

Probability Versus Representativeness in Infancy: Can Infants Use Naïve Physics to Adjust Population Base Rates in Probabilistic Inference?

Stephanie Denison
University of Waterloo

Pallavi Trikutam and Fei Xu
University of California, Berkeley

A rich tradition in developmental psychology explores physical reasoning in infancy. However, no research to date has investigated whether infants can reason about physical objects that behave probabilistically, rather than deterministically. Physical events are often quite variable, in that similar-looking objects can be placed in similar contexts with different outcomes. Can infants rapidly acquire probabilistic physical knowledge, such as *some leaves fall* and *some glasses break* by simply observing the statistical regularity with which objects behave and apply that knowledge in subsequent reasoning? We taught 11-month-old infants physical constraints on objects and asked them to reason about the probability of different outcomes when objects were drawn from a large distribution. Infants could have reasoned either by using the perceptual similarity between the samples and larger distributions or by applying physical rules to adjust base rates and estimate the probabilities. Infants learned the physical constraints quickly and used them to estimate probabilities, rather than relying on similarity, a version of the representativeness heuristic. These results indicate that infants can rapidly and flexibly acquire physical knowledge about objects following very brief exposure and apply it in subsequent reasoning.

Keywords: cognitive development, infant cognition, learning, physical reasoning, probability

The actions and interactions of physical objects are incredibly variable. Imagine a typical Saturday in the life of a toddler. The day begins with a trip to the grocery store, where she bumps into a table, knocking over a display of jam. As you move her away from the broken glass, she notices that some of the jars have shattered completely, and others did not even crack. When you sit down for lunch later in the day, she playfully throws a handful of spaghetti against the kitchen wall. She sees that some of the strands have stuck to the wall, and others have fallen to the floor. After lunch, you go for a walk outside, and she sees the leaves on the trees rustling in the wind. She notices that some of the leaves flutter to the ground, and others remain on the branches. How does she represent and reason about these physical events? What does she think about the rules that govern the objects involved in these interactions,

and what does she think will happen the next time she encounters them in a different situation? Understanding the nuances of physical events and making predictions about how objects will behave in the future is an incredibly difficult task. How do humans come to understand the stochastic and highly unpredictable behavior of physical objects in our environment?

Extensive research from the last two decades documents the development of physical reasoning in infancy. By all accounts, the experimental findings suggest that infants are extraordinarily sophisticated naïve physicists. This body of work has centered on two aspects of physical reasoning: First, a wealth of experiments has examined and confirmed that very young infants expect objects to behave in accord with the core principle of persistence—that objects continue to exist as they are, in space and time (Baillargeon, 2008). Encompassed in this principle are a variety of rules that infants expect physical objects to obey, including principles of cohesion (objects do not spontaneously break apart), boundedness (objects do not spontaneously merge together), continuity (objects move on continuous paths through space and time), and solidity (objects cannot pass through other solid objects; Aguiar & Baillargeon, 1999; Baillargeon, Spelke, & Wasserman, 1985; Luo & Baillargeon, 2005; Newcombe, Huttenlocher, & Learmonth, 1999; Simon, Hespos, & Rochat, 1995; Spelke, Breinlinger, Macomber, & Jacobson, 1992). A second, and equally impressive, body of research explores what infants understand about the relationships between objects in different event categories, perhaps most notably, occlusion, containment, and covering (see Baillargeon, Li, Ng, & Yuan, 2009, for a review). Research on infants' comprehension of event categories has focused on the ages at which they begin to include an assortment of relevant variables in their representations of physical events. For example, it has been

This article was published Online First June 16, 2014.

Stephanie Denison, Department of Psychology, University of Waterloo; Pallavi Trikutam and Fei Xu, Department of Psychology, University of California, Berkeley.

We thank the infants and parents for participating in the experiments. We also thank Justine Hoch, Kathie Pham, Eduardo Colon, and Madeleine Hanson for their assistance with data collection and recruitment. Stephanie Denison was supported by a graduate fellowship from the Natural Sciences and Engineering Research Council of Canada (NSERC) while completing data collection.

Correspondence concerning this article should be addressed to Stephanie Denison, Department of Psychology, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada. E-mail: stephanie.denison@uwaterloo.ca

documented that infants include height as a relevant variable in occlusion events at 3.5 months of age, but they do not include height as a relevant variable in containment events until 7.5 months of age (Baillargeon & DeVos, 1991; Hespos & Baillargeon, 2001, 2006).

In general, applying the principle of persistence and identifying relevant variables such as height and width result in physical rules that are deterministic and reliable—they do not vary. A basketball will never pass through another basketball, and a short fence will never hide a tall tree. These are hard and fast rules of physical reasoning. However, many aspects of the physical world are variable: *Some* glasses break when knocked over, and *some* strands of spaghetti stick to walls if thrown by a toddler; as adults, we have intuitions about the conditions under which these objects are likely to break or stick and in what proportions. To our knowledge, no research has examined the origins of our sensitivity to such variable physical events. Are infants sensitive to the regularities of objects' behaviors when physical events occur nondeterministically (i.e., probabilistically) with sets of otherwise identical objects?

It has been suggested that the principle of persistence may be part of infants' innate conceptual endowment (Baillargeon, 2008; Spelke et al. 1992), whereas infants must learn which are the relevant variables to consider when reasoning about different event categories (Baillargeon & Carey, 2012; Baillargeon et al. 2009). These theoretical accounts leave open the question of how humans come to know the less obvious probabilistic constraints of the physical environment, those that do not necessarily adhere to regular and reliable laws. Learning these variable constraints is further complicated by the fact that the objects involved (e.g., the spaghetti, the glass jars, the leaves) can be nearly identical in surface appearance and undergo nearly identical physical interactions, yet the outcomes are inconsistent. If these complex physical constraints are acquired through experience, how are they learned? One possibility is that this knowledge is only learnable through explicit, verbal means (e.g., a parent tells a child, "Pasta is sticky. If we throw some spaghetti against the wall, some will stick but not all of it."). Evidence for such a hypothesis could be revealed if infants are unable to reason about objects that behave stochastically before acquiring complex linguistic knowledge. In other words, infants might be unable to apply a probabilistic physical constraint without representing this knowledge linguistically. Another possibility is that infants are sensitive to the statistical input of these variable constraints, allowing them to flexibly learn whether physical properties should apply to all objects in a set or to only a subset of the objects. This second possibility hinges on infants' ability to accurately track observed patterns and statistical regularities in the environment, and their abilities to extract probabilistic constraints and make predictions based on this input.

