

The Bias Bias in Behavioral Economics

Gerd Gigerenzer*

Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany; gigerenzer@mpib-berlin.mpg.de

ABSTRACT

Behavioral economics began with the intention of eliminating the psychological blind spot in rational choice theory and ended up portraying psychology as the study of irrationality. In its portrayal, people have systematic cognitive biases that are not only as persistent as visual illusions but also costly in real life—meaning that governmental paternalism is called upon to steer people with the help of “nudges.” These biases have since attained the status of truisms. In contrast, I show that such a view of human nature is tainted by a “bias bias,” the tendency to spot biases even when there are none. This may occur by failing to notice when small sample statistics differ from large sample statistics, mistaking people’s random error for systematic error, or confusing intelligent inferences with logical errors. Unknown to most economists, much of psychological research reveals a different portrayal, where people appear to have largely fine-tuned intuitions about chance, frequency, and framing. A systematic review of the literature shows little evidence that the alleged biases are potentially costly in terms of less health, wealth, or happiness. Getting rid of the bias bias is a precondition for psychology to play a positive role in economics.

Keywords: Behavioral economics, Biases, Bounded Rationality, Imperfect information

Behavioral economics began with the intention of eliminating the psychological blind spot in rational choice theory and ended up portraying psychology as the source of irrationality. Inspired by the heuristics-and-biases program of the 1970s (Tversky and Kahneman, 1974), a variety of deviations from rational choice theory were interpreted as systematic flaws in the human mind rather

*I would like to thank Florian Artinger, Wolfgang Gaissmaier, Andrew Gelman, Perke Jacobs, Joshua B. Miller, Hans-Jörg Neth, Andreas Ortmann, and Peter Todd for helpful comments.

than in the theory. Subsequently, an attempt was made to explain the deviations by adding free parameters to expected utility theories (Camerer *et al.*, 2003). In this vision, Homo economicus remained the unquestioned ideal, while Homer Simpson was teasingly proposed as a more apt description of Homo sapiens, given his bouts of ignorance, laziness, and brilliant incompetence (Thaler and Sunstein, 2008).

The field would have progressed in an entirely different direction had it followed Herbert Simon's original vision of behavioral economics, in which psychology comprised more than mental quirks and theory more than expected utility maximization. Simon (1979, 1989) urged economists to move away from as-if expected utility models and study how people actually make decisions in realistic situations of uncertainty as opposed to under risk. I use the term *uncertainty* for situations where the exhaustive and mutually exclusive set of future states and their consequences is not known or knowable. In these frequent real-world situations where the assumptions of rational choice theory do not hold, individuals and institutions can nevertheless make smart decisions using psychological tools such as fast-and-frugal heuristics (Gigerenzer and Selten, 2001). Frank Knight, John Maynard Keynes, and Vernon Smith similarly distinguished situations where rational choice theory applies and where its assumptions are violated, along with the need for a complementary framework for decision making under uncertainty.

Whereas Simon aimed at developing an alternative to neoclassical economics, the line of behavioral economics shaped by the heuristics-and-biases program largely limited itself to studying deviations from the neoclassical paradigm, or what it took this paradigm to be. Experimenters aimed at demonstrating “anomalies” and “biases” in human behavior. These biases have attained the status of truisms and provided justification for new paternalistic policies, popularly known as *nudging*, adopted by governments in the UK, US, and elsewhere. The argument leading from cognitive biases to governmental paternalism—in short, the *irrationality argument*—consists of three assumptions and one conclusion:

1. *Lack of rationality.* Experiments have shown that people's intuitions are systematically biased.
2. *Stubbornness.* Like visual illusions, biases are persistent and hardly corrigible by education.
3. *Substantial costs.* Biases may incur substantial welfare-relevant costs such as lower wealth, health, or happiness.
4. *Biases justify governmental paternalism.* To protect people from their biases, governments should “nudge” the public toward better behavior.

The three assumptions—lack of rationality, stubbornness, and costs—imply that there is slim chance that people can ever learn or be educated out of their biases; instead governments need to step in with a policy called *libertarian paternalism* (Thaler and Sunstein, 2003). In this variant of soft paternalism, governments use nudges that exploit people’s cognitive deficiencies in order to steer them toward proper behavior (Rebonato, 2012). This new paternalism is not aimed at protecting people from imperfections of the markets or from criminals, but from the enemy within, their own irrationality. The irrationality argument exists in many versions (e.g. Conley, 2013; Kahneman, 2011). Not only has it come to define behavioral economics but it also has defined how most economists view psychology: Psychology is about biases, and psychology has nothing to say about reasonable behavior.

Few economists appear to be aware that the bias message is not representative of psychology or cognitive science in general. For instance, loss aversion is often presented as a truism; in contrast, a review of the literature concluded that the “evidence does not support that losses, on balance, tend to be any more impactful than gains” (Gal and Rucker, 2018). Research outside the heuristics-and-biases program that does not confirm this message—including most of the psychological research described in this article—is rarely cited in the behavioral economics literature (Gigerenzer, 2015). To correct this picture, I discuss the three assumptions of the irrationality argument in the first part of this article. In the second part, I address the first assumption (lack of rationality) in more detail and show that behavioral economics is tainted by a “bias bias,” the tendency to spot biases even if there are none (Brighton and Gigerenzer, 2015). I will not deal in any depth with the positive alternative to current behavioral economics, a topic beyond the scope of this article; other authors and I have elsewhere written about how to rethink behavioral economics based on Simon’s perspective (e.g., Gigerenzer, 2016a,b; Gigerenzer *et al.*, 2011; Gigerenzer and Selten, 2001).

1 Part I: The Irrationality Argument

1.1 *The Bias Bias*

Economists, including experimental economists, have traditionally trusted in the rationality of people, at least in the aggregate and when people are given the time and opportunity to learn from experience (Binmore, 2009; Smith, 2007). The claim that people systematically violate the laws of probability (Tversky and Kahneman, 1974) was new and surprising given that throughout the 1960s, experimental psychologists arrived at the conclusion that people are fairly good intuitive statisticians (e.g., Edwards, 1968; Peterson and Beach, 1967; Piaget and Inhelder, 1975). It also conflicts

with contemporary cognitive scientists' view of people's conscious *and* unconscious cognitive processes as largely consistent with Bayesian statistics (e.g., Anderson, 1990; Chater *et al.*, 2006) and with neuroscience research that speaks of the Bayesian brain (e.g., Friston, 2010). Nevertheless, the bias message single-handedly shaped the emerging field of behavioral economics. According to Conlisk (1996), people

display intransitivity; misunderstand statistical independence; mistake random data for patterned data and vice versa; fail to appreciate law of large number effects; fail to recognize statistical dominance; make errors in updating probabilities on the basis of new information; understate the significance of given sample sizes; fail to understand covariation for even the simplest 2X2 contingency tables; make false inferences about causality; ignore relevant information; use irrelevant information (as in sunk cost fallacies); exaggerate the importance of vivid over pallid evidence; exaggerate the importance of fallible predictors; exaggerate the ex ante probability of a random event which has already occurred; display overconfidence in judgment relative to evidence; exaggerate confirming over disconfirming evidence relative to initial beliefs; give answers that are highly sensitive to logically irrelevant changes in questions; do redundant and ambiguous tests to confirm an hypothesis at the expense of decisive tests to disconfirm; make frequent errors in deductive reasoning tasks such as syllogisms; place higher value on an opportunity if an experimenter rigs it to be the "status quo" opportunity; fail to discount the future consistently; fail to adjust repeated choices to accommodate intertemporal connections; and more (p. 670).

The argument is not that people sometimes make mistakes, engage in wishful thinking, or lack education in statistical thinking; that would be nothing new. Rather, these biases are implied to be engrained in our brains like hard-wired visual illusions so that virtually everyone errs in the same systematic way. According to Thaler (1991), "mental illusions should be considered the rule rather than the exception" (p. 4).

