Fuzzy universality of probability judgment

Valerie F. Reyna¹ and Charles J. Brainerd

Human Neuroscience Institute, Cornell University, Ithaca, NY 14853

Probability judgment is essential to navigating through life, from understanding risks to making decisions under uncertainty. However, probability theory is relatively recent. How has humanity managed to get along before this development? In PNAS, Fontanari et al. (1) provide exciting new evidence of probabilistic cognition among Mayan indigenous groups, both adults and children, whose performance was generally indistinguishable from that of Western adults.

On their face, these findings clash with the extensive literature on biases in probability judgment in which Westerners' judgments violate elementary rules of probability theory (2). Even statisticians who are familiar with the probability calculus systematically violate its rules (3). Nevertheless, Mayans without formal education succeeded at tasks such as judging proportions, combining probabilities, and updating beliefs that college-educated people fail. That failure occurs in simple

choice tasks that require neither numerical estimates nor computation. The mystery deepens when we consider the developmental origins of probability judgment: paradoxically, Fontanari et al. claim that infants make accurate probabilistic predictions in tasks that closely resemble those that older preschoolers fail. Why are infants so smart and statisticians so dumb?

Order can be brought to this seeming chaos by considering successive generations of theories of the development of probability judgment, which have each ruled out compelling misconceptions. Many studies show that young children pass "probability-judgment" tasks that only require discrimination of greater-or-lesser magnitude, have clear and memorable instructions, and do not require memory for problem information (i.e., they use external stores) (4–6).

Early studies often had confounded tasks that made them easier to pass. These tasks could be passed without understanding probability by simply picking whichever option had more of the winning color or tokens, a relative-magnitude judgment. Although understanding ratios is not the only criterion for understanding probability, it has long been considered an essential criterion (7, 8). Only the tasks shown in the right panel of figure 3 B and C of Fontanari et al. satisfy this criterion: a larger ratio but smaller number of winning tokens is pitted against a smaller ratio but larger number. Choosing 9/12 over 12/48 reflects understanding the importance of both numerator and denominator in probability judgments, a harder test to pass than relative-magnitude judgments. Thus, substituting relative magnitude for true probability judgments presumably accounts for high performance in some tasks in Fontanari et al.

However, the posterior and updating tasks go beyond simple magnitude judgments because they involve responding to changes. Educated Westerners and even professionals fall prey to well-known errors in judging posterior and updated probabilities. For example, told that 10% of patients have a disease (the prior) and a patient tested positive for the

Fig. 1. Illustrations of effects of formatting on probability judgment for lower (Left) and higher (Right) numerosities of equivalent probabilities. In A and B, the items are ordered, but in C and D, they are scrambled. Harder problems with both ratios less than one half are shown. Children and sometimes adults prefer box B because it has more winners (numerosity effect or denominator neglect). Higher numerosity and ordered (as opposed to scrambled) arrangements (B) facilitate gist pop-out: the ratio for box A is greater (vertical gist) but the relative magnitude of winners is greater for B (horizontal gist). E and F illustrate effects of making sets discrete: 1–2 odds are more clearly better than 2–5 odds, compared with 1/3 vs. 2/7.

Author contributions: V.F.R. and C.J.B. wrote the paper.

The authors declare no conflict of interest.

See companion article on page 17075.

¹To whom correspondence should be addressed. Email: vr53@cornell.edu.

disease given a test with 80% accuracy, only 32% of 82 physicians realized that the posterior probability was closer to 30% than 70% (9). Contrary to initial reports, frequency formats do not reduce such errors (10).

Current theory reconciles the apparent conflict between Fontanari et al.'s results and the larger literature by pinpointing the source of the difficulty in posterior and updating tasks as coming from confusion about overlapping sets (11). Disentangling overlapping sets for subjects, as is done in their cardboard displays, reduces errors substantially, as predicted by carefully tested theory (12). Thus, base-rate neglect as illustrated in our example with physicians and other forms of posterior probability biases do not reflect a lack of understanding of probability but, rather, are about the mechanics of keeping sets of events distinct in working memory (13). Displays work with both children and adults as long as the displays keep sets separate, revealing underlying probabilistic (ratio) competence that emerges around 5–6 y of age (14).

In particular, fuzzy-trace theory (FTT) demystifies the paradox in the probability-judgment literature of competent children growing up to be incompetent adults. FTT explains how some specific task formats confuse overlapping sets, whereas others make sets distinctive and, thus, why performance of adults can look worse than that of children who are given different tasks (Fig. 1). In addition, FTT accounts for results showing that advanced probabilistic understanding is intuitive: emphasis on verbal explanations was discarded after the first generation of theories because it was found to be an unreliable gauge of understanding (4). Subsequent research showed that preverbal intuition could reflect advanced understanding (15, 16), but some theories continued to emphasize numbers and computation. However, consistent with Fontanari et al.'s findings, ostensibly numerical tasks such as probability

judgment are typically solved nonnumerically using qualitative gist (17).