A great deal of research suggests that infants are incredibly sensitive to statistical regularities in their environments (e.g., Aslin, Saffran, & Newport, 1998; Gerken, 2006; Gómez, 2002; Gopnik et al., 2004; Marcus, Vijayan, Rao, & Vishton, 1999; Saffran, Aslin, & Newport, 1996; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002; Sobel & Kirkham, 2006). Recent research also suggests that 6- to 15-month-old infants are capable of making rudimentary probabilistic inferences (Denison

& Xu, 2010a, 2010b; Denison, Reed, & Xu, 2013; Gweon, Tenenbaum, & Shulz, 2010; Téglás, Giroto, Gonzalez, & Bonatti, 2007; Téglás et al., 2011; Xu & Denison, 2009; Xu & Garcia, 2008; see Denison & Xu, 2012, for a review). For example, when infants see a sample of, say, four red balls and one white ball being drawn from a large covered box, they expect the box to contain mostly red balls and just a few white balls, rather than mostly white balls and just a few red balls. Infants also appear to integrate substantive domain knowledge in their probabilistic inferences. Most notably, for the purpose of the present experiments, evidence from two looking-time studies reveals that infants can correctly apply their physical knowledge when making probabilistic inferences. Téglás and colleagues (Téglás et al. 2007, 2011) found that 12-month-old infants could integrate knowledge of solidity and complex spatiotemporal information with distributional information when making probabilistic inferences. In a different series of experiments, Denison and Xu (2010a) found that 11-month-old infants could integrate a cohesion constraint in probabilistic inference. In one of these experiments, infants were asked to reason about a large box containing 50% green balls, 40% red balls, and 10% yellow balls. They were shown that the green balls could not be removed from the box. Results showed that infants were able to adjust the base rate of the distribution of balls by excluding all green balls and compute the probability of different samples based on the proportions of the red and yellow balls.

Although each of these experiments suggests that infants can consider physical constraints on objects and apply them in probabilistic inference, no experiment has examined physical constraints that are probabilistic, that is, rules that only apply to a subset of the objects of interest. In the present series of experiments, we took both the research on infant physical reasoning and the research on infant probabilistic inference in new directions. We conducted experiments to examine whether infants can flexibly learn probabilistic or deterministic physical constraints, based solely on the observed statistical regularities of the objects within the sets, and apply those constraints to guide their subsequent expectations of event outcomes. The design of the experiments does not allow infants to learn these physical constraints based on the surface features of the objects, as all items within the sets appear identical. To our knowledge, this is the first study examining whether infants can acquire probabilistic constraints when reasoning about physical events.

The experiments also provide a strong test of whether infants reason using either heuristics or estimations of probabilities in their inferences. In previous experiments on probabilistic inference, it was difficult to discern which algorithm underpinned infants' performance. For example, in Xu and Garcia (2008), infants watched as an experimenter removed a random sample of either one white and four red balls or one red and four white balls from a large covered box on alternating trials. The box was revealed to contain mostly red balls, and infants looked longer at the one red and four white balls sample, suggesting that they found this outcome surprising or unexpected. However, at least one of two reasoning processes could have led to this pattern of looking times: (a) Infants may have estimated the probability of randomly obtaining the particular samples from a box containing 90% red balls, or (b) they may have reasoned about the sampling events based on a well-known heuristic, namely, representative-

ness. Tversky and Kahneman (1974) defined the representativeness heuristic as using the similarity between the surface features of a sample and population in reasoning (i.e., the distribution of objects in a small sample should resemble the distribution of objects in the larger populations from which they are drawn).¹ Adults exhibit this behavior in a variety of contexts. Thus, infants in previous probabilistic inference experiments may have simply compared the surface appearance of the ratios in the samples with the ratio in the population and concluded that the perceptually more congruent sample was more probable. As with all heuristics, this provides a fairly reliable shortcut but also results in incorrect inferences under some conditions. In the experiments that follow, the use of the representativeness heuristic is directly pitted against true probability estimations.

We report three experiments with 11-month-old infants using a violation of expectation (VOE) looking-time method. In the first two experiments, infants were shown boxes containing sets of green and red balls and taught that a physical constraint applies probabilistically to the green balls, which are all identical in appearance (i.e., most, but not all, balls are immobile—they cannot be removed from boxes and containers). They were also shown that all red balls moved freely. On the test trials, an experimenter showed the infants a large opaque box, removed a sample from it (e.g., one red and four green balls), placed them in a container, and then revealed the contents of the box as consisting of a ratio of three green balls to one red ball. On an alternating test trial, infants were shown this same series of events, except that the experimenter removed a sample of one green ball and four red balls. In this case, if infants are sensitive to the physical constraint, and they integrate this constraint with the distribution of balls in the population box, they should look longer at the one red and four green balls sample, as this is an improbable event. If infants cannot learn a probabilistic physical constraint, or if they make inferences based on the similarity in appearance between the population and the samples, they should look longer at the one green and four red balls sample, as this sample has nearly the exact opposite proportions compared with the population box (see Figures 1 and 2A). This design pits probabilistic inference directly against representativeness, as the probable sample is less congruent in appearance to the population box than the improbable sample. Thus, if infants view the scene and apply the heuristic that random samples should look similar to populations, they would incorrectly look longer at the more probable (but perceptually incongruent) sample. If they instead use the physical constraint to adjust the base rate of balls in the population box, they would correctly look longer at the less probable (but perceptually congruent) sample.

In the third experiment, infants were shown a deterministic physical constraint—all green balls were immobile, and all red balls were mobile. Do infants understand, in the context of this task, that some physical constraints do in fact apply deterministically, with no allowable exceptions? Can they apply this constraint to subsequently reason about the probability of different event outcomes? In this case, if infants correctly make inferences based on the physical constraint that no green balls can move, then they should find both the one-red-and-four-green sample and the one-green-and-four-red sample unexpected (see Figure 2B).

Experiment 1

Method

Participants. Twelve 11-month-old infants participated (nine males; $M = 11.33$ months; age range = 10.6–11.99 months). Two additional infants were tested and their data excluded, one due to parental interference and one to infant fussiness. Legal guardians of all infants provided written consent for their child's participation. The guardians of the infants in all experiments were recruited from the San Francisco Bay area, and infants were given a small prize for participating.

Apparatus. For a detailed description of the apparatus, see Denison et al. (2013). Infants in all experiments participated in a VOE looking-time paradigm. Infants sat in a high chair approximately 70 cm from the center of a stage. The parent sat next to the infant facing the opposite direction and was instructed to avoid looking at the stage or interfering with the infant in any way. The experimenter oriented the infant to the outer limits of the viewable area of the stage while the observer, located behind a curtain not visible to the infant, watched the infant's eyes on the TV screen. The experimenter sat behind the small stage to show infants the stimuli. The observer coded looking times using jHab, Java Habituation software (Casstevens, 2007).

Materials. The materials were identical across experiments.

Ping-Pong balls. A total of 168 (84 green and 84 red) balls were used. The green balls had six Velcro strips glued to them (approximately 0.8–1.5 cm); the red balls had no Velcro.

Boxes and containers.

Free play box. A small white box (17.5 cm × 17.5 cm × 8 cm) without a top, containing four green and four red Ping-Pong balls, was used during the free play phase. Three of the four green balls were stuck to the bottom of the container; all the red balls were loose.

Sample container. A Plexiglas container (20 cm × 4.5 cm × 4 cm) was placed at the front left-hand corner of the stage at the beginning of the experiment to display the five-ball samples removed during test trials.

