involved in statistical learning experiments (transitional probabilities) are likely quite different from those investigated here, thus a full discussion of these tasks is outside the scope of this paper.

general principles of Bayesian inference (Bonawitz, Denison, Chen, Gopnik, & Griffiths, 2011; Eaves & Shafto, 2012; Schulz, Bonawitz, & Griffiths, 2007). This claim requires that children represent beliefs in terms of probabilities, and an ability to make probabilistic inferences based on true proportional reasoning early in infancy would be consistent with such a claim.

Evidence for a true understanding of probability can be provided by experiments that test whether infants' computations involve estimates of proportions, not just simple subtraction as in the case of more or less. Imagine you are standing in front of two gumball machines and you have a preference for pink gumballs. Machine 1 has 50 pink gumballs and Machine 2 has 100. But, Machine 1 also has 100 black gumballs and Machine 2 has 500 black gumballs. Which machine should you put your quarter into? If you simply compare quantities of pink gumballs and subtract, you'll put your quarter into Machine 2, but if you reason about proportions and compare the ratios of pink to non-pink gumballs across machines, your odds are doubled if you select Machine 1.

The experiments that follow present 10- to 12-month-old infants with versions of the gumball problem, designed to examine whether infants use a quantity heuristic or proportional reasoning to make probabilistic inferences. Infants in all four experiments were shown two kinds of objects: one pink and attractive, the other black and unadorned. Infants crawled to one of the two objects to demonstrate their preference for one or the other. Then infants saw two jars containing different proportions of objects. The jars were then covered, and one object was randomly and invisibly removed from each jar and each was hidden in a separate location. The question is: Will infants correctly infer which population jar is more likely to yield a preferred object on a single, random draw, and search in the correct location?

If infants reason based on comparisons of the absolute number of their preferred objects in each jar rather than the proportions of preferred to other objects in each jar, then they will not search in the correct location at higher than chance levels in all four experiments. Specifically, Experiment 1 is designed such that there is an equal number of preferred objects in each population, but a different probability of obtaining a preferred object across jars (.75 versus .25). Experiment 2 directly pits proportional reasoning against comparisons of quantity, as infants' preferred object type is more numerous in the less probable population. Experiment 3 tests the possibility that infants use a different quantity heuristic in our tasks, namely that they compare quantities of dis-preferred objects, rather than proportions of preferred objects, to guide their choices. Finally, Experiment 4 tests the level of sophistication in infants' probability computations. Can infants determine which of two populations is most likely to yield an object they prefer when both populations contain a higher proportion of the ones they prefer but one population is more skewed than the other (.80 versus .60 preferred objects)?

Our task is designed to answer two other important questions: First, can infants under one year of age use their estimates of probabilities to guide their prediction and action? Previous work has used the standard looking-time methodology to study probability in infants below 12 months, which does not address questions of prediction and action. Second, can infants use their estimates of probabilities based on proportions to fulfill their own desires and wishes (e.g., getting a highly desirable pink object from a jar)? To our knowledge, the current experiments are the first to ask this question. If the answer is positive, it suggests that probabilistic reasoning in infancy is robust enough to provide a useful tool for navigating the world.

2. Experiment 1

In this experiment, we test whether infants can make an inference about where to search for a desired object type when the absolute number of preferred objects is equated across populations. On test trials, infants saw populations of 12 preferred to 4 other objects versus 12 preferred to 36 other objects.

2.1. Experiment 1 methods

2.1.1. Participants

Ten- to twelve-month-old infants participated in an action measure modeled after a procedure used by Feigenson, Carey, and Hauser (2002). Data from 24 infants were included (10 females; *M* = 11 months, 3 days; *Range* = 10 - months, 1 day–12 months, 30 days); 6 additional infants were tested and their data excluded due to failure to complete the preference trial (2), or the test trial (3), or parental interference (1). Parents in Experiments 1–3 were recruited by phone from the San Francisco Bay Area and infants received a small gift for their participation.

2.1.2. Materials

2.1.2.1. Objects. (132) "Lollipops" covered in construction paper were used as stimuli for the experiment. Lollipop-objects were used so that each object would have a stick, allowing infants to see that just one object was removed from each jar and to track the objects' locations during the sampling portion of the experiment (see below). Half of the lollipop tops were covered in black construction paper and half in pink construction paper to create two types of objects. Each pink lollipop had 3 small gold stars on each side.

2.1.2.2. Population Jars. (4) Cylindrical, transparent jars were used to hold the populations of objects (approx. 1320 cm^3 in volume). Four cylindrical covers made from construction paper were used as covers for the jars.

2.1.2.3. Cups. (2) Opaque cups (10 cm in diameter, 9 cm in height) were used to hold the samples removed from the population jars. Each cup had a cover with a small hole to allow the sticks to remain visible while the lollipops were in the cups and the covers were closed.

2.1.3. Procedure, design and predictions

All infants were tested individually in a forced-choice paradigm. Each infant sat with her parent across from an

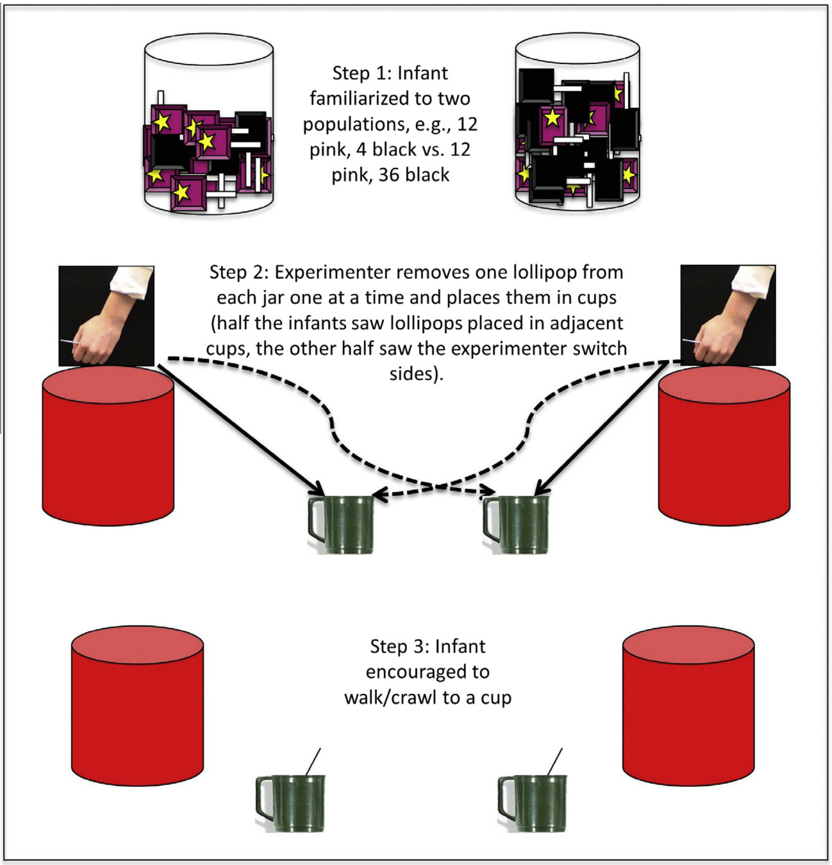

Fig. 1. Schematic representation of a test trial (adapted from Denison & Xu, 2010b).

experimenter. Parents were asked to hold their infant in front of them and to refrain from talking, pointing, or influencing their child in any way.