A bias is a systematic discrepancy between the (average) judgment of a person or a group and a true value or norm. Most of the biases in this list are defined as deviations from statistical principles. At issue is not that people make errors—we all do—but whether these statistical or logical principles are actually sensible norms of behavior, which would qualify deviations as mental illusions. As every practicing statistician knows, a statistical principle does not automatically qualify as a general norm for all situations; rather, a careful analysis of the problem at hand and the assumptions made is necessary. Herein lies the problem called the "bias bias" (Gigerenzer, 2016a).

Bias bias: The tendency to see systematic biases in behavior even when there is only unsystematic error or no verifiable error at all.

But why would researchers exhibit a bias bias? One motive is an academic agenda to question the reasonableness of people. Unlike in Simon's vision, behavioral economics à la Kahneman and Thaler defined its agenda as uncovering deviations from rational choice theory; without these deviations it would lose its *raison d'être*. A second possible motive is a commercial agenda to discredit the judgment of jurors, as the Exxon Valdez oil spill in Alaska illustrates. In 1994, an Alaskan federal jury had awarded \$5.3 billion to fishermen and others whose livelihoods had been devastated by the spill. Exxon quietly funded a new line of research using studies with mock juries that questioned juror's cognitive abilities. Based on the results, Exxon argued in an appeal "that jurors are generally incapable of performing the tasks the law assigns to them in punitive damage cases" (Zarembko, 2003). This argument served Exxon well in court. A third possible motive—or more likely, unintended consequence—is to promote trust in abstract logic, algorithms, and predictive analytics and distrust in human intuition and expertise. For instance, the COMPAS algorithm has been used in U.S. courts to predict the probability of recidivism for over one million defendants, influencing pretrial, parole, and sentencing decisions. Apparently, the algorithm remained unquestioned for almost two decades until a study disclosed that COMPAS predicted no better than ordinary people without any experience with recidivism and had a racial bias to boot (Dressel and Farid, 2018). In the book jacket copy of his biography on Kahneman and Tversky, Lewis (2017) states that they "are more responsible than anybody for the powerful trend to mistrust human intuition and defer to algorithms."

1.2 *Stubbornness and Visual Illusions*

In contrast to experimental economists' reports that people learn to make better decisions over time (Binmore, 2009; Smith, 2007), many behavioral economists are more pessimistic. Biases are said to exhibit "stubborn persistence" (Tversky and Kahneman, 1983, p. 300), which is attributed to a biologically old "System 1" that is "not readily educable" (Kahneman, 2011, p. 417), so that "our ability to de-bias people is quite limited" (Thaler, cited in Bond, 2009, p. 1191). The rhetoric is that "we are not only irrational, but *predictably* irrational" (Ariely, 2008, p. xviii).

Again, at issue is not that people make errors, such as when native English speakers make grammatical mistakes in French, or students untaught in calculus try to solve calculus problems. That would have hardly attracted attention and served as the foundation of behavioral economics. The claim is that biases are firmly imbedded in our brains.

Visual Illusions

By way of suggestion, articles and books introduce biases together with images of visual illusions, implying that biases (often called “cognitive illusions”) are equally stable and inevitable. If our cognitive system makes such big blunders like our visual system, what can you expect from everyday and business decisions? Yet this analogy is misleading, and in two respects.

First, visual illusions are not a sign of irrationality, but a byproduct of an intelligent brain that makes “unconscious inferences”—a term coined by Hermann von Helmholtz—from two-dimensional retinal images to a three-dimensional world. In Kandel’s (2012) words, “In fact, if the brain relied solely on the information it received from the eyes, vision would be impossible” (p. 203). Consider Roger Shepard’s fascinating two-tables illusion, which Thaler and Sunstein (2008) present as “the key insight that behavioral economists have borrowed from psychologists,” namely that our judgment is “biased, and predictably so” (p. 20). Yet that is not the insight drawn by Shepard and other psychologists who study visual illusions; they consider our visual system a marvel of unconscious intelligence that manages to infer three-dimensional objects (here, tables) from a two-dimensional retinal picture. Because any two-dimensional projection is consistent with infinitely different three-dimensional realities, it is remarkable how seldom we err. In Shepard’s (1990) words, “to fool a visual system that has a full binocular and freely mobile view of a well-illuminated scene is next to impossible” (p. 122). Thus, in psychology, the visual system is seen more as a genius than a fool in making intelligent inferences, and inferences, after all, are necessary for making sense of the images on the retina.

Second, the analogy with visual illusions suggests that people cannot learn, specifically that education in statistical reasoning is of little efficacy (Bond, 2009). This is incorrect, as can be shown with one of the most important items on Conlisk’s list: making errors in updating probabilities.

Bayesian Reasoning Can Be Taught in 2 Hours

Consider the claim that people fail to make forecasts that are consistent with Bayes’ rule: the human mind “is not Bayesian at all” (Kahneman and Tversky, 1972, p. 450; see also Kahneman, 2011, pp. 146–155). This claim is correct if people have no statistical education *and* are given the relevant information in conditional probabilities. But another picture emerges if people receive a short training and learn to use proper representations. To illustrate, please attempt to solve this problem without the help of a calculator.

Conditional probabilities:

A disease has a base rate of .1, and a test is performed that has a hit rate of .9 (the conditional probability of a positive test given

disease) and a false positive rate of .1 (the conditional probability of a positive test given no disease). What is the probability that a random person with a positive test result actually has the disease?

There is reliable experimental evidence that most people cannot find the Bayesian posterior probability, which would be $(.1 \times .9) / [(.1 \times .9) + (.9 \times .1)] = .5$. However, when people can learn from experience instead, research reports that people are approximate Bayesians (e.g., Edwards, 1968). On the basis of this result, one can present the information as the tally of experienced frequencies, known as *natural frequencies* (because experienced frequencies are not normalized like conditional probabilities or relative frequencies):

Natural frequencies:

Among every 100 people, 10 are expected to have a disease. Among those 10, nine are expected to correctly test positive. Among the 90 people without the disease, nine are expected to falsely test positive. What proportion of those who test positive actually have the disease?

Now it is easier to see what the answer is: $9 / (9 + 9) = 1$ of 2. The theoretical reason is that natural frequencies facilitate Bayesian computations. Since the first demonstration of the de-biasing effect of natural frequencies (Gigerenzer and Hoffrage, 1995, 1999; Hoffrage *et al.*, 2000), many studies have documented this effect among both laypeople and experts. For instance, only 21% of 160 gynecologists could correctly infer the posterior probability that a woman has breast cancer after a positive screening mammogram when the information was given in conditional probabilities. This number increased to 87% after the gynecologists were trained how to translate conditional probabilities into natural frequencies (Gigerenzer *et al.*, 2007). A similar effect of teaching, where the percentage of those making correct inferences increased from 23% to 88%, was observed among medical students in their final year (Jenny *et al.*, 2018). When college students were given a 2-hour course in natural frequencies, the number of correct Bayesian inferences increased from 10% to 90%; most important, this 90% rate was maintained 3 months after training (Sedlmeier and Gigerenzer, 2001). Meta-analyses have also documented the “de-biasing” effect, and natural frequencies are now a technical term in evidence-based medicine (Aki *et al.*, 2011; McDowell and Jacobs, 2017). These results are consistent with a long literature on techniques for successfully teaching statistical reasoning (e.g., Fong *et al.*, 1986). In sum, humans can learn Bayesian inference quickly if the information is presented in natural frequencies.

1.3 Lack of Evidence of Costs

The argument that biases are costly is essential for justifying governmental paternalism. The irrationality argument provides a convenient rhetoric to attribute problems caused by flawed incentives and system failures to flaws inside people's minds, detracting from political and industrial causes (e.g., Conley, 2013; Thaler, 2015; Thaler and Sunstein, 2008). Nicotine addiction and obesity have been attributed to people's myopia and probability-blindness, not to the actions of the food and tobacco industry. Similarly, an article by the *Deutsche Bank Research* "Homo economicus – or more like Homer Simpson?" attributed the financial crisis to a list of 17 cognitive biases rather than the reckless practices and excessive fragility of banks and the financial system (Schneider, 2010).