Demonstration containers. Two Plexiglas containers (28.5 cm × 4.5 cm × 4 cm) were used during Demonstration Phase 1. One container held eight red balls, the other eight green balls. Six

¹ The question of whether infants engage in heuristic or analytical processing is separate from the concern that infants in previous probability experiments might not have considered probability or random sampling at all. For example, one might surmise that infants would look longer at a display containing a collection of one red ball and four white balls paired with a mostly red box than a display containing a collection of one white ball and four red balls paired with a mostly red box, simply because the mismatched ratios are more interesting to look at or require more processing time. Xu and Garcia (2008) ruled out this interpretation by presenting infants in control experiments with these samples and populations but removing the random sampling element. They found that when the experimenter removed these samples from her pocket, rather than the population box, infants did not look longer at the one-red-ball-and-four-white-balls sample paired with the mostly red box. Therefore, it is known that infants are reasoning about random sampling in these experiments, but it is unknown whether they are reasoning using the shortcut that samples should resemble populations in perceptual appearance or using true probability estimations.

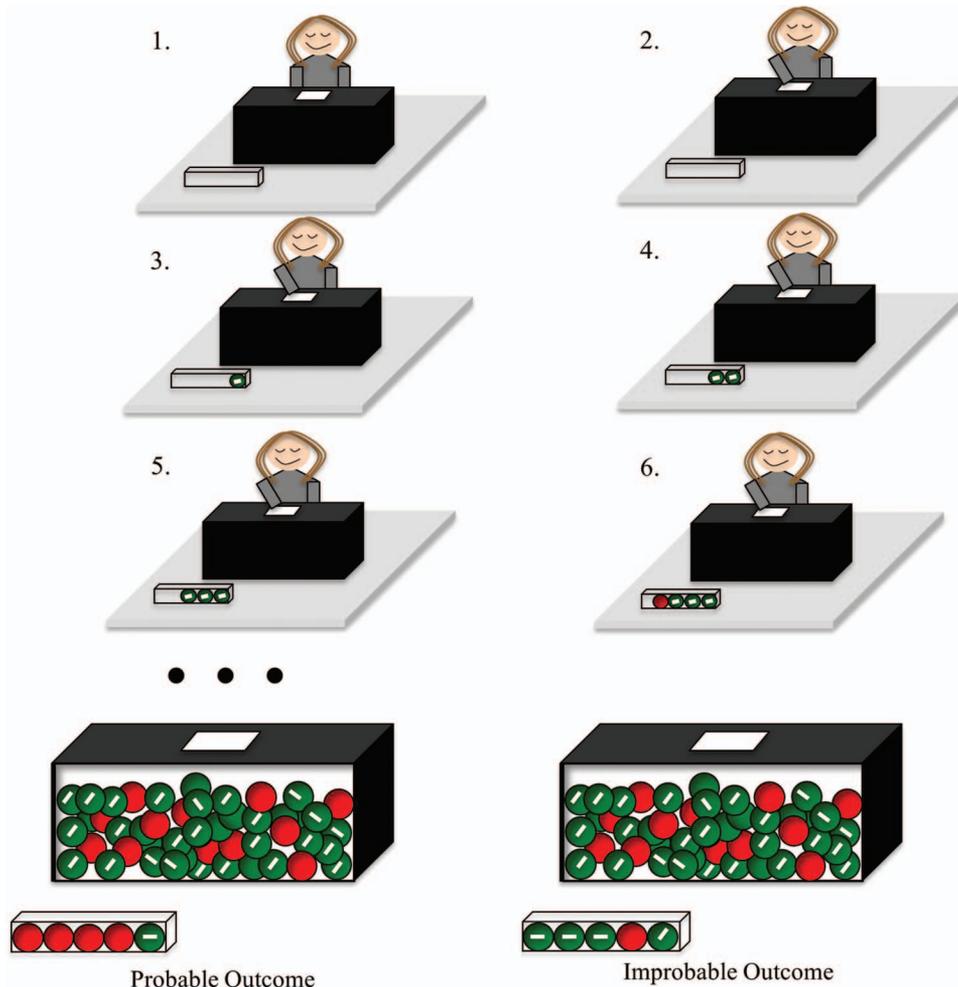


Figure 1. Schematic representation of a test trial. See the online article for a color version of this figure.

green balls were glued to the container (balls in Positions 1, 3, 4, 6, 7, and 8).

Demonstration box. A small white box (27 cm × 16 cm × 13 cm) was used in Demonstration Phase 2. The top of the box had a cutout for the experimenter to reach in to access the balls, and the front surface of the box was replaced with Plexiglas to provide a transparent front. Affixed to the top of the front of the box was a purple curtain that could be raised and lowered to hide or reveal the contents of the box. This box contained 14 Ping-Pong balls: seven green and seven red. Six of the green balls were glued to the inside surfaces of the box.

Population box. A large box (39 cm × 34 cm × 22 cm) was used to display the population on the familiarization and test trials. The box was a white rectangular cube, with the inside divided into three parts: Two Plexiglas containers were inserted into the front and back of the box, each containing 60 Ping-Pong balls, and a hidden center compartment used to hold the samples to be removed from the box during test trials. The “mostly red” side of the box contained 45 red and 15 green balls (ratio of red:green = 3:1); the “mostly green” side contained the opposite ratio (ratio of red:green = 1:3). When viewed from the front or the back, the box

appeared to be one large box filled with Ping-Pong balls. The Plexiglas on each side was covered with black fabric curtains (secured to the top of the box with Velcro) that could be lifted to reveal the contents of the box through the window. From the perspective of the infants, there were two boxes, which appeared identical when the curtains were lowered but contained different populations. The top of the box had a cutout that allowed the experimenter to reach into the center compartment of the box.

Procedure.

Free play phase. Infants began by playing with the balls in the free play box to familiarize them with the objects that they would observe throughout the session. The experimenter picked up the balls and encouraged the infants to play with balls of each color, giving them the opportunity to notice that the red balls were easily removed and most of the green balls (three out of four) were stuck.

Demonstration Phase 1. The experimenter went behind the stage, where she remained for the duration of the experiment. She brought out one of the two display containers, in counterbalanced order. For the container with red balls, she picked up each ball and then put it back in the container, starting from the leftmost ball. For the container with green balls, she picked up the movable balls and