2.1.3.1. Preference trial. The experimenter brought out one pink lollipop and one black lollipop. She drew the infant's attention to each one, "See this one?". She then drew the infant's attention to both lollipops at the same time again and said, "Do you want to come pick one?" while placing them approximately 1 meter apart on the floor, equal distance from the infant. The parent was instructed to let go of the infant. Infants were encouraged to crawl or walk to an object of their choice ("Come get one!")—the color chosen was considered their preference. The experimenter clapped for the infant and said, "Good job; you found the one you like!"

2.1.3.2. Test trial. For infants preferring the pink object on the preference trial, the experimenter presented them with two transparent jars: one was filled with 12 pink and 4 black lollipops (3:1) and the other was filled with 12 pink and 36 black lollipops (1:3). Analogously, infants preferring the black object, saw one jar with 12 black and 4 pink lollipops and a second jar with 12 black and 36 pink lollipops. Thus the number of preferred objects was equated across the jars but the probability of drawing a preferred object differed between them.[2]

The general procedure of the test trial was the same for each infant (see Fig. 1). Half of the infants completed a standard test trial during which the experimenter brought out one set of the large covered transparent jars and placed them 1 meter apart on the floor. She also brought out two opaque cups and placed them next to the jars. She began the trial by removing the covers from both jars simultaneously to reveal the populations to the infant. Next, the experimenter drew the infant's attention to each jar by lifting the jar off the floor, shaking it, and rotating it around, always beginning with the jar on the right side. After placing each jar back on the floor, she simultaneously replaced the covers on the jars, closed her eyes and reached into the

[2] Does the perception of these populations reflect the actual numbers used? One might wonder whether the jar with 12 pink and 4 black lollipops appears to have a greater number of pink lollipops than the jar containing 12 pink and 36 black lollipops, due to the large number of black lollipops possibly occluding the pink lollipops in the 12:36 jar. We asked adults ($N = 12$) whether the 12:4 jar had more pink lollipops, the 12:36 jar had more pink lollipops or the jars had approximately equal numbers of pink lollipops (in counterbalanced order). Adults rated the jars as having equal numbers of pink lollipops at above chance levels. 9/12 participants rated the jars as equal, ($p = .007$, binomial test), 2 rated the 12:4 jar as having more pink lollipops and 1 rated the 12:36 jar as having more pink lollipops.

jar on the right. She pulled out one object such that the infant could see the stick but could not see the color of the object, as it was occluded by the experimenter's hand. She placed the object in the cup next to the jar and closed the cover on the cup. She repeated this action with the jar on the left and placed the object in the cup next to the jar. Finally, the experimenter lifted both cups simultaneously and said, "Come choose one!" She placed them back down and instructed the parent to let go of their infant. Once the infant crawled or walked to a cup the experiment was over.

The other half of the infants completed a test trial identical to the one described above except that when the experimenter removed the lollipops from the jars, instead of placing each lollipop into the cup next to the jar from which the object was removed, she placed them in the opposite cups.

While infants were making choices on both the preference and test trial, the experimenter looked directly at him/her, not directing her own attention to either object or cup. If the infant was hesitant or looked to the experimenter, she simply smiled, nodded her head up and down and said, "Yeah, you can have one!" For the more hesitant infants, the experimenter sometimes had to show the lollipops twice on the preference trial, or move them closer to the infant.

2.1.3.3. Design. Approximately half of the infants saw the pink object on the right during preference trials ($n = 13$) and half of the infants saw the mostly pink jar on the right during test trials. The placement of the two lollipops removed from the jars on the test trial was counterbalanced (next to the jar from which it was removed or the jar from which it was not removed).

2.1.3.4. Predictions. If infants use absolute quantity to guide their predictions, they should crawl to the two cups equally often because the two jars contained the same number of preferred objects. If infants instead use proportions to inform their decisions, they should choose the cup containing a sample from the population containing 12 preferred and 4 other objects (3:1).

2.2. Experiment 1 results and discussion

In all experiments, infants' choices on the preference and test trials were retrieved from the video records. For all experiments, two researchers coded the data; inter-rater reliability was 100%. Eighteen out of 24 (75%) infants selected the correct cup, reliably different from chance, binomial test, $p = .024$; 95% Confidence interval [53,90].[3]

This experiment was designed to address the question of whether infants rely on a quantity heuristic or proportional reasoning to make probabilistic inferences. Infants were able to make correct inferences when the distinguishing feature between the populations was the relative proportion of preferred to other objects, and not the number of preferred objects themselves, which were equated. These results suggest that infants can predict the outcome of a random sampling event using proportions when quantity is controlled. Experiment 2 presents infants with a more challenging task, where quantity and proportions are pitted against each other.

3. Experiment 2

In Experiment 2, we pitted absolute quantity and probability against each other by increasing the number of preferred objects in the less probable jar to exceed the number of preferred objects in the more probable jar.

3.1. Experiment 2 methods

3.1.1. Participants

Participants were 10- to 13-month-old infants. A total of 24 infants (17 females; *mean age* = 11 months, 16 days; *Range* = 10 months, 16 days–13 months, 7 days) completed the experiment and were included in data analyses. Nine additional infants were tested and excluded due to failure to complete the preference trial (4), or failure to complete the test trial (5).

3.1.2. Materials

Materials were identical to Experiment 1 except that the objects were altered to make the pink ones even more appealing. This change was made in an attempt to induce a more consistent and stronger preference for pink objects across all of the infants, so they will be more motivated to find the pink one on the test trial. We used brighter colored paper, silver stars and glitter for the pink object. The pink object for the preference trial had a small LED light in it that the experimenter switched on and off twice while showing infants the objects.

3.1.3. Procedure, design and predictions

The experimental procedure was the same as Experiment 1.[4] We used populations of 16 pink and 4 black lollipops (4:1) versus 24 pink and 96 black lollipops (1:4). We chose these ratios specifically because 10-month-old infants could perceive the difference between 16 and 24 pink lollipops, giving them the opportunity to make the inference based on absolute quantities, as infants this age discriminate 2:3 ratios (Xu & Arriaga, 2007). Infants should choose the sample from the 24:96 population if they compare absolute quantities of preferred objects (24 > 16) and the 16:4 population if they correctly estimate proportions (4:1 > 1:4).[5]

[3] For all experiments, ANOVAs revealed no differences in performance based on gender, whether the experimenter switched sides when placing the lollipops in the cups, whether the pink or black lollipop was on the right/left side during the preference trial, whether the mostly pink or mostly black jar was on the right/left side during test trials or whether the lollipop for the preference trial and the cup most likely to contain that color lollipop on the test trial were placed on the same side or opposite sides.

[4] Four infants were excluded for having a black lollipop preference. We only included infants with pink preferences in this experiment so that all infants would view the same populations.

