

End Notes

How I Got Started: Teaching Physicians and Judges Risk Literacy

GERD GIGERENZER*

Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Berlin, Germany

Like many unfortunate first-year psychology students, my first encounter with statistics began in a crowded lecture theater (in my case at the University of Munich), watching a flaky statistician scribble line after line of derivations on the blackboard. I dutifully copied the equations into my notebook, only to see him eventually reach for a sponge and wipe them away. We students had no clue what the equations were good for. We asked him to explain, but he knew only the math. Finally, after months of mutual suffering, he was fired. In the second semester, a psychology professor took over the job and began by hiring tutors. Thinking that teaching would be the best way to learn, I applied. At the interview, he showed me a chi-square equation and asked what it was. I had no idea, mistook the crudely drawn chi for an x, and responded that the equation decomposes an x on the left side into its additive components on the right side. That was good enough, I was hired. For many semesters, I was guilty of teaching other students what I naively believed to be statistical method, rejecting null hypothesis at the 5% significance level—a ritual I later called ‘mindless statistics’ (Gigerenzer, 2004).

All that changed thanks to a regulation at the University of Munich requiring doctoral candidates to undergo additional oral examinations in two external fields. Virtually all of my peers chose psychiatry and education. I, instead, chose statistics in the mathematics department (and, admittedly, education). Those three years in the company of statisticians were an eye-opener. I had believed that statistics is statistics, but now I learned the difference between statistical *thinking* and *rituals*. My eyes were opened even wider at the Center for Interdisciplinary Research in Bielefeld, as a member of a one-year research group studying the *Probabilistic Revolution*. From morning to midnight, I discussed the topic with historians of science, evolutionary biologists, physicists, economists, and statisticians (there was not much else to do in Bielefeld). You can find the exciting story about how probability changed science and everyday life in *The Empire of Chance* (Gigerenzer et al., 1989) and in *Cognition as Intuitive Statistics* (Gigerenzer & Murray, 1987). Once more, I learned that there is more to statistics than the math.

HOW TO MAKE COGNITIVE ILLUSIONS DISAPPEAR

One insight from the Probabilistic Revolution project proved highly relevant for my research. Probability, I’d discovered,

*Correspondence to: Gerd Gigerenzer, Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Berlin, Germany. E-mail: sek@gigerenzer@mpib-berlin.mpg.de

has multiple meanings: relative frequency (counting in order to measure probability), propensity (designing a die or slot machine to determine the probability), and degrees of belief (measured by one’s willingness to bet). And when I picked up a copy of *Judgment Under Uncertainty* (Kahneman, Slovic, & Tversky, 1982), published during my year in Bielefeld, I read the chapters from a different perspective than most of my peers.

One fascinating chapter was on overconfidence (Lichtenstein, Fischhoff, & Phillips, 1982). In this research, people were asked questions (e.g., ‘Which city has more inhabitants: Islamabad or Hyderabad?’) and subsequently asked to rate their confidence that their answer was correct. Surprisingly, the average confidence was higher than the average proportion of correct answers—a phenomenon called overconfidence, attributed to people’s self-delusion or other motivational deficits. My training, however, made me aware that the question asked was about a degree of belief (confidence) in a single event (the answer being correct or not), not about frequency. Thus, my colleagues and I conducted experiments in which we instead asked participants to estimate how many correct answers they gave. The result was another eye-opener. Overconfidence magically disappeared; estimated frequencies were as accurate as they could be. This result showed that what was called ‘overconfidence’ is not a matter of personality or motivation but is based on cognitive inference mechanisms, which we spelled out in the probabilistic mental models theory (Gigerenzer, Hoffrage, & Kleinbölting, 1991). Extending these findings to other so-called cognitive illusions (Gigerenzer, Fiedler, & Olsson, 2012), we sparked a fruitful controversy about the nature of cognitive illusions.

Now that I understood the key to making cognitive illusions disappear, the next step was to leave the lab in order to share this knowledge and show experts, not only laypeople, how to understand risks.

HOW TO MAKE PHYSICIANS RISK LITERATE

A most fertile discovery was that natural frequencies facilitate Bayesian reasoning. For years, some psychologists claimed that people are lousy Bayesians because they ignore base rates and labeled this the base rate fallacy. In experiments, individuals were as a rule presented with conditional probabilities, such as hit rates and false alarm rates, and generally floundered when asked to estimate the Bayesian posterior probability. In our experiments, however, Ulrich Hoffrage and I showed for the first time that natural frequencies help students to make Bayesian inferences

(Gigerenzer & Hoffrage, 1995, 1999; Kleiter, 1994). The problem was not simply, as had been suspected, in the human mind, but also in the representation of the information.