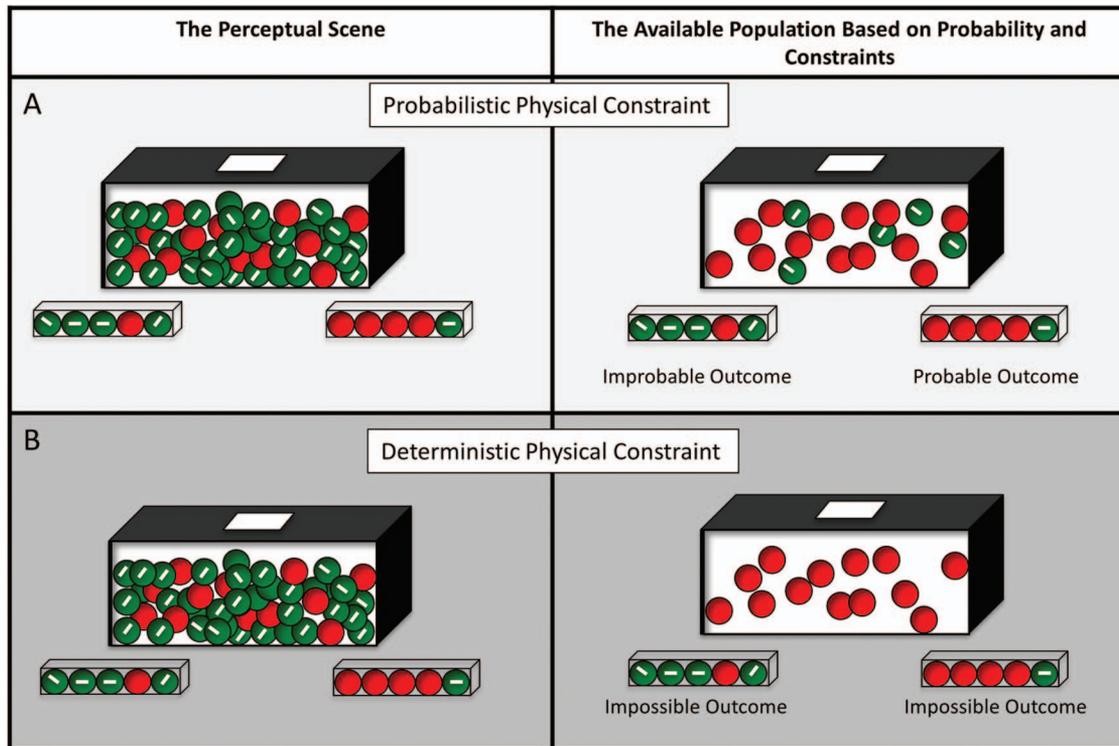


Figure 2. A: Illustration of predictions for Experiments 1 and 2 (probabilistic constraint). B: Illustration of predictions for Experiment 3 (deterministic constraint). See the online article for a color version of this figure.

placed them back in the container, just as she did with the red balls, but when she got to a stuck green ball, she attempted to lift it, and it caused the entire container to move.

Familiarization trials (four trials). On each trial, the experimenter placed the population box on the stage with the front curtain closed, placing either the mostly red or mostly green side facing the infant, in counterbalanced order. The experimenter shook the box back and forth a few times, saying, “What’s in the box?” Then, with one hand, she lifted the front cover of the box, and with the other, she simultaneously lowered the backdrop of the stage while saying, “Look, [baby’s name], look!” The observer began timing upon hearing the second “look,” as this was when the population of either mostly red or mostly green balls became visible and the experimenter became concealed. The trial ended when the infant looked away for 2 consecutive seconds. The four trials alternated between the mostly green and the mostly red populations, thus each infant saw each side twice.

Demonstration Phase 2. The experimenter brought the demonstration box onto the stage with the purple curtain closed. She opened the curtain to reveal the seven balls of each color, flipped the box upside down and then right side up, turned the box side to side and shook it, and placed it back on the stage. This allowed infants to see that the green balls that were stuck did not move even when the box was shaken and turned. She reached into the box to pick up the balls, her hand visible to the infant, starting with either the red or green balls in counterbalanced order. For example, she picked up each of the four red balls to the top of the box, then lowered it, and left it on the floor of the box. Then she grasped and

lifted each of three stuck green balls, which resulted in the entire box moving, and then the one movable green ball, which she lifted to the top of the box and lowered just as she had done with the red balls. She repeated the entire sequence once more, this time grasping four stuck green balls, rather than three stuck balls and one movable ball.

Test trials (four trials). On each trial, the experimenter placed the population box on the stage, with its front curtain closed. She shook the box a few times, closed her eyes, turned her head away, and reached into the box, drawing out three balls and placing them into a narrow transparent container (the sample container) on the front right corner of the stage. She then repeated this action, drawing out two more balls. The experimenter lifted the front curtain of the box to reveal the population, always of mostly green balls for every infant and lowered the backdrop while saying, “Look, [baby’s name], look!” (Figure 1). Timing proceeded as in the familiarization trials. The samples alternated between the one red and four green balls and the one green and four red balls, resulting in two pairs of test trials. The order of the samples drawn on the test trials (one green and four red balls or one red and four green balls first) was counterbalanced across infants.

Predictions. The infants could make a variety of different inferences, each of which predicts a different pattern of looking. First, infants might appropriately adjust the base rate of the population based on the newly taught probabilistic physical constraint (i.e., most but not all green balls were unmovable) and estimate the probabilities of the two sampling outcomes given this constraint. That is, many more green balls were present in the population box,

but the majority of these balls were unavailable for sampling; thus, infants should look longer at the one-red-and-four-green sample than at the one-green-and-four-red sample (Figure 2A). Second, if infants are unable to learn and integrate this probabilistic physical constraint, or if they base their judgments on the similarity in appearance between the population and the sample, the proportion of balls in the box should predict their looking behavior. That is, infants should look longer at the one-green-and-four-red sample than at the one-red-and-four-green sample, given the ratio of Ping-Pong balls in the box. This experiment pitted probability estimations against similarity judgments by assessing whether infants would use probability estimations that require an adjustment to the base rate of a population or assess the similarity in appearance between the ratios of the population and samples. One final possibility was that infants would ignore the probabilistic nature of the physical constraint and apply the physical constraint to all green balls. This could occur if infants have difficulty following the procedure and do not notice that there are some green balls that move. In this case, infants should look approximately equally at both samples, as removing any green balls from the population box would be impossible.

In mathematical terms, infants saw a 3:1 population (75% green balls, 25% red balls) and 4:1 samples. According to the demonstrations, on average, 79% of the green balls were stuck in the box: 75% (free play), 75% (Demonstration 1), and 86% (Demonstration 2); thus, $45^{*}.79 = 35$ green balls were stuck, converting the population available for sampling to 10 green to 15 red balls (2:3). The probability of drawing one green ball and four red balls was $5!/4!(3/5)4(2/5) = 0.259$; the probability of drawing one red ball and four green balls was $5!/4!(2/5)4(3/5) = 0.0768$. In other words, out of 100 draws, one was four times more likely to draw a sample of one green and four red balls than a sample of one red and four green balls. If infants did not learn the constraint that the majority of the green balls were stuck, their probability of drawing one green and four red balls was $5!/4!(1/4)4(3/4) = 0.0146$; their probability of drawing one red and four green balls was $5!/4!(3/4)4(1/4) = 0.396$. That is, out of 100 draws, one was 45 times more likely to draw a sample of one red and four green balls than a sample of one green and four red balls. If instead, infants applied the rule deterministically to all the green balls, the probability of both samples equaled zero.

Results

Preliminary analyses found no effects of gender, order of familiarization trials (mostly red or mostly green box first), or order of test trials (expected vs. unexpected first); subsequent analyses collapsed across these variables. A second observer, unaware of the order of the trials, timed 50% of the familiarization and test trials. Interobserver reliability averaged 94%.

An alpha level of .05 was used for all analyses. An analysis of variance (ANOVA) examined the effects of trial pair (1, 2) and outcome (one-green-and-four-red samples vs. one-red-and-four-green samples). There was a main effect of outcome, $F(1, 11) = 9.585$, $p = .015$, effect size (η_p^2) = .542. Infants looked reliably longer at the one-red-and-four-green samples ($M = 13.71$ s, $SD = 7.15$) than the one-green-and-four-red samples ($M = 8.63$ s, $SD = 5.067$). There were no other main effects or interactions. Eleven of 12 infants looked longer at the one-red-

and-four-green outcomes, Wilcoxon signed-ranks test: $z = 2.748$, $p = .006$.

