

# Infants Use Statistical Sampling to Understand the Psychological World

Henry M. Wellman  
*University of Michigan*

Tamar Kushnir  
*Cornell University*

Fei Xu  
*University of California, Berkeley*

Kimberly A. Brink  
*University of Michigan*

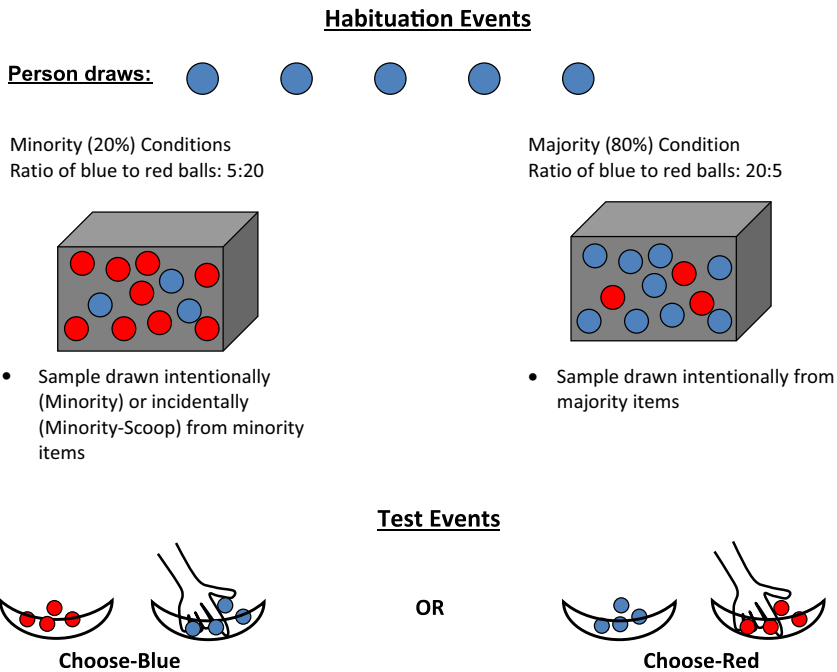
Preverbal infants engage in statistical and probabilistic inference to learn about their linguistic and physical worlds. Do they also employ probabilistic information to understand their social world? Do they infer underlying causal mechanisms from statistical data? Here, we show, with looking-time methods, that 10-month-olds attend to statistical information to understand their social–psychological world and plausibly infer underlying causal mechanisms from violations of physical probabilities.

In recent decades, a crucial advance has occurred in our understanding of childhood development: demonstrations that even young infants use statistical learning (e.g., Safra, Aslin, & Newport, 1996; Téglás et al., 2011; Xu & Garcia, 2008) to infer the structure of language and the physical world. Infants also live in a social world, full of intentional agents acting in accordance with their goals, desires, and beliefs. To what extent is statistical learning similarly instrumental for social understandings in infancy?

To adults, overt human intentional actions make manifest unobservable causes: to fulfill desires, agents deliberately manipulate the overt, observable world. For example, a haphazard handful from the Halloween candy bowl is unlikely to produce our favorites, so if someone deliberately chooses just five Snickers, we can infer she preferred those. In contrast, selecting from a bowl full of Snickers gives little information for inferring a preference for Snickers over other candies. Thus, intentional acts that

violate physical probabilities can inform us about psychological causes where the same act in the absence of relevant statistical information is less informative.

Recent studies (Kushnir, Xu, & Wellman, 2010; Ma & Xu, 2011) demonstrate that preschoolers and toddlers use violations of physical probabilities like these to infer agents' preferences. Consider Figure 1. The Minority condition agent removes five blue balls from a box of 80% *red* ones. An observer sensitive to the relation between sample and population could infer the person drew this nonrandom, low-probability sample because of some sort of desire or preference for blue balls. If the person takes five blue balls from a box of 80% *blue* ones, this largely reflects the probabilities of the underlying population and as such provides ambiguous information about her preference. It is intriguing that even toddlers (20-month-olds) infer a psychological cause—a preference for one type of object over another—from this statistical pattern, even when the preference differs from their own. But by 20 months children have accumulated considerable information about person's actions and desires (Repacholi & Gopnik, 1997), including verbal information from others (Bartsch & Wellman, 1995; Ruffman, Slade, & Crowe, 2002). Conceivably, encountering and using words like “want” scaffolds children's social-statistical understandings. Thus, data from toddlers alone leave a theoretical gap as to the origins and nature of statistical learning in understanding human action, a gap that requires data from preverbal infants to fill.