The claim that biases incur substantial costs is often based on anecdotal evidence. Surely there must be hard evidence for this important hypothesis? Arkes *et al.* (2016) searched for such evidence through articles that demonstrated biases. In over 100 studies on violations of transitivity, the search found not a single one demonstrating that a person could become a money pump, that is, be continually exploited due to intransitive choices. In more than 1,000 studies that identified preference reversals, arbitrage or financial feedback made preference reversals and their costs largely disappear. Similarly, in hundreds of studies on the Asian Disease Problem and other framing effects, little evidence was found that "irrational" attention to framing would be costly. All in all, little to no evidence was found that violations of these and other logical rules are either associated or causally connected with less income, poorer health, lower happiness, inaccurate beliefs, shorter lives, or any other measurable costs. Lack of evidence for costs does not mean evidence for lack of costs; however, this striking absence suggests that the ever-growing list of apparent fallacies includes numerous "logical bogeymen," as psychologist Lola Lopes once put it.

There are two ways to understand this negative result. One reason is the bias bias, that is, many of the alleged biases are not biases in the first place. The second reason was mentioned above: The biases are typically defined as deviations of people's judgments from a logical or statistical rule in some laboratory setting, and therefore may have little to do with actual health, wealth, or happiness.

1.4 Less Paternalism and More Risk Savvy Citizens

The previous analysis showed that the three assumptions in the argument leading from biases to governmental paternalism are based on questionable evidence. Moreover, systematic departures from rational choice may either increase or decrease policy makers' rationale for paternalism, depending on the specific assumptions made; specifically, the aggregate consequences of

individual departures may be Pareto-inferior or superior (Berg and Gigerenzer, 2007).

The alternative to paternalism is to invest in citizens so that they can reach their own goals rather than be herded like sheep. Immanuel Kant (1784) called upon people to use their own minds without external steering or coercion, which is equally important today, where the combination of nudging with big data technology provides opportunities for “Big Nudging,” that is, the massive manipulation of people’s goals and preferences using large amounts of data (Helbing *et al.*, 2017). To that end, measures need to be taken to hone the skills of the general public in dealing with risks and making decisions, for instance by teaching techniques such as natural frequencies (see above) and effective decision-making heuristics to laypeople and experts in matters such as health, finance, and digital media (Gigerenzer, 2014). The behavioral science unit of the German chancellery has implemented this approach successfully since 2015 (Bundesregierung [Federal Government of Germany], 2018). It aims at “boosting” rather than “nudging” (Hertwig and Grüne-Yanoff, 2017).

In the second part of this article, I provide evidence for the bias bias. I use several biases that have attained the status of truisms to illustrate the general point. Once again, the message is not that people do not make errors; it is that researchers need to be more careful when attributing systematic biases to people.

2 Part II: Case Studies in the Bias Bias

There is no room here to cover all items on Conlisk’s list, not to speak of the 175 or so biases listed on Wikipedia. Instead I will focus on intuitions about randomness, frequency, and framing, which play a role in several alleged biases. Specifically, I will deal with three common sources of the bias bias:

1. *Small sample statistics are mistaken for people’s biases.* The statistics of small samples can systematically differ from the population parameters. If researchers ignore this difference, they will mistake correct judgments about small samples for systematic biases.
2. *Unsystematic errors are mistaken for systematic biases.* A bias is a systematic error, not an unsystematic one.
3. *Logically equivalent frames are mistaken as informationally equivalent.* To pay attention to how a message is framed is not a logical error but an essential feature of human intelligence.

For instance, Point 1 is relevant for critically evaluating the claim in Conlisk’s list that people “misunderstand statistical independence” and “mistake

random data for patterned data and vice versa,” Point 2 for the claim that people “display overconfidence,” and Point 3 for the claim that people “use irrelevant information” and “give answers that are highly sensitive to logically irrelevant changes in questions.” In each case, the problem lies in the norm, against which a judgment is evaluated as a bias. I hope that these analyses encourage readers to look more critically at other biases.

2.1 *People’s Intuitions About Randomness*

A large body of psychological research has concluded that people have good intuitions about chance. For instance, children are reported to have a fairly solid understanding of random processes, including the law of large numbers (Piaget and Inhelder, 1975) and the production of random sequences (Kareev, 1992), and adults are reported to be good intuitive statisticians (Farmer *et al.*, 2017; Peterson and Beach, 1967).

In contrast to these findings, one of the most fundamental cognitive biases reported in the heuristics-and-biases program is that intuitions about chance depart systematically from the laws of chance. Consider experiments on the perception of randomness in which a fair coin is thrown many times, and the outcomes of heads (H) and tails (T) are recorded. Two key experimental findings are:

1. *Law of small numbers:* People think a string is more likely the closer the number of heads and tails corresponds to the underlying equal probabilities. For instance, the string HHHHHT is deemed more likely than HHHHHH.
2. *Irregularity:* If the number of heads and tails is the same in two strings, people think that the one with a more irregular pattern is more likely. For instance, the string HHTHHT is deemed more likely than HTHTHT.

Both phenomena have been called systematic errors and are often confused with the gambler’s fallacy. But that is not so.

Law of Small Numbers

To simplify the presentation, I will use short strings of length $k = 3$. If a fair coin is thrown n times ($n \geq k$), which of the two strings is more likely to be encountered?

HHH
HHT

Most people think that HHT is more likely, as documented in both recognition tasks (as above) and production tasks (in which people are asked to produce random strings; Hahn, 2014). This intuition has been deemed a fallacy

H	H	H	H	H	H	H	H	T	T	T	T	T	T	T	T
H	H	H	H	T	T	T	T	H	H	H	H	T	T	T	T
H	H	T	T	H	H	T	T	H	H	T	T	H	H	T	T
H	T	H	T	H	T	H	T	H	T	H	T	H	T	H	T
✓	✓+	+	+	-	-	-	-	✓	+	-	-	-	-	-	-

Figure 1: If the length of the string k is smaller than that of the observed sequence n , then a pure string of heads (H) is less likely to be encountered than one with an alternation. This is shown here for the case of $k = 3$ and $n = 4$. There are 16 possible sequences of four tosses of a fair coin, each equally likely. In three of these, there is at least one string HHH (“check mark”), while HHT occurs in four of these (“cross”). In this situation, people’s intuitions are correct (Hahn and Warren, 2009).

because each of the two strings is said to have the same probability of occurring. The alleged bias was explained by people’s belief in the “law of small numbers,” which means that people expect the equal probability of H and T to hold in small samples as well: HHT is more “representative” than HHH (Kahneman and Tversky, 1972). The belief in the law of small numbers has been seen as a key example of how people depart from perfect rationality, posing a radical challenge to neoclassical economic theory (Rabin, 1998), and was modeled as mistaking an urn with replacement for one with limited replacement (Rabin, 2002). Rabin and Vayanos (2010) concluded that people “view randomness in coin flipping as corresponding to a switching rate of 60% instead of 50%” (p. 735), which corresponds to misunderstanding statistical independence in Conlisk’s list. Whatever the explanation, the alleged bias shows “stubbornness,” namely that people show little insight and fail to overcome it: “For anyone who would wish to view man as a reasonable intuitive statistician, such results are discouraging” (Kahneman and Tversky, 1972, p. 445).

Let us have a closer look. Consider a simple experiment:

John takes a coin and flips it *four* times in a row. Is it more likely that John will encounter a string of HHH or of HHT?

Both appear to be equally likely. Surprisingly, however, the answer is that HHT is more likely. More specifically, the relative frequency of encountering at least one HHT is $4/16 = .25$ and that of HHH is $3/16 = .19$. The 16 possible sequences are shown in Figure 1; they are all equally probable. In four of 16 possible sequences, John will encounter at least one HHT, and in three sequences, John will encounter at least one HHH. The reason is that the empirical probability in a small sample is not the same as the true probability. Another way to understand the result is that two strings of HHH overlap in the first column of Figure 1, whereas no such overlaps occur (or are even possible) for HHT. Thus, people’s intuition that an HHT is more likely to be encountered than an HHH in a sequence of four flips is correct.

