Statistics: What Seems Natural?

Which statistical data seem easier to understand, 10 cases in 100, or 10%? In their Policy Forum "Communicating statistical information" (*Science*'s Compass, 22 Dec., p. 2261), U. Hoffrage and colleagues offer persuasive evidence that both experts and novices find it to be the former. When prevalence, sensitivity, and false positive rates are given as probabilities (e.g., 10%), most physicians misinterpret the information in a way that could be potentially disastrous for the patient, but when they are presented as "natural frequencies" (e.g., 10 cases in 100), the physicians' performance is dramatically better. The authors suggest ways to improve both communication of statistical information and medical education by using frequencies rather than probabilities.

The discussion by Hoffrage *et al*. leaves open the question as to why this is the case. Elsewhere, Gigerenzer and Hoffrage suggest that "humans seem to be developmentally and evolutionarily prepared to handle natural frequencies" (*1*, p. 430) by accumulating examples of the category in question. However, this would not, in itself, explain why this accumulation is preferentially represented as frequencies rather than being transformed into some other representation, such as rate or probability.

Frequencies (e.g., 10 cases in 100) can be thought of as a subcollection (with a numerosity of 10) in a collection (with a numerosity of 100). I have suggested that we are born with a specialized capacity for representing collections and their numerosities (2). The evidence for this comes from a range of studies showing that infants, even in the first week of life, are sensitive to changes in the numerosity of a collection of visual objects (3) and that, at 6 months, they are able to form arithmetical expectations on the basis of adding an object to a collection or taking it away (4). The almost universal use of fingers as the representative collection in counting and arithmetic suggests that collections and numerosities form the basis of later representations also (2). This suggestion has been supported by recent brain-imaging evidence showing that key number areas are closely connected to the finger circuit in the intraparietal sulci (5).

Of course, the big developmental gap between the capacities of young children and the performance of adult decision-makers is typically filled by an education system that teaches children about collections and numerosities far more than about probability. It is thus plausible that educational practices are, in part, responsible for the biases Hoffrage *et al*. report. However, there is indirect evidence that probability concepts are intrinsically difficult for humans. Although the computational techniques required by probabilities of the type described by the authors would have been available to the ancient Greeks, an understanding of the concepts began only with Girolamo Cardano's *Liber de ludo aleae* (1525, published in 1663) and in the correspondence between Pascal and Fermat about games of chance in 1654.

Brian Butterworth
Institute of Cognitive Neuroscience,
University College London,
Alexandra House,
17 Queen Square,
London WC1N 3AR, UK;
e-mail: b.butterworth @ucl.ac.uk

References and Notes

1. G. Gigerenzer, U. Hoffrage, *Psychol. Rev.* **106**, 425 (1999).
2. B. Butterworth, *What Counts: How Every Brain Is Hardwired for Math* (Free Press, New York, 1999) [in the UK, *The Mathematical Brain* (Macmillan, London, 1999)].
3. S. E. Antell, D. P. Keating, *Child Devel.* **54**, 695 (1983); P. Starkey, R. G. Cooper Jr., *Science* **210**, 1033 (1980); E. Van Loosbroek, A. W. Smitsman, *Devel. Psychol.* **26**, 916 (1990).
4. K. Wynn, *Nature* **358**, 749 (1992).
5. S. Dehaene, E. Spelke, P. Pinel, R. Stanescu, S. Tsivkin, *Science* **284**, 970 (1999); B. Butterworth, *Science* **284**, 928 (1999).

Response

Butterworth suggests that natural frequencies facilitate reasoning because people are born with a specialized capacity for processing collections of discrete objects, rather than probabilities and fractions. There is certainly evolutionary and developmental evidence supporting this explanation, but there is also a second, more specific reason: Bayesian computations are simpler when information is represented in natural frequencies than in probabilities, percentages, or relative frequencies (1, 2).

With natural frequencies, people can calculate the conditional probability of a hypothesis H (e.g., the occurrence of cancer) given data D (e.g., a positive test) simply:

where *a* is the natural frequency of people with cancer who tested positive and *b* is that of people without cancer who tested positive. In the colorectoral cancer example used in our Policy Forum, *a* equals 15 and *b* equals 300 people, respectively. Communicated this way, it is easy to see that 15 out of the 315 people who tested positive actually have cancer. In contrast, when the same information is communicated in terms of conditional probabilities, as is common practice, the calculation is complicated:

Equation 2 is known as Bayes' rule. In our example, applying this rule would require the nontrivial computation (0.003)(0.5)/[(0.03)(0.5) + (0.997)(0.03)]. The reason why natural frequencies facilitate Bayesian inference is because they retain information about base rates (e.g., of cancer), whereas conditional probabilities are normalized with respect to these base rates. As a consequence, the probabilities in Eq. 2 (0.5 and 0.03) need to be multiplied by the base rates (0.003 and 0.997) in order to reintroduce base rate information. In other words, natural frequencies facilitate Bayesian reasoning because part of the calculation is already "done" within the representation itself.

Ulrich Hoffrage,*
Max Planck Institute for Human Development,
Lentzeallee 94,
Berlin 14195, Germany;
Samuel Lindsey,
Department of Psychology,
University of Virginia,
Charlottesville, VA 22903, USA
Ralph Hertwig,
Max Planck Institute for Human Development,
Lentzeallee 94,
Berlin 14195, Germany;
Gerd Gigerenzer
Max Planck Institute for Human Development,
Lentzeallee 94,
Berlin 14195, Germany;

*To whom correspondence should be addressed.
E-mail: hoffrage@mpib-berlin.mpg.de

References and Notes

1. G. Gigerenzer, U. Hoffrage, *Psychol. Rev.* **102**, 684 (1995).
2. G. Gigerenzer, U. Hoffrage, *Psychol. Rev.* **106**, 425 (1999).