[5] Exactly as in Experiment 1, 12 adults rated whether either jar had more pink lollipops than the other, or if they were equal. 8/12 participants rated the 24:96 jar as having more pink lollipops ($p = 0.035$, binomial test), with the other 4 participants rating them as having equal numbers of pink lollipops.

3.2. Experiment 2 results and discussion

Nineteen out of 24 (79%) infants selected the correct cup, reliably different from chance, binomial test, $p = .007$; 95% Confidence interval [58, 93]. We take this evidence as suggestive of infants using proportions and not comparisons of absolute quantity when making predictions about the outcome of a random draw, when these two strategies lead to opposing inferences. Together, we interpret the evidence from Experiments 1 and 2 to suggest that infants are able to use the proportions of preferred lollipops to guide their prediction and action.

However, two alternative interpretations of the data are still left open. First, infants could have used a different quantity heuristic to succeed at both of these tasks, which would not require proportional reasoning: they may have *avoided* the population containing a greater number of objects that were not selected as their preferred object-type on the preference trial. There are at least two possible motivations that could result in infants completing this task by avoiding the objects that they did not choose on the preference trial. The first, albeit unlikely, possibility is that infants selected their "preferred" object on the preference trial exclusively as a means of avoiding the object that they did not like, and then continued this strategy when completing the test trials. This seems unlikely as, if avoidance was the goal for the infants, a better course of action would be to not choose any object at all, which infants of this age regularly do (picture an infant sitting in a highchair with her mouth tightly closed, eyes pinched shut, and head shaking "no"). Infants in this experiment readily chose an object, and those who did not make a choice did not complete a test trial. A second possible motivation, which could also result in infants avoiding the unselected objects from the preference trial on the subsequent test trials, is that their choice on the preference trial biased them to develop a dis-preference for the unselected object-type on test trials. The experimenter praised the infant for choosing, for example, the pink object, and this might have driven down the value of the black objects, causing the infant to avoid those on test trials.

Second, recall that Experiments 1 and 2 were designed to prevent infants from succeeding at the tasks via simple quantity comparisons of their preferred objects across populations. However, one can argue that infants in Experiments 1 and 2 could have succeeded by making a series of quantity comparisons, without considering ratios or proportions. This requires a multi-step process, proceeding as follows: Infants first perceive the quantity of preferred objects in an individual jar and the quantity of dis-preferred objects in the same jar and then compare these two quantities. Next, infants repeat this process with the second jar. Then, infants apply the simple heuristic that if there are more preferred objects in an individual container, a preferred object is the "winner" and if there are more dis-preferred objects in a container, then the dis-preferred object is the "winner". This process would allow infants to choose between populations without comparing the proportions across populations. To illustrate, in Experiment 2 infants could begin by comparing the quantity of preferred objects to dis-preferred objects in the 16:4 population and mark a preferred object as the "winner". Then they could compare the quantity of preferred to dis-preferred objects in the 24:96 container and mark a dis-preferred object as the "winner". This shortcut would allow infants to make the correct inference without comparing the ratios/proportions across populations, as they then simply approach the sample from the container with a preferred object marked as the winner.

To address these concerns, we conducted a third experiment, in which comparisons of quantity of either preferred *or* dis-preferred objects would result in incorrect judgments, and only reasoning based on proportions of preferred objects could yield the correct judgment.

4. Experiment 3

In Experiment 3, infants must choose between a sample from one of two populations with three types of objects: preferred, dis-preferred and neutral. The neutral objects were lollipops covered in green construction paper and were undecorated. These objects should not have any valence associated with them a priori, as they are not included in the preference trial and thus should not be particularly preferred or dis-preferred.

4.1. Experiment 3 methods

4.1.1. Participants

Participants were 10- to 12-month-old infants. A total of 24 infants (14 females; *mean age* = 11 months, 8 days; *Range* = 10 months 8 days–12 months, 7 days) completed the experiment and were included in data analyses. Four additional infants were tested and excluded due to parental reports of non-typical development (2) and fussiness, which resulted in not completing the preference trial (2).

4.1.2. Materials

Materials were identical to Experiment 2, except that the set of green lollipops was included. Populations consisted of 8 preferred, 12 dis-preferred and 2 neutral lollipops versus 8 preferred, 8 dis-preferred and 64 neutral lollipops. In this experiment, we did not exclude infants with a black lollipop preference, reverting back to the design from Experiment 1. We did this because, despite the changes made to the lollipops to increase the salience of the pink ones, some infants in Experiment 2 still preferred the black lollipops and we were forced to exclude these infants in Experiment 2 because the populations required for infants with a black preference were not available. Rather than excluding infants based on preferences in Experiment 3, we created two sets of populations, as in Experiment 1. Thus, infants preferring pink objects saw populations of 8 pink, 12 black, and 2 green versus 8 pink, 8 black, and 64 green objects and infants preferring black objects saw populations containing 8 black, 12 pink, and 2 green versus 8 black, 8 pink and 64 green objects.

4.1.3. Procedure, design and predictions

Before completing the preference trial, infants were familiarized with three lollipops, one pink, one black and one green for approximately 1–2 min in the reception room. She held each one out individually to show the infant 2–3 times. This was done to allow infants to examine all three object-types, such that none would be completely novel and allowed them to see the experimenter handle all types of objects.

In this experiment, infants first chose between a black and a pink object on the preference trial. On test trials, they saw one population containing 8 preferred, 12 dis-preferred and 2 neutral objects and a second population containing 8 preferred, 8 dis-preferred and 64 neutral objects.[6] In this case, if infants are motivated by increasing their likelihood to obtain a preferred object, and they achieve this based on comparisons of proportions, then they should choose the sample drawn from the 8:12:2 population, as this population is more likely to yield a preferred object than the 8:8:64 population (8/22 versus 8/80). If infants are instead motivated by *avoiding* dis-preferred objects from the preference trials, either based on comparisons of absolute quantity (8 versus 12) or proportions (12/22 versus 8/80), then they should choose the sample drawn from the 8:8:64 population. Finally, if infants make choices based on comparisons of absolute quantity of preferred objects across populations (8 versus 8), then they should perform at chance. Thus, infants should only make the inference to search in the location containing a sample from the 8:12:2 location if they are (a) motivated by increasing the probability of obtaining a preferred object and (b) reasoning based on comparisons of proportions of preferred objects to all other objects across the populations (8/22 = .36 versus 8/80 = .10).

4.1.3.1. Post-test preference trial. Infants also completed a post-test preference trial, to ascertain whether infants in these experiments maintain a consistent preference between their choice on the first preference trial and their preferences at the conclusion of the experiment. On the test trial, infants were not allowed to see which color object was in the cup they chose, as the cup was taken away as soon as the infant made contact with it. Following this, the infant completed a second preference trial, wherein they chose again between the black and pink lollipops. The side that each object was on was counterbalanced.

4.2. Experiment 3 results

On the test trial, 18 out of 24 (75%) infants selected the correct cup, i.e., the cup containing a sample from the 8:12:2 population, reliably different from chance, binomial test, $p = .024$; 95% Confidence interval [53,90].