**Figure 1** Schematic of experimental events. In habituation, infants saw a person draw five blue balls from a transparent container holding either 20% blue balls (Minority and Minority-Scoop Conditions) or 80% blue balls (Majority Condition). In test, infants saw the same person seated between two transparent bowls, one of blue and one of red balls. The person leaned and grasped either a blue or a red ball.

Previous research extensively and conclusively indicates that preverbal infants infer preferences from intentional actions that do not violate physical probabilities (e.g., Csibra, Gergely, Biró, Koós, & Brockbank, 1999; Phillips, Wellman & Spelke, 2002; Sommerville, Woodward, & Needham, 2005; Woodward, 1998). These studies typically involve a repeated presentation (to habituation) of a person reaching for one of two objects, and infants then look longer (show a violation of expectation) when the person later reaches for the previously ignored object (seemingly changing her preference or goal). It may be critical that two objects are visible to the agent (that the action reflects the choice of one alternative over another, see Luo & Baillargeon, 2005), but choosing one of two objects does not itself constitute a violation of physical probabilities (50/50 chance).

Only a few recent studies have addressed whether infants connect preferences to violations of physical probabilities. These studies build upon prior research showing infants' sensitivities to the probabilistic relations between populations and samples in the physical world. For example, Xu and Garcia (2008) showed 8-month-olds a box of white and red balls, in an 80-20 proportion. With her eyes closed, an adult drew some balls from the box. Assuming the draw was randomly generated, the distribution of balls in the sample should approximate the distribution of the balls in the box. Indeed, infants looked longer when a sample of mostly red balls was drawn from this box of mostly white balls. Such data indicate infants are sensitive to statistical relations between samples and populations but do not indicate whether infants make causal inferences about the events nor consider the person drawing the balls in terms of her intentions, desires, and preferences.

Nevertheless, a human acting on objects was central to these methods and it was important (at least to adults) that her eyes were closed. Indeed, in Xu and Denison (2009) if infants saw the experimenter draw the sample with her eyes open, 11-month-olds no longer expected the sample to reflect the statistical properties of the population. Specifically, if 11-month-olds saw an agent draw a sample from a hidden box with her eyes closed (e.g., drew five red balls from the hidden box), then the infants expected the box to have almost all red balls rather than almost all white balls—looking longer if the box was revealed to have mostly white balls. However, when the agent first established that she preferred red balls (by initially drawing red ones from a set of three red and three white, i.e., from a 50/50 set as in the studies described above) and then drew the sample of red balls with her eyes *open*, infants looked equally if the hidden box was revealed to have almost all red or almost all white balls. Thus, infants' expectation of statistical sampling was overridden if the sample was drawn by an eyes-open agent with an expressed preference.

While these studies establish that infants can infer that agents with known preferences may cause violations of physical probabilities, they leave open many questions about whether and how young infants link together agents, intentions, and statistical probabilities to learn about the social world. Importantly, can infants also use statistical/probabilistic information to infer persons' preferences? To help address these questions, we asked whether infants could use violations of probabilities as a mechanism for learning about psychological causes such as preferences. As depicted in Figure 1, we tested 10-month-olds in a violation of expectation paradigm to determine whether they make the inferences that toddlers and preschoolers do from statistical patterns of intentional actions.

## METHODS

Seventy infants ( $M = 10.19$  months, range: 9.6–11.43) participated: 26 in the Minority Condition, 24 in the Majority Condition, and 20 in a Minority-Scoop condition; 63% were European American, 23% bi- or multiracial, 6% African American, and 8% Asian American, Pacific American, or other. Twenty-four additional infants did not habituate within the maximum eight trials; six others were excluded for fussiness.

