Current Directions in Psychological Science
22(1) 28–32

DOI: 10.1177/0963721412469396
http://cdps.sagepub.com

Infants Are Rational Constructivist Learners

Fei Xu[1] and Tamar Kushnir[2]
[1]Department of Psychology, University of California, Berkeley, and [2]Department of Human Development, Cornell University

Abstract

What is the nature of human learning, and what insights can be gained from understanding early learning in infants and young children? This is an important question for understanding the human mind, the origins of knowledge, scientific reasoning, and how to best structure our educational environment. In this article, we argue for a new approach to cognitive development: *rational constructivism*. This view characterizes the child as a rational constructive learner, and it sees early learning as rational, statistical, and inferential. Empirical evidence for this approach has been accumulating rapidly, and a set of domain-general statistical and inferential mechanisms have been uncovered to explain why infants and young children learn so fast and so well.

Keywords

rational constructivism, learning mechanisms, cognitive development

A new approach to cognitive development—*rational constructivism*—has emerged in recent years. This is against the background of two traditional classes of theories in developmental psychology—nativism and empiricism, with the former emphasizing innate concepts and core knowledge systems (Chomsky, 1987; Fodor, 1981; Spelke, 1994) and the latter emphasizing perceptual primitives (i.e., basic building blocks) and the role of associative learning mechanisms (Elman et al., 1996). The new perspective on cognitive development has been dubbed "rational constructivism" (Xu, 2007; Xu, Dewar, & Perfors, 2009; Xu & Griffiths, 2011) because it blends elements of a constructivist account of development with an account of learning as rational statistical inference—the same type of learning that underlies probabilistic models of cognition (Chater & Oaksford, 2008; Griffiths, Chater, Kemp, Perfors, & Tenenbaum, 2010; Tenenbaum, Kemp, Griffiths, & Goodman, 2011). In this article, we explicate what is meant by *rational*, what is meant by *constructivist*, and what the state of the evidence is for this view, focusing on infancy.

What is Meant by "Rational"?

Rational learners integrate prior beliefs, knowledge, and biases with new evidence provided by the environment. They do so by implicitly assessing both the prior probabilities of a set of hypotheses under consideration and how strong the evidence is and how it was generated, and then combining these assessments to generate the posterior probabilities of the hypotheses (according to Bayes' rule; see Perfors, Tenenbaum, Griffiths, & Xu, 2011, for a nontechnical introduction to these ideas). What are some specific ways in which infants have been shown to be rational, statistical, and inferential learners?

First, recent studies have shown that 6- to 12-month-old infants are sensitive to probabilistic relations when making inferences from samples to populations, and vice versa (Denison, Reed, & Xu, in press; Xu & Garcia, 2008). When 6- and 8-month-old infants were shown a box of ping-pong balls—80% of which were red and 20% of which were white—and an experimenter closed her eyes and randomly drew out a handful of balls, the infants found a sample of four red balls and one white ball more probable than a sample of one red ball and four white balls (Xu & Garcia, 2008; see Fig. 1). That is, they looked reliably longer at the latter sample than the former. Thus, infants can estimate probabilities in simple statistical-inference tasks (see Teglas, Girotto, Gonzalez, & Bonatti, 2007, for converging evidence using a different experimental paradigm).

Second, when infants are given prior constraints, they can integrate them into their statistical computations. In other words, infants' sensitivity to probabilistic relations is not part of an automatic, bottom-up learning mechanism. For psychological reasoning, for example, if 11-month-old infants were given evidence that the experimenter's goal was to pick out only red balls, and the experimenter had visual access when

Corresponding Author:
Fei Xu, Department of Psychology, University of California, 3210 Tolman Hall, Berkeley, CA 94720
E-mail: fei_xu@berkeley.edu