Twenty-one out of 23 infants maintained a consistent preference between the preference trial and preference post-test, binomial test, $p < .001$. One infant did not complete this trial due to experimenter error.

4.3. Experiment 3 discussion

Infants in Experiment 3 chose to search for an object in the cup that contained a sample from the 8:12:2 population. This suggests that infants did not attempt to avoid obtaining an object of the type that they had rejected in the preference trial but rather they attempted to obtain a preferred object. In this experiment, infants had to compare which of two populations contained a higher ratio of preferred objects, in a case where the preferred objects were always in the minority. Specifically, infants had to compare proportions in a jar containing 8/22 preferred objects versus a jar containing 8/80 preferred objects. Thus the concern affecting Experiments 1 and 2 does not, in principle, apply to this experiment. Infants could not use the shortcut of marking one population as having a preferred object as a "winner" and marking the other population as having a dis-preferred object as a "winner" and then simply approaching the sample from the container with a preferred object marked as "winner". In this experiment, infants were required to use proportional reasoning to determine that a draw of their preferred object was more probable in one population than the other, even though it was not highly probable in either case.

One may wonder if the infants really considered the green lollipops neutral in this experiment. Because the green objects were not present during the preference trial, we assumed that they were neutral to the infants – that infants likely did not prefer or dis-prefer them. Of course there is a chance that infants did not have neutral preferences for these green objects, but this is unlikely for a number of reasons. First, if infants did prefer or have positive feelings towards the green objects or were curious about them (perhaps based on their relative novelty), then they should have chosen the samples from the 8:8:64 population at greater than chance levels, which they did not. Additionally, experimental findings from the infancy literature suggest that infants are more likely to approach objects that an adult emotes positively towards than other objects in an array (Baldwin & Moses, 1994; Hornik, Risenhoover, & Gunnar, 1987). Infants in our experiments were praised for making a selection on the preference trials. The experimenter clapped for them, and said, "Yes, you found the one you like! It is really neat!" to help ensure that the infants continued to be motivated to try to obtain that object-type on the test trial. This positive emoting towards the preferred object should have increased its value even further, presumably above the un-praised green objects.

Second, we collected some additional data to investigate the status of the green objects. Fourteen 10- to 12-month-old infants completed a short experiment which began with a preference trial (between the pink and black objects) and then, in counterbalanced order, a trial that pitted their preferred object against the green object and a trial that pitted their dis-preferred object against the green object.[7] The experimental procedure of these two trials un-

[6] We again obtained ratings to determine whether adults perceive the number of pink lollipops as more numerous in one of the two jars, or as equal across jars. 11/12 adults rated the jars as having equal numbers of pink lollipops ($p < .001$, binomial test). 1 adult rated the 8 pink:12 black:2 green jar as having more pinks. We did the same with a different group of 12 adults for the black lollipops, using the same jars. 9/12 rated the 8 pink:12 black:2 green population as having more black lollipops than the other jar ($p = .007$, binomial test), the other three participants rated them as equal.

[7] Participants included in the reported data ranged in age from 10.2 to 12.4 months (Mean age = 11.31 months; 7 females). Two infants were excluded for either not choosing anything (1) or experimenter error (1). 13/14 infants preferred the pink object on the initial preference trial.

folded in the same way as the preference trial. When the green object was paired against the infant's preferred object, 11/14 infants chose their preferred object (significantly greater than chance, $p = .0574$). When the green object was paired against their dis-preferred object, 6/14 infants chose their dis-preferred object. These data suggest that the green objects were in fact neutral, and that infants were motivated to obtain their preferred object over the green object. Infants appeared to be more or less indifferent between the dis-preferred object and the green object. We take this to suggest that infants are likely neutral toward both the dis-preferred object and the green object, as infants may not dis-prefer the unselected object from the preference trial but simply prefer it less than the one they selected. Thus the "dis-preferred" objects might more aptly be referred to as simply unselected objects, but we will continue to refer to them as dis-preferred for the sake of simplicity.[8]

An additional concern regarding the results and interpretation of Experiment 3 involves whether infants can correctly perceive and represent the ratios in the competing populations. Although McCrink and Wynn (2007) found that 6-month-old infants can abstract and differentiate ratios of 4:1 versus 2:1, and our ratios of 8:14 versus 8:72 are more discrepant than this, the stimuli in their experiments only included two types of elements. No previous research has examined whether infants can abstract ratios and/or compare quantities when three sets of elements are present in a task with stimuli similar to ours. Although our data suggest that infants did represent and discriminate these ratios (as they performed above chance in the task), it is difficult to rule out alternative explanations of infants' performance without independent evidence to confirm that infants can perform the first step in the reasoning process required by the task (i.e., abstracting the ratios). There is also reason to question whether infants encoded that there was a larger quantity of dis-preferred objects, relative to preferred objects, in the 8:12:2 population, which further complicates the matter.[9] If the infants perceived these quantities as *equivalent*, then a version of the previously described heuristic can be applied to the design of this experiment. Specifically, infants might mark as a "winner" both the preferred and dis-preferred objects in the 8:12:2 population, whereas in the 8:8:64 population, infants will mark a neutral object as the "winner". If this strategy is applied, infants could choose a sample from the 8:12:2 population without comparing ratios, as the preferred object type is marked as one of the "winners" in one population and a different object (the neutral) is marked as the "winner" in the other population. The heuristic applies imperfectly here, as which object infants should expect to find in the 8:12:2 population if both a preferred and dis-preferred object are marked as "winners" is unclear (one might assume that, at best this would produce something close to chance responding). However, if infants were unsure about which object should be labeled as the "winner" in the 8:12:2 population but clearly marked the neutral object as the "winner" in the 8:8:64 population, they might use this to simply avoid the 8:8:64 population.

Due to the fact that it is unknown as to whether infants can abstract ratios when three types of elements are present and because these results could be explained by a modified version of a heuristic, we conducted one final experiment.

5. Experiment 4

Experiment 4 provides another strong test of whether infants compute probabilities via comparisons of proportions, or via comparisons of absolute quantities. In this experiment, we asked infants to reason about just two object types in the populations, to increase the likelihood that they would abstract and encode the correct ratios. We also used competing ratios that infants of this age are highly likely to discriminate: 4:1 versus 1.5:1, as previous research suggests that 6-month-old infants perceive the difference between ratios that vary by a factor of 2 when two types of elements are present (McCrink & Wynn, 2007). We showed infants one population containing 60 preferred and 15 neutral objects and another population containing 60 preferred and 40 neutral objects. This design ensures that: (a) the absolute number of preferred objects are equated across populations; (b) no dis-preferred objects are included, preventing infants from possibly avoiding dis-preferred objects; (c) drawing a preferred objects is the most likely outcome in both populations and (d) previous research suggests that infants are likely to discriminate these ratios. Therefore, if infants use the more straightforward heuristic described in the Discussions of Experiments 2 and 3, they will mark a preferred object as the "winner" in both populations, and will not be capable of choosing between them. Infants in this experiment must compare the ratios of the preferred to neutral objects to make the correct choice in this design.