For example, comparing 9/12 (box A) to 12/48 (box B), children and adults extract the qualitative gist of relations among sets, such as fewer winners in box A than B ($9 < 12$), mostly win for box A ($9/12 > \text{about half}$), and mostly lose for box B ($12/48 < \text{about half}$) (14). Note that the relative magnitude of numerators erroneously favors box B, called the numerosity effect or denominator neglect, which even adults are prone to when denominators are similar and overlapping sets are confusing (18). However, children are able to pass critical tests of ratio competence, such as 9/12 (box A) vs. 12/48 (box B), when the two ratios fall above vs. below a half (5, 14). Thus, the high performance of all subject groups for the ratio task depicted in figure 3C of Fontanari et al. can be reconciled with the literature on numerosity effects (and denominator neglect) because the ratios differ categorically (likely vs. unlikely). Typical tests of numerosity effects compare two low or two high probabilities (18).

In sum, although earlier theories of probability judgment emphasized formal rules and numerical computation as the zenith of development, Fontanari et al.'s results are consistent with contemporary theory. FTT emphasizes "approximate comparison over precise assessment of possibilities," placing gist-based intuition at the center of advanced cognition (ref. 1, p. 17077). Given a clear task, it is no wonder that prenumerate groups without formal schooling can perform subtle probabilistic judgments using qualitative gist. Westerners generally perform these tasks in the same way using nonverbal intuition. Performance of children and adults is modulated by predictable features of the task, notably interference from nested or overlapping sets that must be inhibited; otherwise, reasoners tend to neglect denominators and focus only on relative magnitudes of numerators. Together, these theoretical principles account for the sometimes perplexing heterogeneity of findings regarding probability judgment.

- 1 Fontanari L, Gonzalez M, Vallortigara G, Girotto V (2014) Probabilistic cognition in two indigenous Mayan groups. *Proc Natl Acad Sci USA* 111:17075–17080.
- 2 Kahneman D (2003) A perspective on judgment and choice: Mapping bounded rationality. *Am Psychol* 58(9):697–720.
- 3 Tversky A, Kahneman D (1971) Belief in the law of small numbers. *Psychol Bull* 76(2):105–110.
- 4 Brainerd CJ (1981) Working memory and the developmental analysis of probability judgment. *Psychol Rev* 88(6):463–502.
- 5 Reyna V, Brainerd C (1994) The origins of probability judgment: A review of data and theories. *Subjective Probability*, eds Wright G, Ayton P (Wiley, Oxford, UK), pp 239–272.
- 6 US Dept. of Education (2008) *Foundations for Success: The Final Report of the National Mathematics Advisory Panel* (US Dept. of Education, Washington, DC). Available at www2.ed.gov/about/bdscomm/list/mathpanel/report/learning-processes.pdf.
- 7 Acredolo C, O'Connor J, Banks L, Horobin K (1989) Children's ability to make probability estimates: Skills revealed through application of Anderson's functional measurement methodology. *Child Dev* 60(4):933–945.
- 8 Hoemann HW, Ross BM (1982) Children's concepts of chance and probability. *Children's Logical and Mathematical Cognition*, ed Brainerd C (Springer-Verlag, New York).
- 9 Reyna VF (2004) How people make decisions that involve risk: A dual-processes approach. *Curr Dir Psychol Sci* 13(2):60–66.
- 10 Barbey AK, Sloman SA (2007) Base-rate respect: From statistical formats to cognitive structures. *Behav Brain Sci* 30(3):287–297.
- 11 Reyna VF, Brainerd CJ (2008) Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learn Individ Differ* 18(6):89–107.
- 12 Wolfe CR, Reyna VF (2010) Semantic coherence and fallacies in estimating joint probabilities. *J Behav Decis Making* 23(2):203–223.
- 13 Johnson-Laird PN (2010) Mental models and human reasoning. *Proc Natl Acad Sci USA* 107(43):18243–18250.
- 14 Spinillo AG, Bryant P (1991) Children's proportional judgments: The importance of "half." *Child Dev* 62(3):427–440.
- 15 Anderson NH (1996) *A Functional Theory of Cognition* (Erlbaum, Hillsdale, NJ).
- 16 Huber BL, Huber O (1987) Development of the concept of comparative subjective probability. *J Exp Child Psychol* 44(3):304–316.
- 17 Reyna VF, Brainerd CJ (2011) Dual processes in decision making and developmental neuroscience: A fuzzy-trace model. *Dev Rev* 31(2–3):180–206.
- 18 Garcia-Retamero R, Galesic M, Gigerenzer G (2010) Do icon arrays help reduce denominator neglect? *Med Decis Making* 30(6):672–684.