Cognition 168 (2017) 164–175

Contents lists available at ScienceDirect

Cognition

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

Original Articles

Children's quantitative Bayesian inferences from natural frequencies and number of chances

Stefania Pighin [a,*], Vittorio Girotto [b], Katya Tentori [a]

[a] CIMeC, University of Trento, Italy
[b] DCP, University IUAV of Venice, Italy

ARTICLE INFO

Article history:
Received 18 January 2017
Revised 16 June 2017
Accepted 27 June 2017
Available online 7 July 2017

Keywords:
Children
Chances
Natural frequencies
Frequency format
Probabilistic reasoning
Bayesian reasoning

ABSTRACT

Zhu and Gigerenzer (2006) showed that an appreciable number of Chinese children aged between 9 and 12 years old made correct quantitative Bayesian inferences requiring the integration of priors and likelihoods as long as they were presented with numerical information phrased in terms of natural frequencies. In this study, we sought to replicate this finding and extend the investigation of children's Bayesian reasoning to a different numerical format (chances) and other probability questions (distributive and relative). In Experiment 1, a sample of Italian children was presented with the natural frequency version of five Bayesian inference problems employed by Zhu and Gigerenzer (2006), but only a tiny minority of them were able to produce correct responses. In Experiment 2, we found that the children's accuracy, as well as the coherence between their probability judgments, depended on the type of question but not on the format (natural frequency vs. chance) in which information was presented. We conclude that children's competence in drawing quantitative Bayesian inferences is lower than suggested by Zhu and Gigerenzer (2006) and, similarly to what happens with adults, it relies more on a problem representation that fosters an extensional evaluation of possibilities than on a specific numerical format.

1. Introduction

What kinds of probabilistic inferences emerge at earlier ages and what others require, instead, more time to be mastered? When are children able to engage in proper Bayesian reasoning? Does this depend on the way in which the relevant information is presented?

Converging empirical evidence indicates that preverbal infants (Téglás, Girotto, Gonzalez, & Bonatti, 2007; Téglás et al., 2011) as well as nonhuman primates (Rakoczy et al., 2014) possess at least some basic probabilistic intuitions that enable them to make *implicit probability inferences* like those necessary to determine which of two elementary events is more likely to occur. For example, looking time suggests that 12-month-old infants correctly expect a yellow ball, rather than a blue one, to exit from a container in which three yellow balls and only one blue ball are bouncing (Téglás et al., 2011). Probabilistic competence has also been claimed to guide optimal choices already in 10- to 14-month-old infants (Denison & Xu, 2010, 2014), even if this result has been replicated only with children older than 4 years (Girotto, Fontanari, Gonzalez, Vallortigara, & Blaye, 2016).

Explicit qualitative probabilistic inferences like those required to make predictions in accordance with prior probability have been documented at a later stage of development (Brainerd, 1981; Yost, Siegel, & Andrews, 1962; Téglás et al., 2007/Study 3; see also Sobel, Tenenbaum, & Gopnik, 2004 and Griffiths, Sobel, Tenenbaum, & Gopnik, 2011 for similar investigations concerning children's capacity to make probabilistic inferences about the causal properties of objects). For example, Girotto and Gonzalez (2008) showed that 5-year-old children correctly judged that a black token was more likely than a white one to be drawn from an opaque bag containing four black round tokens, one black square token, and three white square tokens. At the same age, children are also able to update their judgments based on new evidence. In the above example, if the children were told the shape of the token (before the outcome of the extraction was revealed), they proved able to integrate this new piece of information into their judgment: that is, when the extracted token was round, they kept betting on black, while when it was square, they changed their bet to white. Similar results have been obtained with preliterate and prenumerate indigenous Maya groups living in remote areas of Guatemala (Fontanari, Gonzalez, Vallortigara, & Girotto, 2014). The findings that 5-year-old children as well as preliterate and prenumerate individuals can make sound