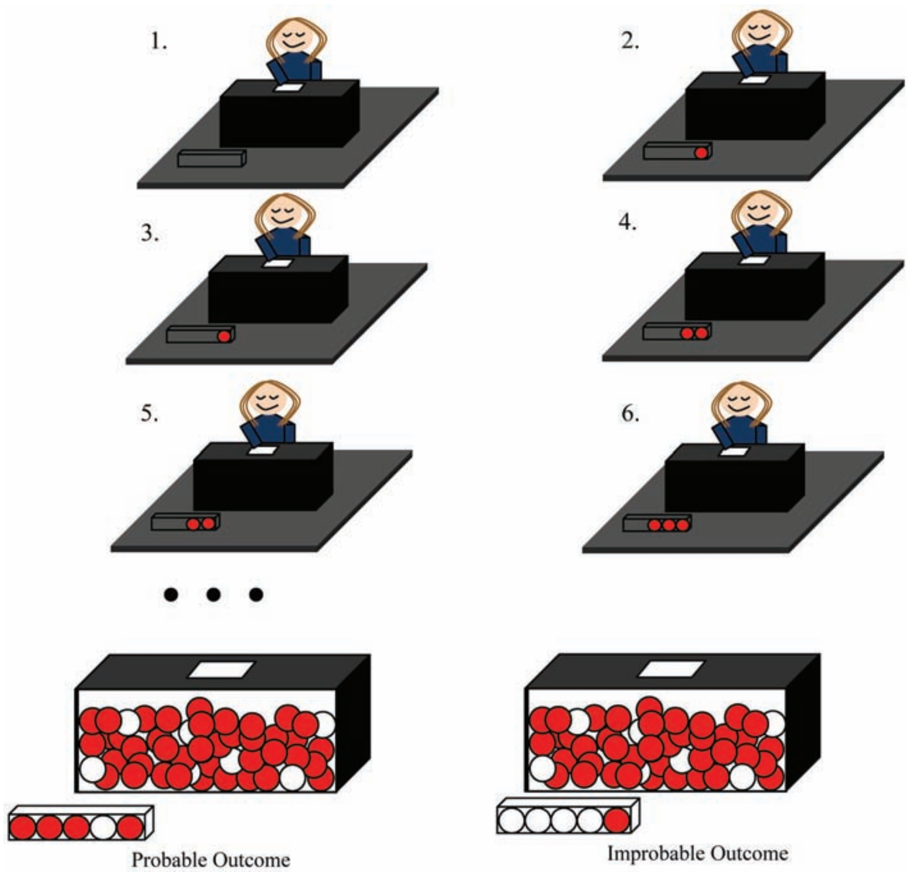

Fig. 1. Schematic representation of the experimental procedure in Xu and Garcia (2008). The experimenter shook a box and, with her eyes closed, drew one ping-pong ball from the box and placed it in a small, transparent container. This sequence was repeated four times. On the test trial, the front cover of the box was lifted to reveal that it contained mostly red balls. For the probable outcome, the experimenter had drawn four red balls and one white ball, and for the improbable outcome, the experimenter had drawn four white balls and one red ball. Infants saw the two outcomes on alternate trials.

she reached into a box of red and white ping-pong balls, the infants expected her to pull out a sample that was consistent with her goal, regardless of the proportions of red and white balls in the box. In contrast, if the same experimenter had a goal of choosing only red balls from the box but was blindfolded when reaching into it, infants expected her to pull out a sample that approximated the proportions inside the box (Xu & Denison, 2009). A similar type of integration occurs with physical constraints: If 11-month-old infants were given evidence that green balls could not be moved from a box filled with green, red, and yellow balls and the experimenter drew out a random sample with her eyes closed, they expected the sample to approximate the proportions of red and yellow balls in the box, excluding the green ones (Denison & Xu, 2010a; see also Teglas et al., 2007; Teglas et al., 2011). Infants' ability to take into account psychological and physical constraints when estimating probabilities is a hallmark of rational, inferential learning.

Third, this sensitivity to probability can be used to make predictions and guide actions. In a choice task, 10- to 14-month-old infants were given a preference trial and a test trial. On the preference trial, each infant was shown two lollipops, one pink and one black, and was allowed to choose the one he or she liked most. The experimenter then showed the infant two transparent jars of lollipops (e.g., one jar had 12 pink and 36 black lollipops, a 1:3 ratio, and the other jar had 12 pink and 4 black lollipops, a 3:1 ratio). The experimenter closed her eyes and drew out one lollipop from one of the jars (with only the stick showing) and placed it in an opaque cup. She drew out another lollipop from the second jar and placed it in a second opaque cup. The question was which cup the infant would crawl to in order to maximize his or her chances of getting the preferred (pink) lollipop. Results showed that most infants crawled to the correct cup—that is, the one that was more likely to yield a pink lollipop given the proportions of pink to black lollipops in the two jars (Denison & Xu, 2010b). Thus, infants do not only "react" to events that show improbable outcomes, as in looking-time experiments; they are also able to generate predictions based on probability estimates to guide their own actions. Infants spontaneously assess the probabilities of certain events and outcomes, an ability that constitutes a useful tool for navigating the world.