Discussion

Infants in this experiment were exposed to three instances indicating that most green balls in a set had the property of being immovable from boxes and containers and that all the red balls moved. On test trials, they looked longer at events in which an experimenter drew samples of one red and four green balls rather than one green and four red balls from a box with a ratio of three green balls to one red ball. This provides evidence that infants can quickly acquire a probabilistic physical constraint, namely, that most but not all balls in a set (approximately 80%) were immobile. It also suggests that infants were able to generalize this new constraint to a new set of balls—the box of balls used during test trials. Finally, infants integrated this constraint with the overall distribution of balls in the box in order to infer that sampling one red and four green balls was fairly unlikely and sampling one green and four red balls was more likely from the remaining population. It is even more impressive that infants produced this pattern of looking, given that the perceptual appearance of the test trial outcomes was in direct opposition with the proportions of the population. It appears that infants can learn and integrate a complex rule and that they do so even when this directly pits perceptual similarity against probability.

Although infants' looking behavior is consistent with this interpretation, an alternative interpretation of these data exists. It is possible that infants looked longer at the one-red-and-four-green-ball sampling events not because they understood and integrated the constraint but because it was less similar to the events they had seen in the demonstration phases than the one-green-and-four-red-ball sampling events. That is, infants may have looked longer when a relatively large number of green balls were removed from the box on test trials because they had seen fewer green balls being removed from boxes during the demonstration phases than red balls. The experimenter lifted eight red balls but only two green balls from the container in Demonstration Phase 1, and she lifted eight red balls but only one green ball in the box on Demonstration Phase 2. This design was necessary to illustrate that just a small proportion of green balls were movable. However, a stronger design would control for the absolute number of times each color of ball was lifted from the containers on these demonstrations. Experiment 2 addressed this concern.

Experiment 2

Method

Participants. Twenty-four 11-month-old infants participated in this experiment (17 females; mean age = 11.07 months; age range = 10.5–11.6 months). Four additional infants were tested and their data excluded, one for parental interference and three for infant fussiness.

Procedure, design, and predictions. The procedure was the same as Experiment 1, except for the following changes made to the demonstration and familiarization trials:

In each demonstration phase, the experimenter equated the number of times that she lifted red and green balls out of the

container. In Demonstration Phase 1, she grasped the six immobile green balls one at a time and then picked up the two movable balls one at a time and repeated her actions with the two movable balls four times each, such that infants saw eight balls of each color being removed from the container. She then grasped all six green immobile balls again to remind the infant that most of the balls were stuck. In Demonstration Phase 2, she lifted the one movable green ball four times, again to equate this with the number of times red balls were lifted, and grasped the stuck green balls without being able to lift them.

Each infant received two familiarization trials, both times observing the mostly green box. In Experiment 1, infants observed two familiarization trials with a population containing a ratio of three red balls to one green ball and two familiarization trials with a population containing a ratio of three green balls to one red ball. This design equates the amount of red and green balls visible to infants until test trials began and potentially primes infants to attend to distributional information. These familiarization boxes were identical in appearance when the front curtains were closed, and the population contents were not visible (in fact, they were actually two sides of the same box). If infants did not pay attention to the contents of the box when revealed on test trials and simply assumed that it alternated as it did in familiarization, they might have assumed that the experimenter was sampling from the mostly red population. Although previous results suggest that this is unlikely (e.g., Denison & Xu, 2010a; Xu & Garcia, 2008), the concern with this design is that if the infant assumes that the experimenter is sampling from the mostly red box on half of the test trials, the one-green-and-four-red-balls sample should be expected under any and all of the following potential belief states of the infant: (a) The infant did not understand the constraint at all and thought all balls were movable, (b) infant understood the constraint but failed to integrate it with the assumed ratio of three red balls to one green ball in the box, or (c) infant correctly integrated the constraint with the assumed ratio of three red balls to one green ball. To rule out this unlikely interpretation, in Experiment 2 (and Experiment 3) we removed the mostly red box from the familiarizations.

The predictions for Experiment 2 were the same as those for Experiment 1. Infants should look longer at the one-red-and-four-green-balls sample than the one-green-and-four-red-balls sample if they comprehend and integrate the probabilistic constraint.

Results

Preliminary analyses found no effects of gender or order of test trials; subsequent analyses collapsed across these variables. A second observer timed 50% of the familiarization and test trials and interobserver reliability averaged 95%.

We performed an ANOVA to examine the effects of trial pair (1, 2) and outcome (one-green-and-four-red samples or one-red-and-four-green samples). There was a main effect of outcome, $F(1, 23) = 6.129$, $p = .022$, effect size (η_p^2) = .235. Infants looked reliably longer at the one-red-and-four-green samples ($M = 11.62$ s, $SD = 5.51$) than the one-green-and-four-red samples ($M = 9.12$ s, $SD = 4.15$). There were no other main effects or interactions. Eighteen of 24 infants looked longer at the one-red-and-four-green outcomes, Wilcoxon signed-ranks test: $z = 2.432$, $p = .015$.

We also compared infants' looking times across Experiments 1 and 2 to assess whether infants performed similarly on these two experiments, despite the procedural differences. As expected, there were no significant differences between infants' looking times on the expected events across experiments, $t(34) = 0.29$, $p = .775$, or the unexpected events, $t(34) = 0.97$, $p = .339$. These analyses suggest that infants' performance in Experiment 1 was unlikely to be driven exclusively by the concerns regarding the display phases and familiarization trials outlined in the Discussion, as performance did not differ across the two experiments.

Discussion

As in Experiment 1, infants looked longer at samples of one red and four green balls than at samples of one green and four red balls, indicating that they had learned the constraint that the majority of green balls were stuck in boxes and integrated this new constraint with the three-green-to-one-red ratio in the population box.² Infants in this experiment could not have looked longer at this sample simply because they had become familiar with observing more instances of red balls being removed from boxes in the demonstration phases, as the number of times balls were lifted was equated across colors in this experiment.

Another possible interpretation of these findings is that infants simply showed their color preferences on test trials, as color was not counterbalanced across infants. This is unlikely because previous experiments using identical stimuli did counterbalance color, and no effect of color was found (Denison & Xu, 2010a). Furthermore, the data in Experiment 3 suggest that infants do not show color preferences on test trials.