* Corresponding author at: CIMeC, University of Trento, Corso Bettini, 31, 38068 Rovereto (TN), Italy.
E-mail address: stefania.pighin@unitn.it (S. Pighin).

http://dx.doi.org/10.1016/j.cognition.2017.06.028
0010-0277/

qualitative Bayesian inferences based on priors and/or likelihoods strongly suggest that such inferences do not depend on formal education. This is also coherent with the findings that abstract knowledge of number and basic numerical skills (i.e., comparing and adding numerical quantities) precede (Barth, La Mont, Lipton, & Spelke, 2005; Barth et al., 2006) and are independent from (Butterworth, Reeve, Reynolds, & Lloyd, 2008; Pica, Lemer, Izard, & Dehaene, 2004) schooling.

Before the acquisition of symbolic number knowledge, young children therefore seem able to make (implicit or explicit) qualitative probabilistic inferences by considering and comparing the various ways in which an outcome may occur (*extensional reasoning*, Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni, 1999). Such judgments express a more-less relationship between two quantities, and may rely on a bare enumeration of relevant possibilities without requiring precise calculations. *Quantitative inferences* are more complex than qualitative ones because they involve the ability to manipulate and estimate exact relative amounts (*proportional reasoning*). Accordingly, much of the research that has been conducted on children's capacity to calculate probabilities has been confined almost exclusively to that ability (for a review, see Bryant & Nunes, 2012). Despite minor disagreements, problems that require proportional reasoning prove to be hard for children up to the age of roughly 10 years, both when numerical information concerns odds (e.g. 1:2) and, to a greater extent, conventional fractions (e.g., 1/3; see Fischbein & Gazit, 1984; Fujimura, 2001; Noelting, 1980; Nunes & Bryant, 1996; Pitkethly & Hunting, 1996; Schwartz & Moore, 1998).

The present study focuses on the developmental course of children's performance in a more sophisticated reasoning ability. In particular, we are interested in understanding when children become able to draw proper *quantitative Bayesian* inferences which involve a numerical integration between different values, as priors and likelihoods. As far as we can tell from the literature, Zhu and Gigerenzer (2006) is the only study which has investigated this issue. A sample of children attending an ordinary elementary school in Beijing (China) was presented with a number of Bayesian problems whose content was suited to children. There follows an example (p. 289):

> Pingping goes to a small village to ask for directions. In this village, 10 out of every 100 people will lie. Of the 10 people who lie, 8 have a red nose. Of the remaining 90 people who don't lie, 9 also have a red nose. Imagine that Pingping meets a group of people in the village with red noses. How many of these people will lie? ___ out of ____.

When problems were phrased in terms of frequencies (like the one above), the rate of Bayesian responses (averaged across two experiments and weighted by their sample sizes) provided by fourth, fifth, and sixth graders (aged between 9 and 12) were 18.7%, 39%, and 53.5%, respectively (p. 294). However, no child could provide any Bayesian response when problems were phrased in terms of percentages. Zhu and Gigerenzer (2006) interpreted their results as supporting the hypothesis that the human mind is not designed for probabilities or percentages, but needs natural frequencies[1] to make sound Bayesian inferences (Gigerenzer, 1996; Gigerenzer & Hoffrage, 1995, 1999).

Zhu and Gigerenzer's (2006) study has had a considerable impact on the literature. The result that a substantial number of children aged between 9 and 12 years old make sound Bayesian inferences when provided with natural frequency but not single-event probability information has been interpreted as crucial evidence supporting the claim that natural frequencies enable humans to reason the Bayesian way, while other formats prevent it (Brase, 2008; Galesic, Gigerenzer, & Straubinger, 2009; Gigerenzer, 2015; Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2008). However, to our knowledge, these results have never been replicated.[2]