Infants saw a live actor remove five blue balls from a transparent box containing blue and red balls (see Figure 1), or saw the actor remove five red balls from the box (not shown in Figure 1). Assignment of infants to these different presentations was counterbalanced. To simplify reporting, stimuli and conditions will be described as if each infant saw the actor remove blue balls from the box.

In *Minority condition* habituation, a screen descended revealing a woman wearing a visor who looked at the box containing five blue and 20 red balls (20% blue). The woman wore a visor so that when her head was lowered it would occlude any further emotion she might display. The woman smiled, said “Hi”, and opened the box saying “Wow.” Then, she removed two blue balls, inspected them briefly while smiling and saying “Oh blue,” and placed them in a row in front of her. She then removed two more blue ones in the exact same manner, then one more, totaling five blue balls. At this point, she said “Look!” lowered her head, looked directly at the five blue balls, and sat like that until the infant looked away for a period of two consecutive seconds, or until 60 total seconds had elapsed ending that trial.

*Majority condition* habituation was identical except that the transparent box contained 20 blue and five red balls (80% blue). In each habituation trial, the woman removed the blue balls in the same manner with the same reactions across both conditions.

By hypothesis, if infants in the Minority condition look longer to the Choose-Red test event (the opposite of the actor’s expressed preference), they are recognizing that intentional actions which deliberately override physical probabilities indicate a preference. Following this reasoning, if infants saw a person produce the same sample (discordant with physical probabilities) but produced that specific sample incidentally rather than deliberately, this would not signal a preference. To test this implication, in a *Minority-Scoop condition*, we employed an action closely parallel to our Minority condition but where a sample of blue balls was taken incidentally rather than intentionally: infants saw a visored woman use a scoop to remove five blue balls from the box in one scooping action. After putting the scoop on the table, the woman then took two blue balls from the scoop (not the box), then two more, and one final one, reacting to the balls at this point and lining them up on the table just as in the other two conditions. By using a scoop and not her hand and by scooping all balls at once, the Minority-Scoop actions should indicate an incidental scoop of blue balls from the box.

Infants in all three conditions saw multiple habituation trials until they looked significantly less on their last three trials than their first three, or until they saw eight habituation trials. For those who habituated, looking-times decreased from  $M = 22.9$  sec for the first three habituation trials to  $M = 7.9$  sec for the last three in the Minority condition, from 20.8 to 7.4 in the Majority condition, and from 17.1 to 12.8 in the Minority-Scoop condition:  $t(25) = 9.67$ ,  $p < .0001$ ;  $t(23) = 10.86$ ,  $p < .0001$ ;  $t(19) = 5.54$ ,  $p < .0001$ , respectively. An 8-trial maximum was set because in pilot testing many infants fussed out of habituation if they had to sit through 10 or 12 trials of

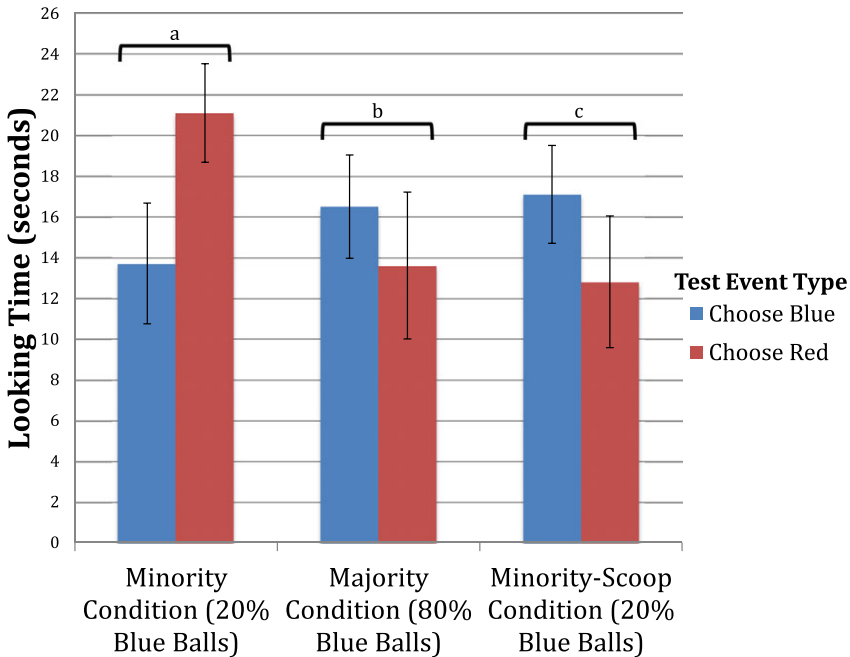
this repeated display. The 8-trial maximum also resulted in the 24 infants who failed to habituate who were noted earlier.

