Cognition 130 (2014) 335–347

Contents lists available at ScienceDirect

Cognition

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

The origins of probabilistic inference in human infants

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

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

ARTICLE INFO

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

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

ABSTRACT

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

1. Introduction

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

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

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

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

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

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

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

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

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

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