Consider now the general case. Let k be the length of the string of heads and tails judged (which is three in the example above), and n be the total sequence (number of tosses) observed ($k \leq n$). If $k = n = 3$, there are eight possible outcomes, all equally likely, and one of these contains HHH and one HHT; thus, both are equally likely to be encountered. In this situation, the intuition that HHT is more likely would be a fallacy. If $k < n$, however, that same intuition is correct. Similarly, it can be shown that HHT is likely to be encountered *earlier* than HHH: The expected waiting time for HHT is eight tosses of a coin, compared with 14 tosses for HHH (Hahn and Warren, 2009).

Now we can specify the general principle under which people's intuition is *ecologically rational*:

If $k < n < \infty$, a string of H's with a single alternation such as HHT is more likely to be encountered than a pure string such as HHH.

The qualifier *ecologically* means that statements about the rationality of judgments need to be qualified with respect to ecological conditions: here, the size k of the string and the size n of the sequence. In sum, the intuition that HHT is more likely to be encountered than HHH is not generally a logical error, but *ecologically* rational if $k < n < \infty$. The condition $k < n$ holds when a person has seen more trials than the short string under question, and the condition $n < \infty$ is always true for mortals. In other words, the ecological principle states only two exceptions where the "belief in the law of small numbers" can be called a fallacy: if $k = n$ and if the window of observations is infinite (which is impossible for a mortal being).

If you are still not convinced, try this bet (Hahn and Warren, 2010), which I will call the *law-of-small-numbers bet*:

You flip a fair coin 20 times. If this sequence contains at least one HHHH, I pay you \$100. If it contains at least one HHHT, you pay me \$100. If it contains neither, nobody wins.

If HHHH and HHHT were equally likely to be encountered, then the two players should break even. But in fact, the person who accepts the bet can expect to lose in the long run. If you watch 20 flips, the probability that you will see at least one HHHH is about 50%, but the chance of an HHHT is around 75% (Hahn and Warren, 2009). For the same reason, a gambler who watches the outcomes of the roulette wheel in a casino for half an hour can more likely expect to see a string of three reds followed by a black than a string of four reds. In this situation, believing in the law of small numbers pays.

The phenomenon that people expect more alternations than probability theory predicts has been sometimes linked to the gambler's fallacy. Yet that assumption is mistaken, as can be deduced from the ecological rationality

condition. The gambler’s fallacy refers to the intuition that after witnessing a string of, say, three heads, one expects that the next outcome will be more likely tail than head. This would be a true fallacy because it corresponds to the condition $k = n$. In other words, a total of four throws is considered, either HHHH or HHHT, and there is no sample k with the property $k < n$.

Irregularity

The second alleged misconception about chance is that people believe that irregular sequences are more likely. Consider the following two sequences:

HTHTHT

HTTHTH

Which sequence is more likely to be encountered? The number of heads and tails are now identical, but alternations are regular in the first string and irregular in the second. Psychological research documented that most people judge the more irregular string as more likely. In Kahneman and Tversky’s (1972) account, among all possible strings of length six, “we venture that only HTTHTH appears really random” (p. 436).

Once again, this intuition was generally declared a fallacy, for the same reason as before: All sequences are assumed to have equal probability. Yet the same ecological analysis shows that if $k < n < \infty$, then the sequence HTTHTH is actually more likely than the regular one. For instance, if John throws a fair coin, the expected waiting time to get a HTTHTH is 64 flips, whereas it is 84 for HTHTHT (and 126 for HHHHHH; see Hahn and Warren, 2009). This can be verified in the same way as with the table above.

In sum, under the ecological condition, people’s belief that irregular alternations are more likely or more quickly to be encountered reflects an astonishingly fine-tuned sensitivity to the statistics of finite samples. The belief is erroneous only in cases where the condition does not hold.

Fine-tuned Intuition Is Mistaken for Bias

In summary, people’s intuitions about chance have been interpreted as systematically flawed because they (i) fall prey to the law of small numbers and (ii) confuse irregularity with chance, both in the recognition and production of randomness. In contrast, an ecological analysis shows that if $k < n < \infty$, which is the typical window of experience, people’s intuitions are in fact correct. The general insight is that the properties of the underlying chance process (such as equal probability) do not match the properties of small samples, a phenomenon known as biased sample estimators in statistics. In other words, people’s intuitions were evaluated against the wrong normative standard, one that does not capture the properties of samples.

In the various justifications for why people commit the gambler's fallacy, the condition $k < n < \infty$ appears to have been overlooked in the bias literature. In the next section, we will see that the same oversight applies to the hot hand fallacy, the mirror image of the gambler's fallacy.

2.2 Professional Intuitions About Randomness: The Hot Hand Fallacy

Most basketball fans can recall magical moments where a player is “on fire,” “in the zone,” “in rhythm,” or “unconscious.” This temporarily elevated performance is known as the “hot hand.” For players and coaches, the hot hand is a common experience. Gilovich *et al.* (1985), however, considered this intuition a cognitive illusion and named it the *hot hand fallacy*. Even the website of the National Collegiate Athletic Association (NCAA) once warned of believing in magic: “Streaks and ‘hot hands’ are simply illusions in sports. And, it is better to be a scientist than be governed by irrational superstition.”

Coaches and players reacted with stubborn disbelief, which some behavioral economists saw as an indication that the hot hand fallacy resembles a robust visual illusion. Thaler and Sunstein (2008) asserted “that the cognitive illusion is so powerful that most people (influenced by their Automatic System) are unwilling even to consider the possibility that their strongly held beliefs might be wrong” (p. 31). Like the gambler's fallacy, the hot hand fallacy was attributed to “representativeness” and formalized by Rabin and Vayanos (2010). It has served as an explanation for various vices and behaviors in financial markets, sports betting, and casino gambling.

Professional coaches and players presumably have no incentives to be wrong, only to be right. Why then would this erroneous belief exist and persist? The answer is the same as with intuitions about randomness.

Gilovich *et al.* (1985) took care to collect multiple sources of data and made their strongest case against the hot hand by looking at free shots, where the other team cannot strategically reduce a “hot” player's number of hits through physical intervention (Raab *et al.*, 2012). Twenty-six shooters from Cornell each did 100 free shots from a fixed distance with varying locations. The authors argued that if there is no hot hand (the null hypothesis), then the frequency of a hit after three hits should be the same as that of a miss after three hits. And that is exactly what they reported for all but one player—this exception could be expected to happen by chance. Thus, their conclusion was that the hot hand does not exist.

Now let us take a closer look at the null hypothesis of chance. To keep it short, I use Hahn and Warren's (2009) analysis of the gambler's fallacy as an analogy, whereas the original analysis of the hot hand by Miller and Sanjuro (forthcoming) and Miller and Sanjuro (2016) was in terms of selection bias. If there is no hot hand, the frequency of a hit after three hits should be *not* equal but actually smaller, and for the reason just discussed. For again we

are dealing with $k = 4$ (the length of the string) and $n = 100$; that is, the ecological condition $k < n < \infty$ is in place. Figure 1 can be used to illustrate the point by replacing the fair coin with a player who shoots from a distance where there is a probability of .50 to score a hit and by replacing the streak of four in the free-shooter experiment with a streak of three. In this situation, is a hit or a miss more likely after two hits?

If a player makes $n = 4$ free shots, there are 16 sequences with equal probabilities. In the absence of a hot hand, we observe a miss (T) after two hits (H) in four of these sequences and a hit after two hits in only three sequences. Thus, Gilovich *et al.*'s null hypothesis that the expected relative frequency of hits and misses after a streak of hits should be the same is not correct in a small sample. The correct null hypothesis is that after a streak of HH, H has an expected proportion of about .42, not .50 (Miller and Sanjuro, forthcoming). Because HHH should be less often observed, finding a relative frequency of .50 instead of .42 actually indicates a hot hand. In a reanalysis of the original data, a substantial number of the shooters showed a pattern consistent with the hot hand (Miller and Sanjuro, forthcoming). Across players, the hot hand boosted performance by 11 percentage points, which is substantial and roughly equal to the difference in field goal percentage between the average and the very best three-point shooter in the NBA.

Many have been puzzled about the relationship between the gambler's fallacy and the hot hand fallacy, given that they refer to contrary phenomena, a negative recency bias and a positive recency bias (Rabin and Vayanos, 2010). Both have been attributed to the usual suspects, the representativeness heuristic or the associative System 1. The present analysis shows a quite different link: Both are a consequence of the same bias bias. Coaches and players have good reason to maintain their belief in the hot hand.