After habituation, infants saw one of two *Test events* where the screen descended revealing the visored woman midway between two transparent bowls, which contained either four red or four blue balls. The woman smiled, looked at each bowl once, said “There,” and lowered her head to look toward and grasp a single ball in one bowl (either red or blue). Then, the action froze until the infant looked away for 2 sec, or 60 total seconds elapsed. Right–left position of the bowls was counterbalanced. For 14/26 infants in Minority, 12/24 in Majority, and 12/20 in the Minority-Scoop conditions, the woman chose blue, for the rest she chose red. Looking-times to these test events are shown in Figure 2.

The looking-times of 15 infants, five in each condition, were recoded from videotapes by a coder completely blind to infants’ test condition. Recodings were within 1 sec of the primary coder’s times for 88.2% of the trials and within 2 sec for 93.2%.

## RESULTS

Preliminary analyses showed that infants in the Choose-Blue and Choose-Red groups did not differ in habituation prior to their test trials. A 2 (Choose-Blue versus



**Figure 2** Test-event looking-times. Data for infants in the three conditions (Minority, Majority, and Minority-Scoop) when they saw the target person choose either a red or a blue ball in the test event in the test situation where the person could freely choose. Error bars represent standard errors of the mean. As detailed in the text, differences between the groups within a condition were as follows: (a)  $p = .056$ ; (b),  $p = .67$ ; (c),  $p = .62$

Choose-Red) by 2 (Minority condition versus Control conditions) ANOVA comparing infants on their looking-times for their last three habituation trials combined showed no effect of Choose-Red versus Choose-Blue groups, no effect of condition, and no interaction, all  $ps > .14$ . More specifically, in the Minority condition looking-times for the Choose-Blue and Choose-Red groups did not differ:  $M = 8.5$  sec average per trial (Choose-Blue) versus 6.3 sec (Choose-Red);  $t(24) = 1.24$ ,  $p = .23$ . Neither did they differ in the Majority condition— $M = 7.9$  sec (Choose-Blue) versus 7.1 sec (Choose-Red),  $t(22) = 0.31$ ,  $p = .76$ —or the Minority-Scoop condition— $M = 9.4$  sec (Choose-Blue) versus 7.2 sec (Choose-Red),  $t(18) = 0.99$ ,  $p = .34$ .

For the central test-event data, following our hypotheses, we expected an interaction between looking-times to Choose-Blue versus Choose-Red test events in the focal Minority condition as opposed to the two control conditions: infants should look longer to the Choose-Red test events (over Choose-Blue test events) in the Minority condition, but not in the Majority or Minority-scoop conditions. A planned contrast comparing Choose-Blue versus Choose-Red test-event looking-times for infants in the Minority condition versus the two control conditions (Majority and Minority-Scoop) yielded the expected interaction— $F(1,64) = 4.66$ ,  $p < .035$ . We explored this interaction further with nonparametric Mann-Whitney  $U$ -tests because, as is typical in infant research, looking-times were not normally distributed. As predicted, in the focal Minority (20%) condition, infants looked longer at the Choose-Red ( $M = 21.1$  sec) than the Choose-Blue ( $M = 13.7$  sec) test event; Mann-Whitney  $U$ -test  $U(N = 26) = 47$ ,  $p = .056$ . In the Majority (80%) condition, they did not;  $M = 13.6$  sec (Choose-Red) versus 16.5 sec (Choose-Blue),  $U(N = 24) = 64$ ,  $p = .67$ . Neither did they in the Minority-Scoop condition,  $M = 12.8$  (Choose-Red) versus 17.1 sec (Choose-Blue),  $U(N = 20) = 41$ ,  $p = .62$ .