5.1. Experiment 4 methods

5.1.1. Participants

Participants were 10- to 12-month-old infants. A total of 24 infants (9 females; *mean age* = 11 months, 12 days; *Range* = 10 months, 11 days–12 months, 28 days) com-

[8] Infants did trend toward choosing the green object over their "dis-preferred" object. This aligns with data from the literature on the origins of cognitive dissonance, which might predict that the dis-preferred objects from the preference trial hold a lower value to the child than the green objects (Egan, Bloom, & Santos, 2010; Egan, Santos, & Bloom, 2007). For example, 4-year-old children and capuchin monkeys rated a number of objects (stickers for the children, M&M's for the monkeys) multiple times until three objects were revealed to have nearly identical ratings in terms of preference. Then they were asked to choose between two of these objects, A and B. If Object A was chosen by the subject, when Object B was later pitted against the previously neutral Object C, both children and capuchins were more likely to choose Object C over the previously unchosen Object B (Egan et al., 2007). Similar results have been shown with adults in myriad experimental settings (Brehm, 1956; Jarcho, Berkman, & Lieberman, 2011; Lieberman, Ochsner, Gilbert, & Schacter, 2001; Lyubomirsky & Ross, 1999). Therefore, it is possible that the value of the unselected object in the preference trial was lower than the green object that was not present on the initial preference trial. If this is the case, then the green objects are "neutral" in the sense that they are less preferred than the object chosen on the preference trial but slightly more preferred than the unselected object.

[9] Findings from Xu and Arriaga (2007) suggest that infants should perceive this difference but the experimental set-up is very different in their experiments. It is possible that infants might not perceive this difference in our experiment, which contained a third set of elements.

pleted the experiment and were included in data analyses. Six additional infants were tested and excluded due to failure to complete the preference trial (2), failure to complete the test trial (3), or making an ambiguous choice on a test trial (1).[10] Infants in Experiment 4 were recruited via phone and email from the Kitchener-Waterloo region in Ontario, Canada.

5.1.2. Materials

Materials were identical to Experiment 3, except that the jars contained different ratios of lollipops. Infants who preferred the pink lollipop on the preference trial saw one jar with 60 pink: 15 green (4:1 ratio) lollipops and a second jar containing 60 pink : 40 green (1.5:1) lollipops. Infants preferring black saw one jar containing 60 black : 15 green lollipops and a second jar with 60 black : 40 green lollipops.

5.1.3. Procedure, design and predictions

Before completing the preference trial, infants were familiarized with the pink, black and green lollipops, as in Experiment 3. This time, the experimenter was more regimented with this familiarization. She ensured that the infant saw her hold each of the three objects individually three times and that the familiarization lasted for at least 90 s. This was done to ensure that the infants did not think that the green objects were less likely to be handled by the experimenter than the other two objects.

Just as in Experiments 1–3, infants completed a Preference trial with the pink and the black objects. On test trials, they saw one population containing 60 preferred and 15 neutral objects and a second population containing 60 preferred and 40 neutral objects.[11] In this case, if infants are motivated by increasing their likelihood to obtain a preferred object, and they achieve this based on comparisons of proportions, then they should choose the sample drawn from the 60:15 population, as this population is more likely to yield a preferred object than the 60:40 population. If infants make choices based on comparisons of absolute quantity of preferred objects across populations (60 versus 60), then they should perform at chance.

Infants completed a post-test preference trial identical to Experiment 3, to establish that they maintained a consistent preference between their choice on the first preference trial and at the conclusion of the experiment.

Counterbalancing was the same as in Experiment 3.

5.2. Experiment 4 Results

On the test trial, 17 out of 24 (71%, $SD = .46$) infants selected the correct cup, i.e., the cup containing a sample from the 60:15 population, marginally significantly different from chance, binomial test, $p = .064$; 95% Confidence interval [49, 87]. We also ran a logistic regression with Age as a continuous predictor variable and Response (coded as 0 for incorrect and 1 for correct) as the dependent variable. Infants' performance significantly improved with age (Odds Ratio = 7.772; $\beta_{Age} = 2.0506$, $SE = 1.0669$; Wald's $\chi^2 = 3.964$, $df = 1$, $p = .0546$).

Seventeen out of 22 infants maintained a consistent preference between the preference trial and preference post-test, binomial test, $p = .0169$. Two infants did not complete this trial due to experimenter error (1) and refusing to complete the trial (1).

5.3. Experiment 4 discussion

As a group, infants in Experiment 4 tended to search for an object in the cup that contained a sample from the 60:15 population. Therefore, infants appear to be engaging in a more sophisticated computation than the heuristic that could potentially explain the results of Experiments 1–3. In this experiment, infants were required to choose between two populations, which both contained a higher proportion of the preferred elements. Although performance in this experiment was only marginally different from chance using a binomial test, when the results of Experiments 3 and 4 are taken together, this provides compelling evidence for probabilistic reasoning based on proportions.

What might account for the difference in performance across age groups in this experiment, as well as the slightly weaker performance overall compared to Experiments 1 through 3? A number of possibilities are left open from the current data, and teasing them apart warrants future exploration. One possibility is that the younger infants in our sample (infants below approximately 11.5 months of age) are not yet capable of using proportions to compute probabilities, at least in our action task. The data from Experiment 3 suggest that this might not be the case as, arguably, infants in this experiment were required to compare proportions to compute probabilities and there were no age differences in this experiment. Nonetheless, due to the poorer performance of the younger infants in Experiment 4, and the potential alternative interpretation of Experiment 3, this possibility cannot be ruled out definitively. Another possibility, which could explain the slightly poorer overall performance in Experiment 4, is that infants were less motivated in general to search for a sample from the more probable population in this experiment, as opposed to the other three experiments. This seems plausible, as this is the only experiment in the series in which both populations had a greater than 50% probability of yielding a preferred object. One might surmise that determining precisely which of two populations affords a better chance of producing a desired object when both options provide a good opportunity to obtain the desired outcome is less motivating than the scenarios presented to infants in the other experiments.

6. Comparison of results across experiments

We assessed a number of factors collapsed across experiments to provide a more detailed breakdown of

---

[10] This infant stood up and turned around to face his parent after the experimenter placed the objects into the cups. As he turned back around, he stumbled forward and landed on top of the correct cup. The initial and second coder deemed this "choice" as clearly unintentional and as grounds for exclusion.

[11] We again obtained ratings to determine whether adults perceive the number of pink lollipops as more numerous in one of the two jars, or equal across jars. 12/12 adults rated the jars as having equal numbers of pink lollipops.

Table 1
Breakdown of infants' performance across Experiments 1–4.