Inspired by a fascinating paper by David Eddy (1982), later consultant to the Clinton administration on healthcare reform, we began to systematically study physicians' ability to understand test results. If you test positive on a medical test, the physician should know how to estimate the probability that you actually have the disease (the 'positive predictive value'). Our first study showed, however, that most physicians were in the dark (Hoffrage & Gigerenzer, 1998). For instance, their estimates of the probability that a patient has colorectal cancer given a positive fecal occult blood screening test ranged from 1 to 99%! Only one out of 24 physicians estimated correctly. But when the information was presented in natural frequencies, most physicians now understood what the positive predictive value was.

In order to get this and other related techniques out into the medical community, my collaborators and I published both in top medical journals (e.g., Gigerenzer & Edwards, 2003; Wegwarth, Gaissmaier, & Gigerenzer, 2011) and in the low-impact-factor journals that doctors actually read. Even more effective in spreading the message were *Calculated Risks* (UK edition: *Reckoning With Risk*, Gigerenzer, 2002) and *Risk Savvy* (Gigerenzer, 2014). These two trade books reached many health care providers, who would never have been reached if I had published exclusively in major psychological journals.

Once interest was raised in the medical field, the next step was to improve doctors' risk literacy. Over the last decade, I have taught about 1000 physicians in their continuing medical education in risk literacy. These sessions covered natural frequencies versus conditional probabilities, relative versus absolute risks, single-event versus frequency statements, and five-year survival rates versus mortality rates. In each of these pairs, the first representation fosters insight, whereas the second supports innumeracy (Gigerenzer, 2014, Ch. 9 and 10). I should note that experienced physicians were one of the most appreciative audiences I have ever had.

I felt honored when natural frequencies began to be recommended by major medical organizations, including the Cochrane Collaboration, the International Patient Decision Aid Standards Collaboration, and the Medicine and Healthcare Products Regulatory Agency (the UK equivalent of the US Food and Drug Administration). It is a rare event that a concept from cognitive psychology makes its way into medicine.

A British investment banker, David Harding, once read *Reckoning With Risk*, bought copies for all of his 200 employees, and in 2007, over dinner, donated a generous sum to fund the *Harding Center of Risk Literacy* at the Max Planck Institute for Human Development in Berlin. At the opening ceremony, Harding jokingly said that he had earned part of his wealth thanks to public statistical illiteracy and now wanted to give some of it back to make the public literate. In the years since, the Harding Center has been influential in motivating the editors of health brochures and medical journals to stop using misleading statistics, which often confuse rather than inform doctors and patients (Gigerenzer & Muir Gray, 2011; Gigerenzer, Gaissmaier, Kurz-Milcke,

Schwartz, & Woloshin, 2007). For instance, in 2009 we helped the German Cancer Care write a new generation of pamphlets in which misleading relative risks and five-year-survival rates were axed and replaced by transparent absolute numbers. The center has also been involved in making statistical literacy and risk communication part of medical departments' core curricula.

TEACHING FEDERAL JUDGES

Doctors are not the only ones who receive inadequate training in understanding statistical evidence. Lawyers and judges get a mostly probability-free education. For that reason, at the O. J. Simpson trial, Alan Dershowitz, a renowned Harvard law professor who advised the defense team, could fool judges and jury about the probability that a man who had battered his wife (as Simpson had) actually murdered her (Gigerenzer, 2002, Chap. 8; Koehler, 1997). My first association with a law school was at the University of Virginia where I taught a course on 'How to understand probabilities.' The students were extremely smart and articulate, but shockingly innumerate, and worse, did not think that it mattered. Yet that attitude changed the moment they realized that understanding how to present statistics would give them a competitive edge—over the prosecution if they were defense lawyers, or vice versa.

Here is an example. You (or your client) have been accused of committing a murder and are standing before the court. There is only one piece of evidence against you, but a damning one: Your DNA matches the traces found on the victim. The court calls an expert witness to explain what that match means. The expert testifies the following:

'The probability that this match has occurred by chance is 1 in 100,000.'

You can already see yourself behind bars. Yet there are always two ways to communicate statistical information: one transparent and the other misleading. The aforementioned statement phrases the evidence in terms of a single-event probability (that this match occurred by chance), which is potentially misleading. In contrast, a frequency statement is transparent:

'Out of every 100,000 people, 1 will show a match.'

If you live in a city with two million adult inhabitants, one can expect 20 whose DNA would match the sample on the victim. On its own, this fact seems very unlikely to land you behind bars. Thus, if you ever stand trial, make sure that the evidence is communicated in frequencies, not in single-event probabilities. The British Court of Appeal recommended that DNA evidence be presented in a frequency format, and other courts have followed (Gigerenzer, 2002).

In 2004 and 2005, I taught risk communication to about 50 US federal judges in a continuing education program organized by the George Mason University School of Law. Today, however, both law and medical students still are not taught to understand evidence adequately. Psychologists

can offer effective tools for risk communication and teach experts how to use them (Hoffrage, Kurzenhäuser, & Gigerenzer, 2000).