Given our habituation paradigm, infants' dishabituation to the test events provides complementary data about their attention and expectations. Because all the test events differed from habituation (the adult went from drawing balls from a single mixed box, to choosing between two single-color bowls of balls), some dishabituation can be expected for all test events. But, if during habituation in the Minority condition, infants inferred the adult had a decided preference for blue balls, then seeing the woman choose a red ball at test should be still further unexpected. Such a habituation-dishabituation comparison also individualizes each infant's test-trial looking-times relative to their habituation looking-times. Difference scores for each infant—the increase in looking from their last habituation trial to their test-event trial—showed that Minority condition infants' increase was significantly larger when the adult chose red ( $M = 15.0$  sec) than blue ( $M = 3.2$  sec);  $U(N = 26) = 41$ ,  $p < .03$ . In contrast, in the Majority condition, there was no difference between infants who saw the adult choose red ( $M = 9.3$  sec) versus blue ( $M = 12.0$ );  $U(N = 24) = 63.5$ ,  $p = .63$ . Nor was there a difference in the Minority-Scoop condition, choose red ( $M = 5.53$  sec) versus blue ( $M = 11.13$ ),  $U(N = 20) = 34.5$ ,  $p = .31$ .

## DISCUSSION

During habituation, infants in all conditions saw very similar acts: the actor took five blue balls out of a box of red and blue ones and looked at each draw of blue balls with pleasure. Moreover, at test, infants in all conditions saw the exact same acts:



either a grasp of a red ball (from its bowl) or a grasp of blue (with the women's visor occluding any emotional reaction to her choice). Only two things differed between conditions: (1) the actor intentionally removed the blue balls from a box containing 20% (Minority condition) or 80% (Majority condition) blue balls, or (2) the actor drew a sample of minority balls intentionally (Minority condition) or incidentally (Minority-Scoop). Thus, either the probability of drawing the samples was different (Minority versus Majority) or the intentionality of the removal was different (Minority versus Minority-Scoop). To be clear, the actor's habituation actions in all conditions were intentional in the overall sense of intentionally drawing balls, looking at them, and placing them in a row. It was the intentionality of the sample achieved that differed critically between Minority and Minority-Scoop conditions: deliberately drawing five blue balls in a series of separate hand movements versus incidentally scooping five blue balls in one apparently haphazard scooping motion.

Our interpretation is that the pattern of looking-times shows that infants in the focal Minority condition inferred a causal intentional state—a desire or preference—from a statistical pattern of action, that is, the agent's deliberate manipulation of the probabilities. An alternative interpretation that infants merely tracked a behavioral regularity—an agent that consistently chooses blue balls will continue that action in test—is ruled out because in that case infants should show the same looking-time pattern in both Minority and Majority conditions. Yet, infants looked longer and dishabituated at test only in the Minority condition. A second alternative interpretation that infants merely reacted to population-sample differences—for example, the nonrandom sample drawn in habituation made the color of the balls more salient in the Minority condition so that a subsequent choice of a red ball during test was especially note-worthy—is also ruled out because in that case infants should show the same looking-time pattern in the Minority and Minority-Scoop conditions. Yet, infants looked longer and dishabituated at test only in the Minority condition. Note that to behave as they did in our conditions requires infants go beyond understanding that the agent is acting intentionally and that intentional actions can override physical statistical probabilities. It further requires using an agent's intentional actions *along with* information from statistical sampling to infer agents' actions and preferences.