Experiment 3

The purpose of Experiments 1 and 2 was to test whether infants can integrate a nondeterministic or probabilistic physical constraint in statistical inference. Therefore, it is important to demonstrate that infants do not draw the same conclusions when given evidence that the constraint applies to all balls in the set. In Experiment 3, we examined infants' inferences in cases where all green balls in the demonstrations were immobile, making both the one-green-and-four-red-balls sample and the one-red-and-four-green-balls sample impossible events. Again, we asked a crucial question about infants' early physical reasoning: Do infants understand the nature of a truly impossible physical event when they observe a novel constraint over just a few trials? That is, would the infants erroneously believe that drawing one green ball and four red balls was more probable than drawing four green balls and one red ball when given evidence that none of the green balls can be moved? Or would the infants correctly reason that the samples were equally improbable and look about equally at each, given that removing any green balls should be interpreted as impossible? If infants looked longer at the one-red-and-four-green sample than at

² Do adults share infants' intuitions? In a pen-and-paper version, we showed 24 adults video clips of Experiment 2 and asked them to rate the test trials. On a 7-point scale (ranging from *highly unexpected* to *highly expected*), adults rated the one-red-and-four-green sample as more unexpected ($M = 2.88$, $SD = 1.85$) than the one-green-and-four-red sample ($M = 4.67$, $SD = 1.95$), $F(1, 23) = 6.877$, $p = .015$, $\eta_p^2 = 0.23$.

the one-green-and-four-red sample, it would suggest that they do not understand what it means for a physical event to be impossible in our experimental setting. This looking pattern would also make the results of Experiments 1 and 2 more difficult to interpret, as it would be unclear whether infants in these experiments correctly integrated the probabilistic constraint or instead thought that the constraint applied deterministically but that the experimenter would still be more likely to draw just one green ball rather than four (see Figure 2).

Method

Participants. Twenty-four 11-month-old infants participated in this experiment (16 females; mean age = 11.36 months; age range = 10.5–11.57 months). Four additional infants were tested and their data excluded, one due to parental interference, two due to fussiness, and one due to experimenter error.

Procedure. As in Experiment 2, infants were familiarized to only the mostly green side of the population box during familiarization trials. In the free play phase and both demonstration phases, infants saw that all green balls were immobile; these trials otherwise proceeded exactly as they had in Experiment 2. The test trials proceeded exactly as in Experiments 1 and 2.

Results and Discussion

Preliminary analyses showed no effects of gender or order of the test trials; subsequent analyses collapsed over these variables. A second observer timed 25% of the familiarization and test trials. Interobserver reliability averaged 95%.

An ANOVA examined trial pair (1, 2) and outcome (one-green-four-red samples vs. one-red-and-four-green samples). There was no main effect of outcome, $F(1, 23) = 0.01$, $p = .942$, effect size (η_p^2) = .000. Infants looked approximately equally at the one-red-four-green samples ($M = 11.68$ s, $SD = 3.856$) and the one-green-four-red samples ($M = 11.74$ s, $SD = 5.121$). Twelve of 24 infants looked longer at the one-red-four-green sample (the unexpected sample in Experiments 1 and 2), Wilcoxon signed-ranks test: $z = 0.69$, $p = .490$.³

Although infants looked for approximately equal amounts of time at each outcome, this result on its own cannot provide evidence that infants found both outcomes equally *unexpected*. Infants in all experiments observed identical test trials but were either taught a probabilistic constraint (Experiments 1 and 2) or a deterministic constraint (Experiment 3). Thus, the looking times in Experiment 3 need to be compared with those of Experiments 1 and 2. We conducted two planned comparisons, collapsing the data from Experiments 1 and 2, as these were not significantly different from one another (see analyses in the Results section of Experiment 2). Infants looked reliably longer at the one-green-four-red sample in Experiment 3 than in Experiments 1 and 2, $t(58) = 2.46$, $p = .0169$. They looked approximately equally at the one-red-and-four-green sample in Experiment 3 as they did in Experiments 1 and 2 combined, $t(58) = 0.39$, $p = .698$. This pattern of results indicates that the infants in Experiment 3 found both samples unlikely (see Figure 3).

Infants' differential performance in these experiments suggests that they were able to rapidly learn either a probabilistic or deterministic physical constraint, use it to adjust the base rate of a

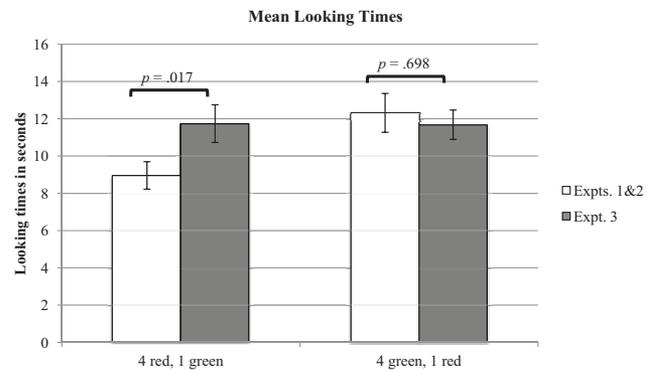


Figure 3. Mean looking times for infants in Experiments 1 and 2 combined (the probabilistic-constraint experiments) and Experiment 3 (the deterministic-constraint experiment). Labels “4 red, 1 green” and “4 green, 1 red” refer to the number of balls in each color used in the experiment. Expt = Experiment.

population, and make inferences accordingly. In doing so, their performance indicates that infants do not simply compare the appearance of a small sample drawn and a larger population to make inferences. Infants appear to be remarkably sophisticated—and flexible—when integrating physical knowledge in probabilistic inference.

General Discussion

The present experiments provide evidence that human infants are able to acquire understanding of physical constraints rapidly and use them to adjust the base rate of a population to estimate the probabilities of two outcomes. Impressively, infants acquired these physical constraints flexibly, learning either a probabilistic or deterministic constraint for a set of identical objects, based solely on the data given. Experiment 1 provides evidence that infants can acquire a probabilistic physical constraint. Infants learned that most, but not all, green balls were physically stuck inside boxes and containers. Informed by this constraint, infants adjusted the base rate of balls in a population box with a 3:1 (green balls: red ball) ratio to correctly infer that a sample of mostly red balls was more probable than a sample of mostly green balls. Experiment 2 provided a replication of this finding and controlled for potential alternative interpretations of infants' looking behavior in Experiment 1. In Experiment 3, infants observed that all green balls had the physical property of being stuck inside boxes (a deterministic constraint) and viewed the same sampling events as in Experiments 1 and 2. It should be noted that under these conditions, infants' looking behavior suggested that they found both the mostly red and mostly green samples equally unexpected, suggesting that they considered the removal of even one green ball an impossible event.

We argue that these data provide compelling evidence that infants can learn a physical constraint that applies either probabi-

³ Again, we showed adults the displays. They rated the samples as equally unexpected, rating the one-red-and-four-green sample as $M = 2.93$ ($SD = 1.83$) and the one-green-and-four-red sample as $M = 2.68$ ($SD = 1.23$), $F(1, 15) = 0.652$, $p = .432$, $\eta_p^2 = 0.042$.

listically or deterministically to a set of objects and generalize this constraint to a new context. However, the current experiments leave open the possibility that the infants did not integrate this constraint in probabilistic inference. In Experiments 1 and 2, infants were shown that most but not all green balls were immobile and that no red balls were immobile. It is possible that infants would have looked longer at the sample containing a larger number of green than red balls based solely on the physical constraint, had the contents of the population box never been revealed, as they might have simply been surprised to see more green than red balls being removed from a box after learning that green balls tend to be immobile?⁴ On this interpretation, infants need not consider either the principles of random sampling or the proportion of balls in the box. This possibility is further exacerbated by the fact that infants participated in the free play phase at the beginning of the experiment, which allowed them to feel the texture difference between the red balls and green balls, which were covered with small strips of Velcro. We placed Velcro markings on the green balls to remind infants that they had a special physical property. However, infants may have noticed this texture difference while handling the balls at the beginning of the experiment and then reasoned that the experimenter would be able to feel this difference while drawing the balls out of the box and would thus not be engaging in random sampling but instead choosing balls based on the texture differences.