Fourth, learners take into account the statistics of input data when evaluating multiple alternative hypotheses; they notice “suspicious coincidences” when making inferences and generalizations. Word-learning studies have shown that preschoolers who are shown three dalmatians as exemplars of the novel (nonsense) word *zav* would only generalize the word to the subordinate-level kind *dalmatian*, even though the evidence would be equally consistent with the basic-level kind *dog*. Why is this inference rational? The key intuition behind it is that if the “teacher” had meant to show the preschooler examples in order to teach him or her the extension of the word *dog*, it would have been odd that the first three examples were all dalmatians (a subordinate-level kind). If the “teacher” had meant to teach the preschooler the extension of the word *dalmatian*, of course there would be nothing suspicious about showing three examples of the subordinate-level kind *dalmatian* (see Xu & Tenenbaum, 2007a, for formal modeling that captures this intuition).

Similar effects have been observed in infants. Gerken (2006) found that if 9-month-old infants heard a string of three-syllable sequences that had varying first and second syllables but all ended with the syllable *da*, the infants would infer that the rule for acceptable sequences was that they included two slots that could be filled with any syllable and a fixed third syllable *da*. Gweon, Tenenbaum, and Schulz (2010) found that this same inferential mechanism operates on 16-month-olds’ inductive inferences about the internal, nonobvious properties of objects. If an agent sampled three yellow toys from a box of mostly blue toys and showed that the yellow toys squeaked, infants generalized the property narrowly (i.e., only to other yellow toys and not to blue ones). Infants were also sensitive to sampling conditions and sample size in property generalization: Intentionally drawn samples led to this narrow generalization, whereas accidentally drawn samples did not. This pattern parallels the findings on sensitivity to sampling conditions and sample size in word learning with preschoolers (Xu & Tenenbaum, 2007b).

What is Meant by “Constructivist”?

The key idea is that infants may start with perceptual (and perhaps protoconceptual) primitives, and they can acquire new concepts and new inductive biases given input data. The newly acquired concepts and learned inductive biases become the elements that constrain subsequent learning and computations. What are the specific ways in which infants have been shown to be constructivist learners?

First, infants engage in hypothesis testing: They entertain multiple hypotheses at once and revise their beliefs as new evidence comes in. In the rule-learning experiment discussed above, 9-month-old infants were given initial exemplars of the rule (e.g., *leledi, wiwidi, jijidi, dededi*) that were equally consistent with both a narrow generalization (any AA*di* triplets with the same final syllable) and a broad generalization (any AAB triplets). When tested, the infants appeared to have learned a narrow rule. When given just one piece of additional evidence that the broad rule was correct, however, these young infants immediately switched to a pattern of responses that was consistent with the broad rule (Gerken, 2010). These findings provided the first demonstration that infants may hold multiple hypotheses in mind and they revise their beliefs when given new evidence.

Second, infants notice anomalous data: Nonrandom sampling may provide a strong cue to learners that a new causal variable is called for. Kushnir, Xu, and Wellman (2010) showed 20-month-olds a box containing 82% toy ducks and 18% toy frogs, and the infants watched as a person picked out five frogs in a row. These infants attributed to the person a preference for frogs over ducks (Fig. 2). This was not just a matter of having seen the person interact with the toy frogs: In a control condition in which the box contained 82% toy frogs and 18% toy ducks and the person picked out five frogs in a row, no preference was attributed to the person. It may be the case that nonrandom sampling is particularly noticeable and that learners are triggered by this type of data to posit new causal variables.

Third, statistical evidence may drive the process of acquiring new concepts. For example, infants are known to assume that everyone else shares their preferences for crackers over broccoli until they are about 18 months old (Repacholi & Gopnik, 1997). In a mini-training study, 16-month-old infants were shown two small bowls of objects, one boring (e.g., white cubes) and the other much more interesting (e.g., colorful mini-slinkies), and were allowed to choose an object from either bowl. Unsurprisingly, the infants chose a slinky for themselves and for an experimenter. The infants were then shown a transparent jar with lots of slinkies and a few white cubes. The experimenter reached in, picked out five white cubes, and asked the infant again to give her what she would like (the choices were between a slinky and a cube). This time, the infants handed her the cube. As before,