Factor	# Of infants correct	Percent correct (%)	Binomial probability
Age split			
Younger (n = 48, M = 10.73mo.s)	35	73	$p = 0.002$
Older (n = 48, M = 11.91mo.s)	37	77	$p < 0.001$
Gender			
Male (n = 46)	35	76	$p < 0.001$
Female (n = 50)	37	74	$p < 0.001$
Color preference			
Pink preference (n = 78)	58	74	$p < 0.001$
Black preference (n = 18)	14	78	$p = 0.031$
Sampling procedure			
No Switch (n = 48)	36	75	$p < 0.001$
Switch (n = 48)	36	75	$p = 0.011$
Consistent side			
Same side (n = 47)	34	72	$p = 0.003$
Different side (n = 49)	38	78	$p < 0.001$

infant performance (see Table 1). Three factors related to the characteristics of the infants: We found above chance performance for both the older half of infants (mean age = 11.91) and the younger half of infants (mean age = 10.73), both the male infants and the female infants, and both the pink-preferring infants and the black-preferring infants. Two factors related to counterbalancing: We found above chance performance for both types of sampling procedures, ("no switch" trials, wherein the experimenter placed the sample in the cup adjacent to the population and "switch" trials, wherein the experimenter placed the sample in the cup adjacent to the opposite population), and for the side infants were required to approach to make correct choices (whether the correct cup for the test trial was presented on the same side as the preference trial or on the opposite side). Additionally, significance tests for the difference between independent proportions did not reveal any reliable differences in performance based on these factors (e.g., younger infants' overall score of 35/48 correct was not different from older infants' score of 37/48 correct).

7. General discussion

These findings provide strong evidence that infants below 12 month of age are capable of rudimentary probabilistic inference – infants can use proportions to predict the outcome of a single, random draw. Furthermore, unlike previous studies that used the looking-time methodology with young infants, where there was ambiguity in whether they could make predictions or just post-dictions after being given possible outcomes (Aslin, 2007), our task provides clear evidence that the format of infants' probabilistic computations is strong enough to support prediction and action. We also provide the first evidence from preverbal infants that they can use their sensitivity to probability based on proportions to fulfill their own desires and wishes, making it a useful tool for navigating the world. In our case, infants were able to estimate the probability of getting a preferred object and fulfill their desires by choosing to go to the correct location to find it. To our knowledge, this is the first demonstration of this sort with infants.

Taken together, the four experiments suggest that infants computed probabilities based on proportions and not alternative heuristics such as comparisons of quantities of preferred and/or dis-preferred objects across populations. It should be noted that one final factor was correlated with the more probable population in all of the experiments reported here: the more probable populations always contained the fewer total number of objects. Although predictions based on the total size of the population would have resulted in the same pattern of results in the current experiments, Denison and Xu (2010b) used populations that were equated in total numbers of objects, and infants made correct inferences in that experiment.

To our knowledge, these findings represent the first demonstration of reasoning about the probability of uncertain future events based on proportions, by children below seven years of age. A large collection of empirical studies from both cognitive psychology and education has examined how children analyze problems that hinge on probabilistic reasoning (e.g., Bryant & Nunes, 2012; Chapman, 1975; Davies, 1965; Falk et al., 2012; Goldberg, 1966; Piaget, 1975; Siegler, 1981). Falk and colleagues (2012) provide a thorough review for the interested reader. In their paper, they define a number of strategies through which children could perform probability computations and then designed tasks that diagnose whether 6- to 11-year-old children use simple rules, such as comparing target objects across sets, or true proportional reasoning, much as we have done here. Results of their experiments suggest that very young children often use shortcuts for probabilistic reasoning that can yield correct answers but do not demonstrate a true understanding of probability based on proportions. Specifically, their findings (and others) suggest that children begin to engage in correct probabilistic reasoning based on proportions at approximately 7–8 years of age (Acredolo, O'Connor, Banks & Horobin, 1989; Schlottmann & Anderson, 1994). However, other experimental findings suggest that children continue to rely on erroneous heuristics until approximately 10 years of age (Hoemann & Ross, 1980; Piaget, 1975).

How do we reconcile these findings with the results of our experiments? First, it should be noted that some experiments find evidence for 4-year-olds passing tasks that require similar computations to Denison and Xu (2010b) and Experiment 1 here, wherein proportion and absolute number of "target" items were either confounded or equated (e.g., Davies, 1965; Yost et al., 1962). In more difficult tasks that do not reveal competence until the ages of 7 to 11, the proportions presented to children were typically much more difficult than the ones we have presented to infants. For example, experiments that included trials wherein the population with a higher proportion of target objects had a lower absolute number of target objects as compared to the other population (similar to Experiment 2 here) had much smaller differences in ratio than our experiment, such as 2:3 versus 3:9 (Falk, Falk, & Levin, 1980). This is also the case in experiments that required children to reason about two proportions on the same side of 50%, asking children to compare, for example, 3:5 versus 1:4 or 3:1 versus 6:3, similarly to our Experiment 4 (Falk et al., 2012). Because we are constrained by the relatively coarse precision of infants' large number representations (Feigenson, Dehaene, & Spelke, 2004; Lipton & Spelke, 2003; McCrink & Wynn, 2007), we used ratios that were more extreme. Nonetheless, many of the papers on probabilistic reasoning in children indicate an increased use of the correct strategy with age. Our data also hint at a potential developmental difference between younger and older infants, as increasing age was positively related to performance in Experiment 4.

Other potentially important differences between our experiments and those undertaken with older children lie in the nature of the tasks. First, our tasks require that infants view two populations, and make an inference about which is most likely to yield a preferred object on a single draw. Most experiments with older children involve two populations or arrays of objects and the children are asked from which they would like to draw to obtain a particular element. Second, with infants, we intentionally avoided any verbal instruction (other than general encouragements such as, "go get one!" or "watch this"). In all experiments with older children, even those that deliberately reduced verbal demands on the child by allowing them to point, the experimenters still engaged in a large amount of explanation, and verbally requested responses from the children. Additionally, nearly all tasks with older children involved countable arrays of objects, as compared to our task, which uses very large numbers of objects presented somewhat briefly and a subject population that cannot yet count. Perhaps if a version of our task, which involves very little verbal instruction and encourages estimation over counting, was presented to older children, it would allow them to rely more heavily on their intuitions and succeed with more difficult computations. Experiments examining the highly related topic of proportional reasoning have revealed that intuitive problems are sometimes easier to solve (Ahl, Moore, & Dixon, 1992). Additionally, findings suggest that an "erroneous counting strategy" can lead children to make incorrect judgments in proportional reasoning tasks (Jeong, Levine, & Huttenlocher, 2007).

In our experiments, we have not yet teased apart whether infants computed proportions based on discrete or continuous variables, as all the lollipops were the same size. In studies on representations of number, researchers are very careful to distinguish whether infants or non-human animals use discrete (e.g., number of elements in a visual-spatial array or number of sounds in a sequence) or continuous quantities (e.g., the total area covered by all the elements in a visual-spatial array or the total duration of a sound sequence) in their computations. This, of course, is because only evidence for the former would constitute evidence for representations of number (e.g., Brannon, Abbott, & Lutz, 2004; Feigenson et al., 2004; Lipton & Spelke, 2003; Xu & Spelke, 2000). Estimating probabilities is interestingly different: one can estimate proportions using either discrete or continuous quantities. For example, if there are about 4 pink lollipops in a jar of about 20 total, then the proportion of pink ones is 0.2; thus if I make a single random draw from the jar, then the probability of drawing a pink one is also 0.2. Similarly, if the length of a straight line is about the length of my hand and a pink portion of the line is about the length of my middle finger, then the proportion of the pink segment is about 0.4; thus if I drop a small object on the line, the probability of it landing on the pink segment is 0.4. Although we cannot determine whether infants made the relevant computations in our experiments using discrete or continuous quantity representations, this does not detract from our finding that infants were sensitive to proportions and they used this information to estimate the probability of getting a desirable lollipop from the jar. Some studies with preschoolers show that computing continuous variables in a proportional reasoning task may be easier than computing discrete variables (e.g., Ahl et al., 1992; Boyer, Levine, & Huttenlocher, 2008; Jeong et al., 2007; Spinillo & Bryant, 1999). Future studies are needed to investigate whether infants are able to compute probabilities using both discrete and continuous variables.