RISK LITERACY

Basic research should be complemented by applied research, whenever possible. Research on risk literacy is a case in point. Heuristic decisions under uncertainty—when risks are unknown—is a second topic I am engaged with (Gigerenzer, Hertwig, & Pachur, 2011). Applying this research to a variety of areas in the real world, from medicine to romance, has been fun and extended my horizon considerably. And it has helped to see the political consequences of what we are doing. For too long a time, some psychologists have argued that cognitive illusions are basically wired into our minds, meaning that there is little promise in trying to educate people (Kahneman, 2011). This stance has recently fueled a new wave of paternalism, according to which we need to be ‘nudged’ into behaving sensibly by the few sane experts on earth (Thaler & Sunstein, 2008). Some kind of nudging can be useful, but not as a general philosophy of the 21st century. The debate between the nudgers and me is covered in a *Nature* article entitled ‘Risk School’ (Bond, 2009).

What my research shows, however, is that we can make people risk savvy—with the proper tools. With natural frequencies, even fourth graders can solve Bayesian problems, just as doctors and patients can learn to better understand health statistics (Gigerenzer, 2014). In teaching people these skills, psychologists can make a difference. Risk literacy needs to be taught not only to ongoing doctors and judges, but beginning at elementary school. Risk-savvy citizens are indispensable pillars of a modern democracy.

REFERENCES

- Bond, M. (2009). Risk school. *Nature*, 461(29), 1189–1192.
- Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: problems and opportunities. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases* (pp. 249–267). Cambridge, UK: Cambridge University Press.
- Gigerenzer, G. (2002). *Calculated Risks: How to Know When Numbers Deceive you*. New York: Simon & Schuster. (UK version: *Reckoning with Risk: Learning to Live With Uncertainty*, London: Penguin).
- Gigerenzer, G. (2004). Mindless statistics. *Journal of Socio-Economics*, 33, 587–606. doi: 10.1016/j.socec.2004.09.033
- Gigerenzer, G. (2014). *Risk Savvy: How to Make Good Decisions*. New York: Viking.
- Gigerenzer, G., & Edwards, A. G. K. (2003). Simple tools for understanding risks: from innumeracy to insight. *British Medical Journal*, 327, 741–744. doi: 10.1136/bmj.327.7417.741
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review*, 102, 684–704. doi: 10.1037/0033-295X.102.4.684
- Gigerenzer, G., & Hoffrage, U. (1999). Overcoming difficulties in Bayesian reasoning: a reply to Lewis & Keren and Mellers & McGraw. *Psychological Review*, 106, 425–430. doi: 10.1037/0033-295X.106.2.425
- Gigerenzer, G., & Muir Gray, J. A. (Eds.). (2011). *Better Doctors, Better Patients, Better Decisions: Envisioning Health Care 2020*. Cambridge, MA: MIT Press.
- Gigerenzer, G., & Murray, D. J. (1987). *Cognition as Intuitive Statistics*. Hillsdale, NJ: Erlbaum.
- Gigerenzer, G., Fiedler, K., & Olsson, H. (2012). Rethinking cognitive biases as environmental consequences. In P. M. Todd, G. Gigerenzer, & the ABC Research Group. *Ecological Rationality: Intelligence in the World* (pp. 80–110). New York: Oxford University Press.
- Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients to make sense of health statistics. *Psychological Science in the Public Interest*, 8, 53–96.
- Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). *Heuristics: The Foundations of Adaptive Behavior*. New York: Oxford University Press.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: a Brunswikian theory of confidence. *Psychological Review*, 98, 506–528. doi: 10.1037/0033-295X.98.4.506
- Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Krüger, L. (1989). *The Empire of Chance: How Probability Changed Science and Everyday Life*. Cambridge, UK: Cambridge University Press.
- Hoffrage, U., & Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Academic Medicine*, 73, 538–540. doi: 10.1097/00001888-199805000-00024
- Hoffrage, U., Kurzenhäuser, S., & Gigerenzer, G. (2000). Wie kann man die Bedeutung medizinischer Testbefunde besser verstehen und kommunizieren? *Zeitschrift für ärztliche Fortbildung und Qualitätssicherung*, 94, 713–719.
- Kahneman, D. (2011). *Thinking Fast and Slow*. London: Allen Lane.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, UK: Cambridge University Press.
- Kleiter, G. D. (1994). Natural sampling: rationality without base rates. In G. H. Fischer & D. Laming (Eds.), *Contributions to Mathematical Psychology, Psychometrics, and Methodology* (pp. 375–388). New York: Springer.
- Koehler, J. J. (1997). One in millions, billions, and trillions: lessons from People vs. Collins (1968) for People vs. Simpson (1995). *Journal of Legal Education*, 47, 214–223.
- Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: the state of the art to 1980. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases* (pp. 306–334). Cambridge, UK: Cambridge University Press.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.
- Wegwarth, O., Gaissmaier, W., & Gigerenzer, G. (2011). Deceiving numbers: survival rates and their impact on doctors’ risk communication. *Medical Decision Making*, 31(3), 386–394. doi: 10.1177/0272989X10391469