2.3 Frequencies: Overestimation of Low Risks and Underestimation of Large Risks?

Frequency is one of the few attributes of objects that seem to be encoded automatically (others being spatial location, temporal information, and word meaning; see Hasher and Zacks, 1984; Barsalou, 1992). For instance, 8-month-old infants learn language from the statistical relationships between neighboring speech sounds (Saffran *et al.*, 1996). Frequency encoding appears to be a fundamental process in learning and encoding information; a meta-analysis of 26 studies with 4,570 participants reported a mean Pearson correlation of .73 between frequency estimates and actual frequencies across many variables, including words, pictures, events, and letters (Kane and Woehr, 2006). As Jonides and Jones (1992) noted in their summary of the psychological literature, "Ask about the relative numbers of many kinds of events, and you are likely to

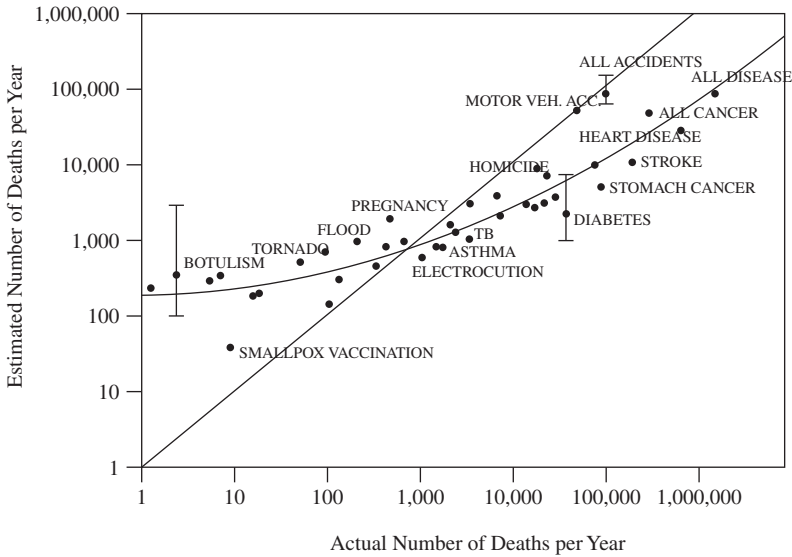


Figure 2: Relationship between estimated and actual number of deaths per year for 41 causes of death in the United States (Lichtenstein *et al.*, 1978). Each point is the mean estimate (geometric mean) of 39 students; vertical bars show the variability (25th and 75th percentile) around the mean estimates for botulism, diabetes, and all accidents. The curved line is the best-fitting quadratic regression line. This pattern has been (mis)interpreted as evidence of overestimation of low risks and underestimation of high risks. (Adapted from Slovic *et al.*, 1982).

get answers that reflect the actual relative frequencies of the events with great fidelity” (p. 368).

Against this evidence, the claim that people show systematic biases in evaluating the frequency of events is surprising. In a classic study, 39 college students were asked to estimate the frequency of 41 causes of death in the United States, such as botulism, tornado, and stroke (Lichtenstein *et al.*, 1978; Slovic *et al.*, 1982). The result (Figure 2) was interpreted as showing two systematic biases in people’s minds: “In general, rare causes of death were overestimated and common causes of death were underestimated” (p. 467).

Why would people overestimate low risks and underestimate high risks? Slovic *et al.* (1982) argued that “an important implication of the availability heuristic is that discussion of a low-probability hazard may increase its memorability and imaginability and hence its perceived riskiness” (p. 456). Kahneman (2011, p. 324) referred to the associative System 1. The two biases correspond to the weighting function in prospect theory and became widely cited as the reason why the public lacks competence in political decisions: People exaggerate the dangers of low-risk technology such as nuclear power plants and underestimate the dangers of high-frequency risks.

Let us now have a closer look at how these biases were diagnosed. Slovic *et al.* (1982) explain: “If the frequency judgments were accurate, they would equal the statistical rates, with all data points falling on the identity line.” (p. 466–467). The identity line is taken as the null hypothesis. Yet that would only be correct if there were no noise in the data, which is not the case, as the wide vertical bars in Figure 2 reveal. In the presence of noise, that is, unsystematic error, the curve is not the identity line but a line that is bent more horizontally, a phenomenon known as *regression toward the mean*. It was discovered in the 1880s by the British polymath Sir Francis Galton, who called it “reversion toward the mean.” Galton observed that the sons of small fathers were on average taller than their fathers, and the sons of tall fathers were on average smaller than their fathers. If we replace fathers and sons with actual and perceived risks, respectively, the result resembles the curved line in Figure 2. However, Galton also observed that the fathers of small sons were on average taller than their sons, and the fathers of tall sons were on average smaller than their sons. The first observation would wrongly suggest that the actual variability of the sons is smaller than that of their fathers, the second observation the opposite. What diminishes in the presence of noise is the variability of the *estimates*, both the estimates of the height of the sons based on that of their fathers, and vice versa. Regression toward the mean is a result of unsystematic, not systematic error (Stigler, 1999).

If the pattern in Figure 2 is due to unsystematic rather than systematic error, plotting the data the other way round as Galton did will show this. In a replication of the original study (Hertwig *et al.*, 2005), when the estimated frequencies were predicted from the actual frequencies, as originally done, the same regression line that fitted the data best was replicated (Figure 3). However, when the actual frequencies were predicted from the estimated frequencies, the regression line showed a reverse pattern, suggesting the opposite biases. For instance, consider low estimates between 10 and 100 per year on the Y-axis. The corresponding actual frequency on the second regression line is higher, between about 50 and 500. This discrepancy now makes it look as though people underestimate, not overestimate low risks. Yet neither is the case: The two regression lines are a consequence of unsystematic noise.

Behavioral economists have repeatedly pointed out that people confuse regression toward the mean with real effects. In this case, it has misled researchers themselves. Milton Friedman (1992) suspected that “the regression fallacy is the most common fallacy in the statistical analysis of economic data” (p. 2131).

Thus, the reanalysis is consistent with the general result in psychological research that people are on average fairly accurate in estimating frequencies, whereas unsystematic error can be large.

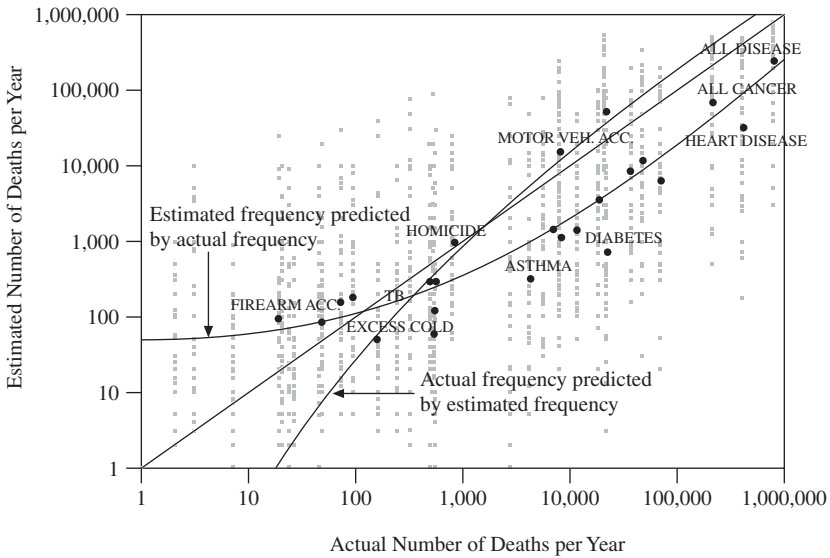


Figure 3: Replication of the causes of death study (Lichtenstein *et al.*, 1978, Figure 2) by Hertwig *et al.* (2005) showing that what appears to be overestimation of low risks and underestimation of high risks is in fact a statistical artifact known as regression toward the mean. The replication used the same 41 causes (7 are not shown because their frequencies were zero in 1996–2000), 45 participants, and no anchor. When the estimated frequencies are predicted from the actual frequencies, the resulting curve (the best-fitting quadratic regression line) is virtually identical with that in the original study (Figure 2). However, when the actual frequencies are predicted from the estimated frequencies, the resulting curve appears to show the opposite bias: Rare risks appear to be underestimated and high risks overestimated. The figure also shows the large unsystematic error in the data (individual data points). (Adapted with permission from Hertwig *et al.*, 2005).