Although this alternative interpretation is possible, we find it unlikely for a number of reasons. First, we know of no evidence suggestive of infants being capable of ascribing texture knowledge from themselves to another individual, based on a short tactile familiarization with objects. Infants would be required to make the following inferences to conclude that the experimenter was engaging in intentional sampling: (a) Encode the difference in texture between the two classes of objects based on a short (approximately 45 s) familiarization with the objects; (b) recognize that the visual objects viewed during the main experiment have the same textural properties as the initially felt objects; (c) infer that the experimenter will have the same knowledge of the texture difference between the two classes of objects; (d) conclude that this negates random sampling, as the experimenter would be able to feel the balls and make choices based on texture while drawing balls under occlusion. The literature on cross-modal matching suggests that both (a) and (b) are in infants' representational repertoire at 11 months of age, although these experiments allow much longer familiarization time for the tactile exploration, as they typically employ full habituation (e.g., Sann & Streri, 2007). The inferences required in (c) and (d) are much more complex, and to our knowledge, there is no research to date to suggest that infants are capable of making these inferences.

Second, the experimenter closed her eyes and turned her head away from the box during sampling, which infants in a variety of previous experiments have taken as strong cues to random sampling (Denison & Xu, 2010a; Denison et al., 2013; Xu & Denison, 2009; Xu & Garcia, 2008). Thus, the experimenter's explicit demonstrations of random sampling were at odds with the arguably much subtler potential concern that the experimenter could feel the balls and was thus sampling intentionally. Furthermore, the adults who participated in a pen-and-paper version of Experiment 2 also participated in the Free play phase, handling the balls and presumably noticing the textural differences between the green

and red balls. These adults completed an open-ended question at the conclusion of the experiment: "What did you think this experiment was testing?" Fifteen of the 21 adults who responded to this question referenced at least one of the words "proportions," "probability," "distributions," "sampling," or "random sampling." Seventeen of the 21 participants mentioned something about the large box (i.e., the population box) in their descriptions, suggesting that they were looking at this box when making judgments and not just attending to the sample display container. No adults mentioned that they thought the experimenter would be able to feel the textural difference between the green and red balls and that she could therefore make intentional choices based on the presence or absence of Velcro on the balls.

Finally, the data suggest that infants did consider the contents of the population box when making inferences in Experiments 1 and 2. Infants were shown that most but not all green balls in a set were immobile. If the infants had no knowledge of the approximate number of green balls in the population box, then they would have no basis for judging whether the green balls were behaving consistently with the learned probabilistic constraint, because the box could contain any number of green balls. That is, when one green ball and four red balls were removed on the expected trial, the population box could have been revealed to contain zero additional green balls, which would violate an infant's expectations of a probabilistic physical constraint. Therefore, we suspect that had the test trials stopped after the sampling phase, and the infants truly had no knowledge of the contents of the population box, they should have looked about equally at each sample, not knowing whether the entire set of green balls had been removed. Additionally, we suspect that infants in Experiment 2 knew the approximate proportion of red and green balls in the box, even before the contents of the population box was revealed on test trials, as they had viewed the population box on the familiarization trials, and it contained the three-green-balls-to-one-red-ball ratio on all trials.

These findings make important contributions to a rich literature on infant physical reasoning and a growing research enterprise on infant probabilistic reasoning. In the real world, the outcomes of physical events are incredibly variable, and adults effortlessly make predictions about complex and stochastic physical events, based on generalizations such as *glass is breakable* and *pasta is sticky*. Yet developmental research had not previously addressed whether infants can reason about physical constraints that apply probabilistically. The findings presented here suggest that infants are also capable of acquiring stochastic physical constraints. Infants might acquire these variable physical constraints through powerful statistical learning mechanisms that they have been shown to possess, time and again. The ability to acquire these constraints rapidly and integrate them to predict the outcomes of future events might help to elucidate how infants come to know so much about the physical world so early in development.

In addition to examining the acquisition of physical constraints, this set of findings also extends and clarifies the growing body of research on infant probabilistic reasoning in important ways. A number of recent experiments have addressed whether infants can integrate substantive domain knowledge into domain-general

⁴ We thank an anonymous reviewer for bringing this alternative interpretation to our attention.

probabilistic inferential mechanisms (see Cesana-Arlotti, Téglás, & Bonatti, 2012, and Denison & Xu, 2012, for reviews). Findings reveal that preverbal infants can reason about both psychological and physical variables when making probabilistic inferences. For example, 11-month-old infants reasoned that a sample should be representative of a population if it is drawn at random, but if an experimenter has a goal for one color object over the other and then draws a sample while looking at the population, the sample should instead reflect her preferences (Xu & Denison, 2009). Additionally, when 12-month-old infants saw a lottery machine containing one blue and three yellow objects, they considered actors such as proximity from a chute and the length of delay before the object exited the chute, along with the ratio of objects, to predict which type of object was most likely to exit (Téglás et al., 2011). The present findings extend this literature, by demonstrating that infants can integrate a probabilistic physical constraint into their probability computations. Infants were able to fully integrate probabilistic inference with physical reasoning, by using the physical constraint to adjust the base rate of balls in the population and make inferences about the likely outcomes of sampling events based on true probabilistic reasoning and not simple heuristics.

One question that remains regarding the nature of the physical constraints learned by the infants in this experiment is whether they interpreted and represented the physical constraint governing the green balls in Experiment 3 as truly impossible (with probability = 0) or as very low probability. The data suggest that infants may have treated these as low-probability—and not impossible—events, as infants' looking times to these outcomes did not differ significantly from looking times to the improbable events in Experiments 1 and 2. The current experiments do not directly address this question, as infants may have simply reached ceiling on the improbable trials in Experiments 1 and 2. On the other hand, young children will state that a variety of extraordinary physical events are impossible (sometimes even when they are not impossible but are just highly improbable; see Shtulman & Carey, 2007), suggesting that they represent some physical events as having zero probability. In general, it is not clear whether infants in VOE looking-time experiments find the “unexpected” events impossible or simply highly unlikely; all they can definitively reveal is whether infants in a particular task find one outcome more or less “expected” than another. Future research is required to determine whether infants make a categorical distinction between highly improbable and impossible events.

Finally, the present experiments have implications for some dual-process models of human reasoning. Most of these models suggest that normatively correct statistical reasoning relies on explicit, effortful, and verbal means and thus increases with age (Evans, 2003; Kokis et al., 2002; Reyna & Brainerd, 1995; Slovic, 1996). Our findings are consistent with an alternative hypothesis: that human infants may begin life with a set of powerful statistical, inferential mechanisms that guide learning, and later on in development, they acquire heuristics, shortcuts, and factual knowledge that sometimes override correct statistical thinking in exchange for expediency and automaticity (see also Cesana-Arlotti et al., 2012; Denison & Xu, 2012, for similar arguments). The infants in our experiments made rational statistical inferences about the likelihood of obtaining different samples by appropriately adjusting the base rates of objects in the populations through

applying a physical constraint. However, they could have reasoned about the sampling events via the representativeness heuristic (Tversky & Kahneman, 1974). Experiments 1 and 2 directly pitted accurate probabilistic reasoning against the representativeness heuristic, as use of representativeness would have predicted the opposite pattern of looking times. Infants did not rely on the representativeness heuristic in this experiment, suggesting that preverbal infants might not be particularly susceptible to reasoning biases. Of course, this set of experiments only scratches the surface of the kind of empirical support that is necessary to make such a strong claim. Much further research is required to assess whether infants use reasoning shortcuts or heuristics in other situations.