Two recent studies complement our findings by suggesting that infants use nonrandomness to infer the presence of agents. Ma and Xu (2011) found that given two samples that were equally probable, 9-month-olds expected a sequenced pattern (e.g., red-red-red-white-white-white-red-red-red), not a seemingly random sequence, to be produced by a human hand and not by an inanimate claw. Similarly, Newman, Keil, Kuhlmeier, and Wynn (2010) found that 12-month-olds inferred the presence of an agent if a disorderly set of objects had been transformed into two neat rows.

Our conclusion that infants use social-statistical reasoning and do so to infer psychological causes has intriguing implications. Arguably (given our data plus that of Ma & Xu, 2011 and Newman et al., 2010), what may be crucial to causal learning in the psychological domain is that intentional actions characteristically violate physical probabilities. Choosing blue balls from a box of mostly red ones, sorting the jumbled socks into their pairs, holding an unsupported spoon in midair represent everyday intentional acts and all violate mere physical probabilities, inertias, and assortments. We suggest that by observing agents repeatedly violating physical probabilities in their intentional actions, infants begin to posit unobservable causal psychological variables—for example, desires or preferences. This may be a crucial process for acquiring

psychological concepts, such as desires, preferences, goals, and eventually beliefs, in the first place.

Several researchable questions remain to confirm and explore prelinguistic infants' ability to use statistical information to infer agents' mental states. Do, or when do, infants infer a general preference—for example, that the agent would also prefer blue balls tomorrow—rather than a more specific desire—for example, the agent wants blue balls for now? Do young infants have some prior (unlearned) notion of preferences and only use statistical information to learn about the particular preference of specific agents? Or, as outlined above, are infants revealing a process whereby they begin to posit unobservable psychological variables (such as preferences) in the first place—that violations of statistical randomness strongly signal the presence of a hidden causal variable (Griffiths & Tenenbaum, 2007)?

Regardless, our demonstration of social-statistical learning in 10-month-olds suggests such learning could be a powerful contributor to childhood development of social cognition. This demonstration in infancy, coupled with parallel demonstrations for toddlers and preschoolers (Kushnir et al., 2010), suggests an important and extended continuity in early social cognition. In this way, our data add to a small but growing set of studies that show that during their first year infants reason statistically not only about physical and linguistic events but also about social–psychological events. These findings demonstrate important commonalities between social, physical, and linguistic learning in infancy, thereby adding to the literature on mechanisms of learning that can propel development further.

#### ACKNOWLEDGMENTS

This research was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development (HD022149) and the McDonnell Foundation Causal Learning Initiative. We gratefully acknowledge the help of the parents and children who participated, and the staff of the Infant Cognition Project headed by Hope Peskin-Shepherd.

#### REFERENCES

- Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. New York, NY: Oxford.
- Csibra, G., Gergely, G., Biró, S., Koós, O., & Brockbank, M. (1999). Goal attribution without agency cues. *Cognition*, 72(3), 237–267.
- Griffiths, T. L., & Tenenbaum, J. B. (2007). Two proposals for causal grammars. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 323–345). New York, NY: Oxford.
- Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science*, 21(8), 1134–1140.
- Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science*, 16(8), 601–608.
- Ma, L., & Xu, F. (2011). Young children's use of statistical sampling evidence to infer the subjectivity of preferences. *Cognition*, 120(3), 403–411.
- Newman, G. E., Keil, F. C., Kuhlmeier, V. A., & Wynn, K. (2010). Early understandings of the link between agents and order. *PNAS*, 107, 17140–17145.
- Phillips, A. T., Wellman, H. M., & Spelke, E. S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition*, 85, 53–78.



- Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires. *Developmental Psychology, 33*(1), 12–21.
- Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development, 73*(3), 734–751.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928.
- Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition, 96*(1), B1–B11.
- Téglás, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. T., & Bonatti, L. L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science, 332*(6033), 1054–1059.
- Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*(1), 1–34.
- Xu, F., & Denison, S. (2009). Statistical inference and sensitivity to sampling in 11-month-old infants. *Cognition, 112*(1), 97–104.
- Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *PNAS, 105*(13), 5012–5015.