Unsystematic Error Is Mistaken for Systematic Error

The classic study of Lichtenstein *et al.* illustrates the second cause of a bias bias: when unsystematic error is mistaken for systematic error. One might object that systematic biases in frequency estimation have been shown in the widely cited letter-frequency study (Kahneman, 2011; Tversky and Kahneman, 1973). In this study, people were asked whether the letter K (and each of four other consonants) is more likely to appear in the first or the third position of a word. More people picked the first position, which was interpreted as a systematic bias in frequency estimation and attributed post hoc to the *availability heuristic*. After finding no single replication of this study, we repeated it with all consonants (not only the selected set of five, each of which has the atypical property of being more frequent in the third position) and actually measured availability in terms of its two major meanings, number

and speed, that is, by the frequency of words produced within a fixed time and by time to the first word produced (Sedlmeier *et al.*, 1998). None of the two measures of availability was found to *predict* the actual frequency judgments. In contrast, frequency judgments highly correlated with the actual frequencies, only regressed toward the mean. Thus, a reanalysis of the letter-frequency study provides no evidence of the two alleged systematic biases in frequency estimates or of the predictive power of availability.

These studies exemplify two widespread methodological shortcomings in the bias studies. First, the heuristics (availability, representativeness, affect) are never specified; since the 1970s, when they were first proposed, they have remained common-sense labels that lack formal models. Second, the heuristic is used to “explain” a bias after the fact, which is almost always possible given the vagueness of the label. The alternative is to use formal models of heuristics that can be tested; to test these models in prediction, not by data fitting after the fact; and to test multiple models competitively, as opposed to testing a single model against chance (Berg and Gigerenzer, 2010; Gigerenzer *et al.*, 2011).

2.4 Subjective Probability: The Overconfidence Bias

Overconfidence is said to be perhaps the most robust finding in the psychology of judgment and choice (DeBondt and Thaler, 1995). “Perhaps the central finding of behavioral economists is that most of us are overconfident when we make predictions” (Silver, 2012, p. 359). “One of the most significant and irrefutable findings of behavioral psychologists is that people are overconfident in their judgments” (Parikh, 2009, p. 142). Discussing these claims is difficult because the term *overconfidence* refers to at least half a dozen different phenomena without any proven theoretical links or empirical correlations (Mousavi and Gigerenzer, 2011; Olsson, 2014). For that reason, I will distinguish the most prominent of these phenomena by numbering them.

Overconfidence 1: Miscalibration

The classic demonstration of miscalibration has been made with general knowledge questions, such as:

Which city is further south? New York or Rome.

How confident are you in your answer? 50%, 60%, 70%, 80%, 90%, 100%.

In the typical study, participants are asked many such questions, and then the average percentage of correct answers is plotted for each confidence level. A typical result is: When people said they were 100% confident, the average

proportion of correct answers was only 90%; when they said they were 90% confident, the accuracy was 80%; and so on. This “miscalibration” was called *overconfidence* (Lichtenstein *et al.*, 1982). Yet when confidence was low, such as 50%, accuracy was actually *higher* than confidence; this phenomenon was called the *hard–easy effect*: Hard questions result in underconfidence, while easy questions result in overconfidence (Gigerenzer *et al.*, 1991).

The resulting miscalibration curves are similar to those in Figure 2, with confidence on the X-axis (range 50% to 100%) and the average percentage of correct answers on the Y-axis. And they are subject to the same argument as stated before: In the presence of noise (quantitative confidence judgments are noisy), the pattern that has been interpreted as overconfidence bias actually results from unsystematic error because of regression toward the mean. This result can be shown in the same way as in the previous section for overestimation of small risks (similar to Figure 3). If one estimates the confidence judgments from the proportion of correct responses (rather than vice versa, as customary), then one should obtain the mirror result, which appears to be underconfidence. Studies that reanalyzed empirical data sets in this twofold manner consistently reported that regression toward the mean accounted for most or all of the pattern that had been attributed to people’s overconfidence (Dawes and Mulford, 1996; Erev *et al.*, 1994; Pfeifer, 1994). For instance, when looking at all questions that participants answered 100% correctly (instead of looking at all questions where participants were 100% confident), it was found that the average confidence was lower, such as 90 %—which then looked like underconfidence. The regression effect also explains the hard–easy effect, that is, why “overconfidence” flips into “underconfidence” when confidence is low (Juslin *et al.*, 2000).

In sum, studies on “miscalibration” have misinterpreted regression toward the mean as people’s overconfidence, just as studies on overestimation of low risks and underestimation of high risks have done. Nevertheless, one might object that overconfidence has been documented elsewhere. Let us then consider a second prominent definition of overconfidence.

Overconfidence 2: Mean confidence > percentage correct

The second meaning of overconfidence corresponds to its definition in the *Behavioral Economics Guide 2014*: “The overconfidence effect is observed when people’s subjective confidence in their own ability is greater than their objective (actual) performance. It is frequently measured by having experimental participants answer general knowledge test questions” (Samson, 2014, p. 21).

Studies reported that across all participants and questions, mean confidence was typically in the order of 10 percentage points above the true percentage cor-

rect. Note that there is no known theoretical relation between Overconfidence 1 and Overconfidence 2. Neither logically implies the other. Miscalibration can be observed with or without a difference in mean confidence and proportion correct. To my knowledge, there is also no empirical evidence that the two phenomena are correlated.

Overconfidence 2 was claimed to be a stable bias in people until it was shown to have emerged from researchers' selection of atypical questions (Gigerenzer *et al.*, 1991). Consider the question of whether New York or Rome lies further south. We know that people use probability cues to make an informed inference. Temperature is such a cue, which has a high validity for predicting latitude and leads to errors only in rare cases. One such case is New York, whose latitude is actually south of Rome.

Gigerenzer *et al.* (1991) found that using representative (random) sets of questions rather than selected sets, such as *all* large cities in a defined region, made Overconfidence 2 disappear. Subsequently, an analysis of 130 studies showed the generality of this result: Overconfidence 2 was limited to studies with selected questions but was practically nonexistent in studies that used representative sets (Juslin *et al.*, 2000).

In sum, by selectively using questions with "surprising" answers, researchers have mistaken intelligent inferences, as modeled by probabilistic mental models theory (Gigerenzer *et al.*, 1991) for overconfidence bias. For both Overconfidence 1 and 2, there is little evidence that the problem is located in people's minds, but evidence for a bias bias. This may explain why people seem to stubbornly stick to their judgments, why debiasing does not help in this case, and why there is a general lack of evidence that Overconfidence 1 and 2 incur substantial costs in terms of less health, wealth, or happiness (Arkes *et al.*, 2016).

One Label, Many Phenomena

Other phenomena that have been named *overconfidence* include the findings that most people believe they are better than average and that people produce too narrow confidence intervals. Neither of these is necessarily a bias; the ecological conditions first need to be specified. Consider the finding that most drivers think they drive better than average. If better driving is interpreted as meaning fewer accidents, then most drivers' beliefs are actually true. The number of accidents per person has a skewed distribution, and an analysis of U.S. accident statistics showed that some 80% of drivers have fewer accidents than the average number of accidents (Mousavi and Gigerenzer, 2011). One can avoid this problem by asking people about percentiles, but that still leaves a second problem, namely that the typical study does not define what exactly the target variable is. In the case of driving, people can pick whatever factor

they consider relevant—driving more safely or faster or without getting tired, or having fewer accidents or taking more risks—, producing the better-than-average effect. With respect to narrow confidence intervals, Juslin *et al.* (2000) and Olsson (2014) have shown how to reverse the phenomenon within the same person, suggesting a rational explanation rather than a systematic bias.