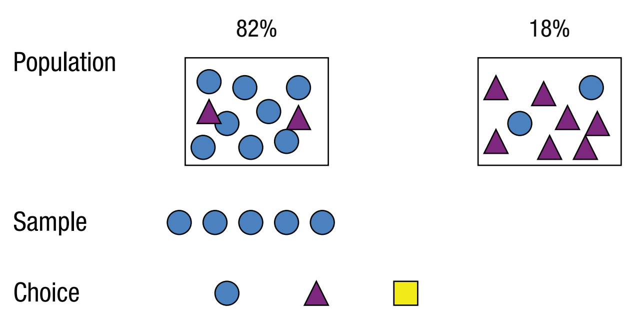

Fig. 2. Schematic representation of the experimental procedure in Kushnir, Xu, and Wellman (2010). One group of toddlers was shown a population box containing 82% toy frogs (represented by circles in the illustration) and 18% toy ducks (represented by triangles); a second group of toddlers was shown a population box containing 18% toy frogs and 82% toy ducks. The experimenter looked inside the box and picked out a sample of five frogs in a row. Toddlers were then presented with three choices–a toy frog, a toy duck, and a new toy (distractor)–and were asked to give the experimenter the toy she liked. Children gave the experimenter the toy frog when the box contained 18% toy frogs but not when the box contained 82% frogs.

this result depended on the proportions in the populations of objects; infants inferred that the experimenter had a preference for cubes only when the box contained a minority of cubes (18%), not when it contained a majority of cubes (82%). With minimal training, infants were able to acquire what was perhaps the beginning of a new concept—in this case, subjective preference (Ma & Xu, 2011).

Fourth, overhypothesis formation (i.e., the ability to make inferences at multiple levels) may be a powerful mechanism for acquiring new inductive constraints. The philosopher Nelson Goodman (1955) first introduced this idea: Suppose you are shown several bags. From the first bag, a few blue marbles are drawn out; from the second bag, a few red marbles are drawn out; from the third bag, a few yellow marbles are drawn out. Now you are shown a fourth bag, and one green marble is drawn out. If I drew another marble out of the fourth bag, what do you think its color would be? With high confidence, people answer "green." Not only do we make the first-order generalization that the first bag most likely contained all blue marbles, and that the second bag all red marbles, and so forth, we also form a second-order generalization or overhypothesis that "bags of marbles are uniform in color" (see Kemp, Perfors, & Tenenbaum, 2007, for formal modeling). This simple example illustrates a powerful idea: Learners make inferences and generalizations at multiple levels, and this allows inductive learning to proceed rapidly with limited amounts of data.

Recent evidence from 9-month-old infants suggests that they can form overhypotheses involving perceptual variables such as shape or color (Dewar & Xu, 2010). Infants were shown a set of boxes. From the first box, the experimenter drew out, with her eyes closed, four small objects of the same shape—say, triangles—but of different colors. From the second box, she drew out four cubes of different colors; from the third box, she drew out four discs of different colors. Then, from the fourth box, she drew out one star. She then drew a second object from the fourth box, either another star of a different color or a triangle of a different color. Infants looked reliably longer at the triangle than the star. Critically, control conditions showed that if the triangle had been drawn from the first box (from which all the other triangles had been drawn), infants looked equally at both outcomes. The same pattern of results was obtained with color. Thus, even preverbal infants can form overhypotheses with limited amounts of data, which suggests that this may be a powerful mechanism that supports later learning (see Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002, and Sim, Yuan, & Xu, 2011, for evidence in the domain of word learning).

Finally, infants are active learners (see Piaget, 1954, for discussion). Eight-month-old infants voluntarily allocate their attention according to information gain: They avoid spending too much time looking at stimuli that are too predictable or too unpredictable; instead, they focus their attention on stimuli that are intermediately predictable (i.e., stimuli with the most potential for information gain; Kidd, Piantadosi, & Aslin, 2012). The studies by Denison and Xu (2010b) also provided evidence that 10-month-old infants spontaneously estimate the probabilities of events and use this knowledge to fulfill their own desires and wishes. Thus, even barely mobile, seemingly passive infants are active learners.

In just a few years, research has uncovered a set of powerful inductive-learning mechanisms in infants and children, and the evidence provides strong support for the emerging view that young humans are rational, constructivist learners. The specific mechanisms identified in all of these studies appear to be domain general, applying to language learning, physical reasoning, psychological reasoning, property induction, and causal learning. In this article, we have selectively reviewed the evidence from studies with infants because the learning mechanisms developed in infancy lay the foundation for later learning.

Future Directions

Many questions remain open. First, how sophisticated are infants' probabilistic-inference abilities? The existing studies have only scratched the surface. Second, if young infants engage in hypothesis testing, where do the hypotheses come from? Work on children's explanations may shed light on this issue, given that explanation is a form of hypothesis generation (Keil, 2006). Third, the experiments reviewed here suggest that with limited amounts of evidence, infants and young children can revise their beliefs and acquire new concepts. But much learning in childhood takes place on a much larger timescale, and the conceptual changes that result from such learning are much more profound (Carey, 1985, 2009; Gopnik & Meltzoff, 1997). These conceptual changes may involve tracking statistical evidence over time and evaluating evidence across subdomains (e.g., the development of intuitive biological knowledge may require integrating evidence for growth, internal organs, and birth). It is an open question whether the same underlying processes can explain these long-term changes and developments.