Given the importance of Zhu and Gigerenzer's (2006) results not only from a developmental perspective but also for their more general implications in regard to reasoning research, confirmatory experimental evidence is conspicuous by its absence. All the more so because the success rate of Zhu and Gigerenzer's (2006) Chinese sixth (and, to a certain extent, fifth) graders exceeds that typically found with naïve and even most educated Western adults (Bramwell, West, & Salmon, 2006; Girotto & Gonzalez, 2001; Hoffrage & Gigerenzer, 1998). For example, in Bayesian problems phrased in terms of natural frequency, accuracy rates of 46% have been reported with Austrian university students (Gigerenzer & Hoffrage, 1995) and experienced German physicians (Hoffrage & Gigerenzer, 1998), 31% with US undergraduates (Sloman, Over, Slovak, & Stibel, 2003, Exp. 1b), below 25% with patients at Spanish hospitals (Garcia-Retamero & Hoffrage, 2013), 2% with US adults recruited using Amazon Mechanical Turk (Pighin, Gonzalez, Savadori, & Girotto, 2016), and 0% with UK midwifes recruited at training events or in maternity wards (Bramwell et al., 2006). Our Experiment 1 was therefore aimed at replicating Zhu and Gigerenzer's (2006) study concerning problems phrased in a natural frequency format, using a different sample of children of the same age (i.e., fourth, fifth, and sixth graders) from another country (Italy).

We also wanted to empirically assess Zhu and Gigerenzer's (2006) conclusion about the advantage of natural frequencies over single-event probabilities on children's Bayesian reasoning. The hypothesis that frequencies could facilitate probabilistic reasoning has a long (and not always linear) history. It originated from the observation that the rate of some well-known biases (e.g., the conjunction fallacy) reduced when the problems were framed in terms of frequencies (Cosmides & Tooby, 1996; Gigerenzer, 1991; Gigerenzer & Hoffrage, 1995, 2007; Gigerenzer, Todd, & ABC Research Group, 1999, but a preliminary version of this hypothesis was already put forward in Tversky & Kahneman, 1983, p. 309). Such a hypothesis, however, has been increasingly challenged by a growing body of evidence showing that frequencies are not inherently easier to process than percentages (Cuite, Weinstein, Emmons, & Colditz, 2008; Waters, Weinstein, Colditz, & Emmons, 2006) and that, once various confounding factors have been eliminated, their advantage disappears as well (Evans, Handley, Perham, Over, & Thompson, 2000; Reyna & Brainerd, 2008; Sloman et al., 2003; Tentori, Bonini, & Osherson, 2004). The advocates of the frequentist hypothesis have rejected these results because they were obtained with problems in which frequencies were typically normalized, and they have reiterated their position with respect to Bayesian updating problems (like, for example, the

[1] Note that, as pointed out by an anonymous reviewer of this paper, the term *natural* does not reflect any biological claim, but only a speculation about an alleged evolutionary advantage. According to this, natural frequencies constitute a cognitively privileged format because they represent the outcomes of the process of counting and classifying the occurrences of events as they are experienced (*natural sampling*, Kleiter, 1994). On the other hand, the human mind "would not be tuned to probabilities or percentages as input format" (Gigerenzer & Hoffrage, 1995, p. 686; but see also Gigerenzer, 1996), because these do not correspond to the typical way in which humans have dealt with statistical information over their evolution.

[2] Luecking (2004) and Multmeier (2012) have been repeatedly mentioned (e.g., by Gigerenzer, 2008; Gigerenzer, 2015; Martignon & Kuntze, 2015) as supporting Zhu and Gigerenzer's (2006) conclusions. However, both these studies are unpublished. Multmeier (2012) is available online, and we note that the data reported therein are only partially coherent with those of Zhu and Gigerenzer (2006). Indeed, Multmeier's first experiment, the only one whose stimuli and participants are directly comparable with those of Zhu and Gigerenzer, did not involve fifth or sixth graders, and the accuracy rate of fourth graders seemed somewhat lower (13%) than that of Zhu and Gigerenzer's (2006) participants.