Berg and Lien (2005) showed that overtrading, another phenomenon called *overconfidence*, may shift a standard market equilibrium into a Pareto-superior equilibrium, with improved liquidity and lower transaction costs. Finally, there is a functional meaning of *overconfidence*. Analysts who earn their money by mostly incorrect predictions such as forecasting exchange rates or the stock market had better be overconfident; otherwise few would buy their advice. Conceptual clarity is desperately needed.

2.5 Framing

Neoclassical economics typically assumes stable preferences and full knowledge of the exhaustive and mutually exclusive set of future states of the world and their consequences. In this world of certainty, the framing of options (such as positive versus negative) should have no impact on choice. However, certainty does not exist for many important decisions, such as whether to undergo risky surgery or to invest in a potential real estate site. Under uncertainty, preferences are not always set and asking for advice and information can be helpful. By framing an option, the speaker can communicate information that is not contained in the verbatim message but which the intelligent listener is able to decode and incorporate into the choice accordingly. This decoding is known as “invited inferences,” which are more intellectually challenging than logical inferences.

Psychologists and psycholinguists have discovered various heuristics people rely on to make such inferences (e.g., Grice, 1989; Hertwig and Gigerenzer, 1999). For instance, listeners expect that what and how the speaker communicates is relevant—the *relevance maxim*. An example is the implication that a speaker is likely making an unspoken recommendation when using a positive frame for an option, whereas a negative frame likely indicates a warning. In general, the ability to listen carefully and pay attention to how messengers frame messages is seen as a form of intelligence, not bias (e.g., Sher and McKenzie, 2006). Similarly, the first study on defaults in organ donation tested whether people follow opt-in or opt-out defaults because they interpret these as recommendations or because following the default demands less effort (Johnson and Goldstein, 2003). They reported that when the effort to opt-in or opt-out was reduced to a single mouse click, the default nevertheless continued to substantially influence choice, consistent with the recommendation explanation and “largely eliminating the effort explanation” (p. 1338). Even so, in some later reports of this classic study, the recommendation aspect is not even

mentioned and people's behavior is attributed to their "inertia" (Thaler and Sunstein, 2008, pp. 186–188), or to "laziness" and as "embarrassing" evidence for cognitive illusions (Kahneman, 2011, pp. 373–374).

Framing has long since been considered an art in mathematics and physics, as the importance of notation and number representation illustrates. Richard Feynman (1967) pointed out the importance of working with different formulations of the same physical law, even if they are logically equivalent. "Psychologically they are different because they are completely unequivalent when you are trying to guess new laws" (p. 53).

Behavioral economists, in contrast, have used neoclassical economics as standard for rational choice, and interpreted attention to framing as a form of irrationality, as a framing bias. In the words of Conlisk in his list of biases, people "give answers that are highly sensitive to logically irrelevant changes in questions" (p. 670). As mentioned above, both phenomena—the effect of framing and defaults—became attributed to people's laziness, decision inertia, procrastination, and status quo bias. Framing effects were said to violate "description invariance," an "essential condition" for rationality (Tversky and Kahneman, 1986, p. S253). People were also said to lack insight: "in their stubborn appeal, framing effects resemble perceptual illusions more than computational errors" (Kahneman and Tversky, 1984, p. 343). Similarly, when a study found that experienced national security intelligence officers exhibited larger "framing biases" than did college students, this finding was interpreted as a surprising form of irrationality among the experienced officers (Reyna *et al.*, 2014).

Consider a classical demonstration of framing. A patient suffering from a serious heart disease considers high-risk surgery and asks a doctor about its prospects. The doctor can frame the answer in two ways:

Positive Frame: Five years after surgery, 90% of patients are alive.

Negative Frame: Five years after surgery, 10% of patients are dead.

Should the patient listen to how the doctor frames the answer? Behavioral economists say no because both frames are logically equivalent (Kahneman, 2011). Nevertheless, people do listen. More are willing to agree to a medical procedure if the doctor uses positive framing (90% alive) than if negative framing is used (10% dead) (Moxey *et al.*, 2003). Framing effects challenge the assumption of stable preferences, leading to preference reversals. Thaler and Sunstein (2008) who presented the above surgery problem, concluded that "framing works because people tend to be somewhat mindless, passive decision makers" (p. 40).

Framing effects have been presented as key justifications for the politics of libertarian paternalism. In the words of Thaler and Sunstein (2008), a "policy is 'paternalistic' if it tries to influence choices in a way that will make choosers better off, *as judged by themselves*" (p. 5, italics in original). Paradoxically,

however, paternalist policy makers rarely try to ascertain individuals' actual preferences, once again on the basis of the framing effect: people's answers or "revealed preferences" will depend on how the question is framed (Thaler and Sunstein, 2003, p. 178).

Intelligent Inferences Mistaken as Biases

Consider the surgery problem again. For the patient, the question is not about checking logical consistency but about making an informed decision. To that end, the relevant question is: Is survival higher with or without surgery? The survival rate without surgery is the *reference point* against which the surgery option needs to be compared. Neither "90% survival" nor "10% mortality" provides this information.

There are various reasons why information is missing and recommendations are not directly communicated. For instance, U.S. tort law encourages malpractice suits, which fuels a culture of blame in which doctors fear making explicit recommendations (Hoffman and Kanzaria, 2014). However, by framing an option, doctors can convey information about the reference point and make an implicit recommendation that intelligent listeners intuitively understand. Experiments have shown, for instance, that when the reference point was lower, that is, fewer patients survived without surgery, then 80–94% of speakers chose the "survival" frame (McKenzie and Nelson, 2003). When, by contrast, the reference point was higher, that is, more patients survived without surgery, then the "mortality" frame was chosen more frequently. Thus, by selecting a positive or negative frame, physicians can communicate their belief whether surgery has a substantial benefit compared to no surgery and make an implicit recommendation.

A chosen frame can implicitly communicate not only a recommendation but also other information. Consider the prototype of framing:

The glass is half full.

The glass is half empty.

Both frames are the same logically, but not psychologically. In an experiment by Sher and McKenzie (2006), a full and an empty glass were put on a table, and the participant was asked to pour half of the water into the other glass and then to place the "half empty glass" on the edge of the table. Most people chose the glass that was previously full. Here, framing conveys unspoken information, and a careful listener understands that half full and half empty are not the same.

The Asian Disease Problem Reassessed

Perhaps the most famous example of a framing effect stems from the “Asian disease problem” (Tversky and Kahneman, 1981). It has been presented as evidence of risk aversion for gains and risk seeking for losses and of the value function in prospect theory:

Imagine that the United States is preparing for an outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:

[Positive Frame]

If Program A is adopted, 200 people will be saved.

If Program B is adopted, there is a $1/3$ probability that 600 people will be saved and a $2/3$ probability that no people will be saved.

[Negative Frame]

If Program A is adopted, 400 people will die.

If Program B is adopted, there is a $1/3$ probability that nobody will die and a $2/3$ probability that 600 people will die.

The authors argued that the two problems are logically equivalent, and that invariance requires that positive versus negative framing should not alter the preference order. Nevertheless, when given the positive frame, most people favored Program A, but favored Program B when given the negative frame. This difference was interpreted as evidence that people are risk averse for gains (positive frame) and risk seeking for losses (negative frame) (Tversky and Kahneman, 1981). In this logical interpretation, the Asian disease problem refutes the assumption of stable preferences and shows that people can be easily manipulated.

Now recall the psychological analysis of the surgery problem: If people notice that part of the information is left out, such as the effect of no surgery, they tend to make inferences. In the Asian disease problem, the “risky” option is always spelled out completely in both frames (e.g., $1/3$ probability that 600 people will be saved *and* a $2/3$ probability that no one is saved), whereas the “certain” option is always incomplete. For instance, it communicates that 200 people will be saved but not that 400 will not be saved. This systematic asymmetry matters neither from the logical norm of “description invariance” nor for prospect theory, given that the framing in terms of loss and gains is preserved. But it should matter if people make intelligent inferences. To test these two competing explanations—logical error or intelligent inference—all that needs to be done is to complete the missing options in both frames. Here is the complete version for the positive frame:

If Program A is adopted, 200 people will be saved and 400 people will not be saved.