Recommended Reading

Perfors, A., Tenenbaum, J. B., Griffiths, T., & Xu, F. (2011). (See References). A nontechnical introduction to Bayesian models and their relevance to cognitive development.

Teglas, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). (See References). A seminal study on how infants view physical reasoning as probabilistic inference, integrating their knowledge about object motion with probability estimations.

Xu, F., & Garcia, V. (2008). (See References). A study demonstrating young infants' sensitivity to probabilistic relations.

Xu, F., & Kushnir, T. (Eds.). (2012). *Rational constructivism in cognitive development: Advances in child development and behavior* (Vol. 43). Waltham, MA: Academic Press. An edited book providing the first synthesis on the rational-constructivist approach to the study of developmental psychology.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

References

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Carey, S. (2009). *The origins of concepts*. New York, NY: Oxford University Press.

Chater, N., & Oaksford, M. (2008). *The probabilistic mind: Prospects for Bayesian cognitive science*. New York, NY: Oxford University Press.

Chomsky, N. (1987). *Language and problems of knowledge*. Cambridge, MA: MIT Press.

Denison, S., Reed, C., & Xu, F. (in press). The emergence of probabilistic reasoning in very young infants: Evidence from 4.5- and 6-month-old infants. *Developmental Psychology*.

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

Denison, S., & Xu, F. (2010b). Twelve- to 14-month-old infants can predict single-event probability with large set sizes. *Developmental Science*, *13*, 798–803.

Dewar, K., & Xu, F. (2010). Induction, overhypothesis, and the origin of abstract knowledge: Evidence from 9-month-old infants. *Psychological Science*, *21*, 1871–1877.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective*. Cambridge, MA: MIT Press.

Fodor, J. (1981). *Modularity of mind*. Cambridge, MA: MIT Press.

Gerken, L. (2006). Decisions, decisions, decisions: Infant language learning when multiple generalizations are possible. *Cognition, 98*, B67–B74.

Gerken, L. (2010). Infants use rational decision criteria for choosing among models of their input. *Cognition, 115*, 362–366.

Goodman, N. (1955). *Fact, fiction, and forecast*. Cambridge, MA: MIT Press.

Gopnik, A., & Meltzoff, A. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences*, *14*, 357–364.

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, USA*, *107*, 9066–9071.

Keil, F. (2006). Explanation and understanding. *Annual Review of Psychology*, *57*, 227–254.

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science*, *10*, 307–321.

Kidd, C., Piantadosi, S., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE*, *7*(5), e36399.

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

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

Perfors, A., Tenenbaum, J. B., Griffiths, T., & Xu, F. (2011). A tutorial introduction to Bayesian models of cognitive development. *Cognition*, *120*, 302–321.

Piaget, J. (1954). *The construction of reality in the child*. New York, NY: Routlege.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 33*, 12–21.

Sim, Z., Yuan, S., & Xu, F. (2011). *Acquiring word learning biases*. In L. Carlson, C. Hoelscher, & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2544–2549). Austin, TX: Cognitive Science Society.

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.

Spelke, E. S. (1994). Initial knowledge: Six suggestions. *Cognition*, *50*, 431–445.

Teglas, E., Girotto, V., Gonzalez, M., & Bonatti, L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proceedings of the National Academy of Sciences, USA*, *104*, 19156–19159.

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

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, *331*, 1279–1285.

Xu, F. (2007). Rational statistical inference and cognitive development. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: foundations and the future* (Vol. 3, pp. 199–215). New York, NY: Oxford University Press.

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

Xu, F., Dewar, K., & Perfors, A. (2009). Induction, overhypotheses, and the shape bias: Some arguments and evidence for rational constructivism. In B. M. Hood & L. Santos (Eds.), *The origins of object knowledge* (pp. 263–284). New York, NY: Oxford University Press.

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

Xu, F., & Griffiths, T. (2011). Probabilistic models of cognitive development: Towards a rational constructivist approach to the study of learning and development. *Cognition*, *120*, 299–301.

Xu, F., & Tenenbaum, J. B. (2007a). Word learning as Bayesian inference. *Psychological Review*, *114*, 245–272.

Xu, F., & Tenenbaum, J. B. (2007b). Sensitivity to sampling in Bayesian word learning. *Developmental Science*, *10*, 288–297.