Cognition 131 (2014) 60–68

Contents lists available at ScienceDirect

Cognition

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

Apes are intuitive statisticians

Hannes Rakoczy [a,*], Annette Clüver [a], Liane Saucke [a], Nicole Stoffregen [a], Alice Gräbener [a], Judith Migura [a], Josep Call [b]

[a] Institute of Psychology & Courant Research Centre "Evolution of Social Behaviour", University of Göttingen, Germany
[b] Max-Planck-Institute for Evolutionary Anthropology, Department of Developmental and Comparative Psychology, Leipzig, Germany

ARTICLE INFO

Article history:
Received 13 March 2013
Revised 12 December 2013
Accepted 19 December 2013

Keywords:
Comparative psychology
Primate cognition
Intuitive statistics
Numerical cognition

ABSTRACT

Inductive learning and reasoning, as we use it both in everyday life and in science, is characterized by flexible inferences based on statistical information: inferences from populations to samples and vice versa. Many forms of such statistical reasoning have been found to develop late in human ontogeny, depending on formal education and language, and to be fragile even in adults. New revolutionary research, however, suggests that even preverbal human infants make use of intuitive statistics. Here, we conducted the first investigation of such intuitive statistical reasoning with non-human primates. In a series of 7 experiments, Bonobos, Chimpanzees, Gorillas and Orangutans drew flexible statistical inferences from populations to samples. These inferences, furthermore, were truly based on statistical information regarding the relative frequency distributions in a population, and not on absolute frequencies. Intuitive statistics in its most basic form is thus an evolutionarily more ancient rather than a uniquely human capacity.

1. Introduction

Much research has suggested that reasoning about probabilities develops late in ontogeny, depends on language and formal education (Piaget & Inhelder, 1975), remains fragile even in adulthood (Tversky & Kahneman, 1974, 1981), and only works under special circumstances (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995). Exciting new research, however, suggests that such reasoning capacities might well be in place in the absence of language. Even preverbal infants engage in some intuitive statistics: they expect randomly drawn samples to reflect the distribution in the population drawn from and vice versa (Denison & Xu, 2010b; Téglás, Girotto, Gonzalez, & Bonatti, 2007; Xu & Garcia, 2008).

This early intuitive statistics seems to be a cognitive capacity that is functionally integrated in humans with other cognitive domains from very early on: For example, infants already integrate information about physical and psychological background conditions into their statistical inferences when judging whether sampling processes are random or non-random. Regarding physical information, for example, infants understand that mechanical constraints (e.g. some kinds of objects in a population cannot be drawn physically in the same way as others) can turn a sampling process into a non-random one such that the sample need not reflect the distribution in the population (Denison & Xu, 2010a; Téglás et al., 2007). Moreover, statistical information is combined with geometrical and temporal information in rather systematic ways to form predictions about future events (Téglás et al., 2011). Regarding psychological information, infants appreciate that when a person draws from a population but has both a preference regarding the different kinds of objects in the population and visual access, her sampling will probably be non-random and her sample will thus not match the distribution of the population (Xu & Garcia, 2008).

* Corresponding author. Address: Department of Developmental Psychology & Courant Research Centre "Evolution of Social Behaviour", University of Göttingen, Waldweg 26, D-37073 Göttingen, Germany. Tel.: +49 551 399235.

E-mail address: hannes.rakoczy@psych.uni-goettingen.de (H. Rakoczy).

0010-0277/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cognition.2013.12.011

Conversely, infants draw inferences in the other direction, from statistical to psychological states of affairs: when confronted with a person who draws samples that are absolutely non-representative of the populations, infants assume the person must have informational access and corresponding preferences (Kushnir, Xu, & Wellman, 2010; Ma & Xu, 2011). Convergent findings have shown such intuitively statistical expectations in infants and toddlers with a number of different measures tapping different types of behaviors: looking time in response to violations of expectations (e.g. Xu & Garcia, 2008), active choice measures (of samples drawn from different populations; (Denison & Xu, 2010b) and actions directed towards others (such as giving them the kind of item they prefer; (Kushnir et al., 2010; Ma & Xu, 2011).