In conclusion, we provide evidence that preverbal infants can rapidly acquire probabilistic physical constraints and apply them when reasoning about the probabilities of subsequent events. The findings shed light on the origins of the ability to acquire probabilistic constraints that govern our physical environment and suggest that infants can do so based solely on observations of the statistical properties of the objects.

References

- Aguiar, A., & Baillargeon, R. (1999). 2.5-month-old infants' reasoning about when objects should and should not be occluded. *Cognitive Psychology, 39*, 116–157. doi:10.1006/cogp.1999.0717
- Aslin, R., Saffran, J., & Newport, E. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*, 321–324. doi:10.1111/1467-9280.00063
- Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science, 3*, 2–13. doi:10.1111/j.1745-6916.2008.00056.x
- Baillargeon, R., & Carey, S. (2012). Core cognition and beyond: The acquisition of physical and numerical knowledge. In S. Pauen (Ed.), *Early childhood development and later outcome* (pp. 33–65). London, England: Cambridge University Press.
- Baillargeon, R., & DeVos, J. (1991). Object permanence in 3.5- and 4.5-month-old infants: Further evidence. *Child Development, 62*, 1227–1246. doi:10.2307/1130803
- Baillargeon, R., Li, J., Ng, W., & Yuan, S. (2009). An account of infants' physical reasoning. In A. Woodward & A. Needham (Eds.), *Learning and the infant mind* (pp. 66–116). New York, NY: Oxford University Press.
- Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in 5-month-old infants. *Cognition, 20*, 191–208. doi:10.1016/0010-0277(85)90008-3
- Casstevens, R. M. (2007). jHab: Java habituation software (Version 1.0.0) [Computer software]. Chevy Chase, MD: Java.
- Cesana-Arlotti, N., Téglás, E., & Bonatti, L. (2012). The probable and the possible at 12 months: Intuitive reasoning about the uncertain future. In J. B. Benson (Series Ed.), F. Xu & T. Kushnir (Vol. Eds.), *Advances in child development and behavior: Rational constructivism in cognitive development* (Vol. 43, pp. 1–26). Waltham, MA: Academic Press.
- 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*, 243–249. doi:10.1037/a0028278
- Denison, S., & Xu, F. (2010a). Integrating physical constraints in statistical inference by 11-month-old infants. *Cognitive Science, 34*, 885–908. doi:10.1111/j.1551-6709.2010.01111.x
- 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. doi:10.1111/j.1467-7687.2009.00943.x
- Denison, S., & Xu, F. (2012). Probabilistic inference in human infants. In

- J. B. Benson (Series Ed.), F. Xu & T. Kushnir (Vol. Eds.), *Advances in child development and behavior: Rational constructivism in cognitive development* (Vol. 43, pp. 27–58). Waltham, MA: Academic Press. doi:10.1016/B978-0-12-397919-3.00002-2
- Evans, J. St. B. T. (2003). In two minds: Dual process accounts of reasoning. *Trends in Cognitive Sciences*, 7, 454–459. doi:10.1016/j.tics.2003.08.012
- Gerken, L. (2006). Decisions, decisions: Infant language learning when multiple generalizations are possible. *Cognition*, 98, B67–B74. doi:10.1016/j.cognition.2005.03.003
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13, 431–436. doi:10.1111/1467-9280.00476
- Gopnik, A., Glymour, C., Sobel, D., Schulz, L., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, 111, 3–32. doi:10.1037/0033-295X.111.1.3
- 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 of the United States of America*, 107, 9066–9071. doi:10.1073/pnas.1003095107
- Hespos, S. J., & Baillargeon, R. (2001). Infants' knowledge about occlusion and containment events: A surprising discrepancy. *Psychological Science*, 12, 141–147. doi:10.1111/1467-9280.00324
- Hespos, S. J., & Baillargeon, R. (2006). Décalage in infants' knowledge about occlusion and containment events: Converging evidence from action tasks. *Cognition*, 99, B31–B41. doi:10.1016/j.cognition.2005.01.010
- 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. doi:10.1016/S0022-0965(02)00121-2
- Luo, Y., & Baillargeon, R. (2005). When the ordinary seems unexpected: Evidence for rule-based physical reasoning in young infants. *Cognition*, 95, 297–328. doi:10.1016/j.cognition.2004.01.010
- Marcus, G. F., Vijayan, S., Rao, B. S., & Vishton, P. M. (1999, January 1). Rule learning by seven-month-old infants. *Science*, 283, 77–80. doi:10.1126/science.283.5398.77
- Newcombe, N. S., Huttenlocher, J., & Learmonth, A. E. (1999). Infants' coding of location in continuous space. *Infant Behavior & Development*, 22, 483–510. doi:10.1016/S0163-6383(00)00011-4
- Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, 7, 1–75. doi:10.1016/1041-6080(95)90031-4
- Saffran, J., Aslin, R., & Newport, E. (1996, December 13). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928. doi:10.1126/science.274.5294.1926
- Sann, C., & Streri, A. (2007). Perception of object shape and texture in human newborns: Evidence from cross-modal transfer tasks. *Developmental Science*, 10, 399–410. doi:10.1111/j.1467-7687.2007.00593.x
- Shtulman, A., & Carey, S. (2007). Improbable or impossible: How children reason about the possibility of extraordinary events. *Child Development*, 78, 1015–1032. doi:10.1111/j.1467-8624.2007.01047.x
- Simon, T., Hespos, S. J., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development*, 10, 253–269. doi:10.1016/0885-2014(95)90011-X
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3–22. doi:10.1037/0033-2909.119.1.3
- Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, 13, 13–19. doi:10.1111/1467-9280.00403
- Sobel, D. M., & Kirkham, N. Z. (2006). Blickets and babies: The development of causal reasoning in toddlers and infants. *Developmental Psychology*, 42, 1103–1115. doi:10.1037/0012-1649.42.6.1103
- Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, 99, 605–632. doi:10.1037/0033-295X.99.4.605
- Téglás, 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. doi:10.1073/pnas.0700271104
- Téglás, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011, May 27). Pure reasoning in 12-month-old infants as probabilistic inference. *Science*, 332, 1054–1059. doi:10.1126/science.1196404
- Tversky, A., & Kahneman, D. (1974, September 27). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131. doi:10.1126/science.185.4157.1124
- Xu, F., & Denison, S. (2009). Statistical inference and sensitivity to sampling in 11-month-old infants. *Cognition*, 112, 97–104. doi:10.1016/j.cognition.2009.04.006
- 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. doi:10.1073/pnas.0704450105

Received January 30, 2013

Revision received January 14, 2014

Accepted April 11, 2014 ■