A related question surrounds how infants represent and compute probabilities in our task and other probability experiments. The literature on infant numerical reasoning carefully examines the particular conditions that result in infants representing arrays of elements as analog magnitudes versus individual object files. The findings in this literature are complex, but the most prominent view is that infants' analog magnitude systems are engaged when they are presented with large numbers of relatively static objects and infants' object tracking systems are engaged when they are presented with small numbers (up to 3) of dynamically presented elements. In the former case, the output of the system is in the form of approximate quantity estimates, whereas in the latter case, the output of the system indirectly provides exact numerical information (see Feigenson et al., 2004; vanMarle, 2013 for evidence and reviews). We speculate that infants in our tasks, as well as the looking-time tasks involving large numbers of elements (e.g., Xu & Garcia, 2008) may represent and compare the quantities of objects in the populations using numerical representations outputted from the ANS. See McCrink and Wynn (2007) for a similar explanation of how infants might abstract ratios with large numbers of objects using the ANS. Infants in the lottery machine probability experiments, which involved small numbers of dynamically moving elements (e.g., Teglas et al., 2007) may instead use representations from object files to compare quantities or enumerate possible outcomes. Future work is required to pinpoint exactly how infants represent and compute probabilities in these tasks, perhaps by varying both the dynamic versus static nature of the stimuli and the numbers of objects in the populations and samples. The current data suggest that infants may be capable of using quantity information provided by either system as input to probability computations, depending on the stimuli. However, this cannot be definitively revealed without also determining whether infants in fact represent the elements as discrete and not continuous quantities.

Future empirical work will also investigate other aspects of probability understanding. For example, do infants and young children understand equivalence, e.g., that the probability of drawing out a pink lollipop is the same between a jar of 4 pink and 8 black lollipops and a second jar of 12 pink and 24 black lollipops? And do infants and young children have some rudimentary understanding of probability distributions? Another important question surrounds the age effect that surfaced in Experiment 4 – older infants outperformed younger infants in this experiment. It should be noted that this age effect occurs in only one of the four experiments reported here, and thus it could be a function of the particular ratios used. This is an empirical question that warrants future research. Are younger infants really incapable of estimating probabilities based on proportions, or do they simply require more discrepant ratios or ratios that do not pit two probable events against one another? Future research could pit improbable and highly improbable events against one another to address this question. Perhaps motivation in this case would be increased. Another intriguing possibility is that infants' probability estimations are undergoing significant development at approximately 10- to 11-months of age. Infants may start out using shortcuts or simpler strategies for estimating probabilities and only later begin making probabilistic inferences based on proportions. Future work can also explore whether younger infants will be more successful at difficult probability problems (such as those presented in Experiment 4) in looking time paradigms or in cases where task demands can be reduced or ratios are made more discrepant.

The present findings are important for understanding the origins of probabilistic reasoning. They argue against the traditional view of probabilistic reasoning, which suggested that children do not comprehend probabilities until middle- to late-childhood (Piaget, 1975). Recent research on probabilistic inference in infants has been interpreted to suggest that infants as young as 6 months have some intuitions about probability. However, none of these experiments teased apart whether infants made these judgments based on comparisons of proportions or simple more-or-less calculations. This is an important distinction to make because proportional reasoning is thought to be one of the hallmarks of a bona fide understanding of probability (Bryant & Nunes, 2012). If infants were only capable of making probabilistic inferences via the quantity heuristic (comparing the number of target items across populations) it would suggest that they do not have an understanding of probability but instead possess a heuristic that is useful but often inaccurate.

Relatedly, our findings add to a growing body of research suggesting that human reasoning may not be as irrational as once thought (Tenenbaum, Kemp, Griffiths, & Goodman, 2011; Tversky & Kahneman, 1974). In fact, it may be the case that human learners, along with other animals (Behrend & Bitterman, 1961; Yang & Shadlen, 2007) have an innate sensitivity for probabilities and the capacity for probabilistic inference. Though many dual-process frameworks of human cognition suggest that heuristic reasoning should be more prevalent early in ontogeny, the present findings suggest that, at least in tasks like ours, infants favor analytic processing over heuristics. As some have argued, use of heuristics may be a later-developing phenomenon – the accumulation of factual knowledge may be the source of these heuristics and they may indeed provide useful shortcuts in real life situations (see Kokis, Macpherson, Toplak, West, & Stanovich, 2002 for a review of dual-process theories as they relate to cognitive development).

The overall picture of early quantitative development has changed drastically in the last decade. Here we show that sensitivity to probability and the ability for carrying out probabilistic inferences are in the preverbal infants' repertoire. These early intuitions might lay the foundation for later development of mathematical thinking and reasoning under uncertainty. As Laplace put it, correctly and insightfully, "The theory of probabilities is at bottom nothing but common sense reduced to calculus; it enables us to appreciate with exactness that which accurate minds feel with a sort of instinct for which ofttimes they are unable to account." (Laplace, 1814).

Acknowledgments

A portion of this research was submitted as part of S. Denison's Dissertation research, and was supported by an NSERC graduate scholarship. We thank Melissa Nikkel, Christie Reed, Eduardo Colon, Pallavi Trikhutum and Nicole Newman for help with data collection, as well as Kathie Pham for secondary coding of data. This research was partially supported by a grant from NSERC (Canada) to F. Xu. Sincere thanks to Sylvia Yuan and Brian Waismeyer for comments on earlier drafts and to three anonymous reviewers for comments that significantly improved the manuscript. We also thank the parents and infants for their participation.

References

Acredolo, C., O'Connor, J., Banks, L., & Horobin, K. (1989). Children's ability to make probability estimates: Skills revealed through application of Anderson's functional measurement methodology. *Child Development*, 933–945.

Ahl, V. A., Moore, C. F., & Dixon, J. A. (1992). Development of intuitive and numerical proportional reasoning. *Cognitive Development, 7*(1), 81–108.

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

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by human infants. *Psychological Science, 9*, 321–324.

Baldwin, D. A., & Moses, L. J. (1994). Early understanding of referential intent and attentional focus: Evidence from language and emotion. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 133–156). Hillsdale: Lawrence Erlbaum Associates, Publishers.

Behrend, E. R., & Bitterman, M. E. (1961). Probability-matching in the fish. *The American Journal of Psychology, 74*, 542–551.

Bell, D. E., Raiffa, H., & Tversky, A. (1988). *Decision making: Descriptive, normative and prescriptive interactions*. New York, NY: Cambridge University Press.