Compared to the information available for human infants, nothing is currently known about the phylogenetic origins and distributions of such intuitive statistics. We do not know how old evolutionarily capacities for intuitive statistics are, and we do not know whether they are shared by any non-human animals. Studies on optimal choice and foraging (Balci, Freestone, & Gallistel, 2009; Kamil, Krebs, & Pulliam, 1987; Stephens, 2008; Stüttgen, Yildiz, & Güntürkün, 2011) and numerosity discrimination (Brannon & Terrace, 1998; Hanus & Call, 2007) have demonstrated that non-human animals share with humans basic cognitive capacities to maximize the amount of food rewards on the basis of perceptual information. Those studies, however, lack some of the crucial features present in intuitive statistics research. Unlike subjects in optimal choice studies, subjects in intuitive statistics studies form expectations and select optimally based on statistical information without any prior training to associate the stimuli and their reinforcement contingencies or any other reliance on past sampling (Téglás et al., 2007, 2011).

It is true that such good first trial performance in the absence of training can also be found in primates' discrimination of absolute set sizes. First, from numerous comparative studies we know that many non-human animals, notably primates, share with humans an analog magnitude system that allows for the approximate discrimination between arbitrarily large sets (Cantlon & Brannon, 2006, 2007; Flombaum, Junge, & Hauser, 2005). The signature limit of this capacity, following Weber's Law, is constituted by the ratios of the sizes of two sets to be discriminated: if a subject can discriminate 4 from 8 objects, it can discriminate 10 from 20, 150 from 300, etc. Second, humans and other primates share an object individuation system that allows for the exact parallel individuation ("subitizing") of small sets (Hauser, MacNeilage, & Ware, 1996). The signature limit here is defined by the absolute set sizes: only sets smaller than 3 (infants) or 4 (monkeys and apes) can be discriminated, such as 1:2, 2:3, and 1:3 (see (Carey, 2009), for review). In contrast to such tasks, however, intuitive statistical problems crucially require representing truly statistical matters, namely relative rather than absolute frequencies – that is, frequencies of items of a given kind in a population (say, winner tickets in a lottery) relative to the frequencies of all kinds of items in the population (all tickets). It is thus an open question whether intuitive statistical reasoning, understood as the capacity to flexibly draw inferences from populations to samples and vice versa, is evolutionarily recent and uniquely human or evolutionarily ancient and shared with other animals. Here we report a series of studies that speaks to that question. These studies with our closest relatives, the great apes, investigated one of the most basic forms of such intuitive statistical capacities: the ability to draw inferences from information about a population to a randomly drawn sample. We used tasks modeled after those developed in recent infant studies (Denison & Xu, 2010b). In these tasks, subjects are confronted with two visible populations with different distributions of items of two kinds (one preferable over the other) and the experimenter randomly draws from each population a 1-object-sample that the subject cannot see. Subjects are then given a choice between the two samples. These tasks thus require the subjects, first, to distinguish between the two populations according to the ratios of the two kinds of objects in their distributions and, second, to form expectations about the probability of sampling events accordingly, that is, expectations as to which sample is more likely to contain an object of the more desirable kind.

Control experiments ruled out alternative explanations such as simpler choice heuristics (Exp. 2 and 3), Clever Hans effects (Exp. 5 and 6) and use of olfactory information (Exp. 7). Most importantly, two experiments (Exp. 4 and 6) tested whether such inferences were truly based on probability information and not just on information about absolute frequencies.

2. Experiment 1: inferences from populations to samples

2.1. Subjects

Participants in all experiments were recruited from a group of four species of Great Apes ($N = 33$; Female $N = 24$): Chimpanzees (*Pan troglodytes verus*, $N = 17$), Gorillas (*Gorilla gorilla*, $N = 5$), Orangutans (*Pongo pygmaeus*, $N = 6$) and Bonobos (*Pan paniscus*, $N = 5$) housed at the Wolfgang Köhler Primate Research Center (WKPRC) in the Leipzig Zoo. Mean age of animals was 16;10 (years; months) with a range of 6;2-30;6. About one third were hand-reared and the remaining two third**s** were mother-reared. All subjects had experience in cognitive studies and were used to receiving food-items as reinforcement (see SI Table 1 for a detailed description of the animals' demographics and background). 28 apes (15 Chimpanzees, 2 Gorillas, 6 Orangutans and 5 Bonobos) were included in the final sample of this experiment. Four further apes (2 Chimpanzees and 2 Gorillas) were tested but excluded from data analysis due to inconsistent item preference during the Preference Test ($N = 1$) or because they did not complete all trials due to lack of motivation ($N = 3$).

2.2. Design and procedure

Populations of banana pellets and carrot pieces were presented in two transparent buckets. Both buckets contained the same absolute amount of food items (80), with