Bonawitz, E. B., Denison, S., Chen, A., Gopnik, A., & Griffiths, T.L. (2011).A simple sequential algorithm for approximating Bayesian inference. In L. Carlson, C. Holscher, & T.F. Shipley (Eds.) *Proceedings of the 33rd annual conference of the Cognitive Science Society* (pp. 885–890).

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

Brannon, E. M., Abbott, S., & Lutz, D. J. (2004). Number bias for the discrimination of large visual sets in infancy. *Cognition, 93*(2), 59–68.

Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *The Journal of Abnormal and Social Psychology, 52*(3), 384.
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>.
Chapman, R. H. (1975). The development of children's understanding of proportions. *Child Development, 46*, 141–148.
Davies, C. M. (1965). Development of the probability concept in children. *Child Development, 36*(3), 779–788.
Denison, S., Bonawitz, E., Gopnik, A., & Griffiths, T. L. (2013). Rational variability in children's causal inferences: The Sampling Hypothesis. *Cognition, 126*(2), 285–300.
Denison, S., Reed, C., & Xu, F. (2013). Probabilistic reasoning in very young infants: Evidence from 4.5- and 6-month-olds. *Developmental Psychology, 49*(2), 243–249.
Denison, S., & Xu, F. (2010a). Integrating physical constraints in statistical inference by 11-month-old infants. *Cognitive Science, 34*, 885–908.
Denison, S., & Xu, F. (2010b). Twelve- to 14-month-old infants can predict single-event probability with large set sizes. *Developmental Science, 13*, 798–803.
Eaves, B., & Shafto, P. (2012). Unifying pedagogical reasoning and epistemic trust. In F. Xu & T. Kushnir (Eds.). *Rational constructivism in cognitive development. Advances in child development and behavior* (Vol. 43, pp. 295–320). Waltham, MA: Academic Press.
Egan, L. C., Bloom, P., & Santos, L. R. (2010). Choice-induced preferences in the absence of choice: Evidence from a blind two choice paradigm with young children and capuchin monkeys. *Journal of Experimental Social Psychology, 46*(1), 204–207.
Egan, L. C., Santos, L. R., & Bloom, P. (2007). The origins of cognitive dissonance evidence from children and monkeys. *Psychological Science, 18*(11), 978–983.
Falk, R., Falk, R., & Levin, I. (1980). A potential for learning probability in young children. *Educational Studies in Mathematics, 11*(2), 181–204.
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.
Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: Object files versus analog magnitudes. *Psychological Science, 13*(2), 150–156.
Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307–314.
Girotto, V., & Gonzalez, M. (2008). Children's understanding of posterior probability. *Cognition, 106*(1), 325–344.
Goldberg, S. (1966). Probability judgments of preschool children: Task conditions and performance. *Child Development, 37*, 158–167.
Gonzalez, M., & Girotto, V. (2011). Combinatorics and probability: Six- to ten-year-olds reliably predict whether a relation will occur. *Cognition, 120*, 372–379.
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.
Hoemann, H. W., & Ross, B. M. (1980). Children's concepts of chance and probability. In C. Brainerd (Ed.), *Children's logical and mathematical cognition: Progress in cognitive developmental research* (pp. 93–121). New York: Springer-Verlag.
Hornik, R., Risenhoover, N., & Gunnar, M. (1987). The effects of maternal positive, neutral, and negative affective communications on infant responses to new toys. *Child Development*, 937–944.
Jarcho, J. M., Berkman, E. T., & Lieberman, M. D. (2011). The neural basis of rationalization: Cognitive dissonance reduction during decision-making. *Social Cognitive and Affective Neuroscience, 6*(4), 460–467.
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.
Koehler, D. J., & Harvey, N. (Eds.). (2004). *Blackwell handbook of judgment and decision-making*. Malden: Blackwell Publishing.
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., & Gopnik, A. (2007). Conditional probability versus spatial contiguity in causal learning: Preschoolers use new contingency evidence to overcome prior spatial assumptions. *Developmental Psychology, 43*(1), 186–196.
Laplace, P. S. (1995). *Philosophical essay on probabilities, a translation of the fifth edition (1825). (A. I. Dale)*. New York: Springer (Original work published 1814).
Lieberman, M. D., Ochsner, K. N., Gilbert, D. T., & Schacter, D. L. (2001). Do amnesics exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychological Science, 12*(2), 135–140.
Lipton, J., & Spelke, E. (2003). Origins of number sense: Large number discrimination in 6-month-old infants. *Psychological Science, 14*, 396–401.
Lyubomirsky, S., & Ross, L. (1999). Changes in attractiveness of elected, rejected, and precluded alternatives: A comparison of happy and unhappy individuals. *Journal of personality and social psychology, 76*(6), 988.
McCrink, K., & Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychological Science, 18*, 740–746.
Pauker, S. G., & Kassirer, J. P. (1980). The threshold approach to clinical decision making. *New England Journal of Medicine, 302*(20), 1109–1117.
Piaget, J. & Inhelder, B. (1975). *The Origin of the Idea of Chance in Children*. Routledge and Kegan Paul. Translation of original work (1951).
Reyna, V. F., & Brainerd, C. J. (1994). The origins of probability judgment: A review of data and theories. In G. Wright & P. Ayton (Eds.), *Subjective Probability* (pp. 239–272). Wiley.
Schlottmann, A., & Anderson, N. H. (1994). Children's judgments of expected value. *Developmental Psychology, 30*(1), 56–66.
Schulz, L. E., Bonawitz, E. B., & Griffiths, T. L. (2007). Can being scared cause tummy aches? Naïve theories, ambiguous evidence, and preschoolers' causal inferences. *Developmental Psychology, 43*(5), 1124–1139.
Siegler, R. S. (1981). Developmental sequences within and between concepts. *Monographs of the Society for Research in Child Development, 46*(2), 1–84.
Spinillo, A. G., & Bryant, P. E. (1999). Proportional reasoning in young children: Part–part comparisons about continuous and discontinuous quantity. *Mathematical Cognition, 5*(2), 181–197.
Teglas, E., Girotto, V., Gonzalez, M., & Bonatti, L. 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.
Teglas, 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*, 1054–1058.
Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science, 331*, 1279–1285.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.
vanMarle, K. (2013). Infants use different mechanisms to make small and large number ordinal judgments. *Journal of Experimental Child Psychology, 114*, 102–110.
Waismeyer, A., Meltzoff, A.N., & Gopnik, A. (April, 2013). Causal learning in an uncertain world. In *Oral presentation at the 2013 Society for Child Development Biennial Meeting*. Seattle Washington.
Xu, F., & Arriaga, R. I. (2007). Number discrimination in 10-month-old infants. *British Journal of Developmental Psychology, 25*, 103–108.
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., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*, B1–B11.
Yang, T., & Shadlen, M. N. (2007). Probabilistic reasoning by neurons. *Nature, 447*(7148), 1075–1080.
Yost, P. A., Siegel, A. E., & Andrews, J. M. (1962). Nonverbal probability judgments by young children. *Child Development, 33*, 769–780.
Zhu, L., & Gigerenzer, G. (2006). Children can solve Bayesian problems: The role of representation in mental computation. *Cognition, 98*, 287–308.