If Program B is adopted, there is a 1/3 probability that 600 people will be saved and a 2/3 probability that no people will be saved.

If any of the proposed explanations given—people’s susceptibility to framing errors, their risk aversion for gains and risk seeking for losses, or the value function of prospect theory—were true, this addition should not matter. But in fact, it changes the entire result. When the information was provided completely, the effect of positive versus negative frames disappeared in one study after the other (e.g. Kühberger, 1995; Kühberger and Tanner, 2010; Mandel, 2001; Tombu and Mandel, 2015). Further studies indicated that many people notice the asymmetry and infer that the incomplete option means that *at least* 200 people are saved because, unlike in Program B, the information for how many will not be saved is not provided (Mandel, 2014). Under this inference, Program A guarantees that 200 *or more* people are saved, as opposed to *exactly* 200. Thus, people’s judgments appear to have nothing to do with unstable preferences or with positive versus negative framing. The incompleteness of Program A alone drives the entire effect. When supplied with incomplete information, people have to make intelligent inferences. Doing so is not a bias.

Is Framing Unavoidable?

The proposal that governments should nudge people by framing messages has been defended by the argument that there is no escape from framing: Results have to be framed either positively or negatively (Thaler and Sunstein, 2008). Yet that is not necessarily so. As the examples above have shown, there is an alternative to positive or negative framing, namely to specify the message in its entirety, such as “Five years after surgery, 90% of patients are alive and 10% are dead,” or “200 people will be saved and 400 not.”

In sum, the principle of logical equivalence or “description invariance” is a poor guide to understanding how human intelligence deals with an uncertain world where not everything is stated explicitly. It misses the very nature of intelligence, the ability to go beyond the information given (Bruner, 1973).

The three sources of the bias bias I identified—failing to notice when small sample statistics differ from large sample statistics, mistaking people’s random error for systematic error, and confusing intelligent inferences with logical errors—generalize to other items in Conlisk’s list as well (e.g., Gigerenzer *et al.*, 2012; Gigerenzer *et al.*, 1988; McKenzie and Chase, 2012). There also appears to be a common denominator to these sources. If one accepts rational choice theory as a universal norm, as in Kahneman and Thaler’s version of behavioral economics, where all uncertainty can be reduced to risk, the statistics of

small samples or the art of framing play no normative role. The normative consequences of uncertainty are easily overlooked, and people's intuitions are likely misperceived as being logical errors.

Taking Uncertainty Seriously

The irrationality argument has become the backbone of behavioral economics, and listing and modeling deviations from rational choice theory has become all that keeps it erect. A new behavioral economics that can stand on its own feet would limit rational choice theory to situations of risk and focus on studying how individuals and organizations actually make decisions under uncertainty, when the assumptions of neoclassical economics do not hold (Gigerenzer and Selten, 2001). This requires formal models of heuristic decision making, and a formal analysis of the ecological conditions under which simple heuristics outperform more complex strategies (Gigerenzer *et al.*, 2011). As the examples in this article illustrate, what appears to be normative under risk is not always normative under uncertainty. Heuristics, framing, defaults, and emotions may be rarely needed in a world of certainty but are essential tools for dealing with uncertainty.

By veering onto the Simonian path, behavioral economics can liberate itself from the confines of the irrationality argument. The key is to take uncertainty seriously, take heuristics seriously, and beware of the bias bias.

References

- Aki, E. A., A. D. Oxman, J. Herrin, G. E. Vist, I. Terrenato, F. Sperati, and et al. 2011. "Using alternative statistical formats for presenting risks and risk reductions". *Cochrane Database of Systematic Reviews*. 3: CD006776.
- Anderson, J. R. 1990. *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.
- Ariely, D. 2008. *Predictably irrational: the hidden forces that shape our decisions*. 1st. New York: HarperCollins.
- Arkes, H. R., G. Gigerenzer, and R. Hertwig. 2016. "How bad is incoherence?" *Decision*. 3: 20–39.
- Barsalou, L. W. 1992. *Cognitive psychology: An overview for cognitive scientists*. Hillsdale, NJ: Erlbaum.
- Berg, N. and G. Gigerenzer. 2007. "Psychology implies paternalism? Bounded rationality may reduce the rationale to regulate risk-taking". *Social Choice and Welfare*. 28: 337–359.
- Berg, N. and G. Gigerenzer. 2010. "As-if behavioral economics: Neoclassical economics in disguise?" *History of Economic Ideas*. 18: 133–165. DOI: 10.1400/140334.

- Berg, N. and D. Lien. 2005. "Does society benefit from investor overconfidence in the ability of financial market experts?" *Journal of Economic Behavior and Organization*. 58: 95–116.
- Binmore, K. 2009. *Rational decisions*. Princeton, NJ: Princeton University Press.
- Bond, M. 2009. "Risk school". *Nature*. 461(2): 1189–1192. DOI: 10.1038/4611189a.
- Brighton, H. and G. Gigerenzer. 2015. "The bias bias". *Journal of Business Research*. 68: 1772–1784. DOI: 10.1016/j.jbusres.2015.01.061.
- Bruner, J. S. 1973. *Beyond the information given: Studies on the psychology of knowing*. Oxford, England: W. W. Norton.
- Bundesregierung [Federal Government of Germany]. 2018. "Wirksam regieren". URL: https://www.bundesregierung.de/Webs/Breg/DE/Themen/Wirksam-regieren/_node.html.
- Camerer, C., S. Issacharoff, L. Loewenstein, T. O'Donoghue, and M. Rabin. 2003. "Regulation for conservatives: Behavior Economics and the case for 'asymmetric paternalism'". *University of Pennsylvania Law Review*. 151: 1211–1254.
- Chater, N., J. B. Tenenbaum, and A. Yuille. 2006. "Probabilistic models of cognition: Conceptual foundations". *TRENDS in Cognitive Sciences*. 10: 335–344.
- Conley, S. 2013. *Against autonomy. Justifying coercive paternalism*. New York: Cambridge University Press.
- Conlisk, J. 1996. "Why bounded rationality?". *Journal of Economic Literature*. 34: 669–700.
- Dawes, R. M. and M. Mulford. 1996. "The false consensus effect and overconfidence: Flaws in judgment or flaws in how we study judgment?" *Organizational Behavior and Human Decision Processes*. 65: 201–211.
- DeBondt, W. F. and R. Thaler. 1995. "Financial decision-making in markets and firms: A behavioral perspective". In: *Handbook in Operations Research and Management Science, Vol. 9, Finance*. Ed. by R. A. Jarrow, V. Maksimovic, and V. T. Ziemba. North Holland: Elsevier. 385–410.
- Dressel, J. and H. Farid. 2018. "The accuracy, fairness, and limits of predicting recidivism". *Science Advances*. 4: eaao5580.
- Edwards, W. 1968. "Conservatism in human information processing". In: *Formal representation of human judgment*. Ed. by B. Kleinmuntz. New York: Wiley. 17–52.
- Erev, I., T. S. Wallsten, and D. V. Budescu. 1994. "Simultaneous over- and under confidence: The role of error in judgment processes". *Psychological Review*. 101: 519–527.
- Farmer, G. D., P. A. Warren, and U. Hahn. 2017. "Who 'believes' in the gambler's fallacy and why?" *Journal of Experimental Psychology: General*. 146: 63–76.

- Feynman, R. P. 1967. *The character of physical law*. Cambridge, MA: MIT Press.
- Fong, G. T., D. H. Krantz, and R. E. Nisbett. 1986. "The effects of statistical training on thinking about everyday problems". *Cognitive Psychology*. 18: 253–292.
- Friedman, M. 1992. "Do old fallacies ever die?" *Journal of Economic Literature*. 30: 2129–2132.
- Friston, K. 2010. "The free-energy principle: a unified brain theory?" *Nature Reviews Neuroscience*. 11(2): 127–138. DOI: 10.1038/nrn2787.
- Gal, D. and D. D. Rucker. 2018. "The loss of loss aversion: Will it loom larger than its gain?" *Journal of Consumer Psychology*: 497–516. DOI: 10.1002/jcpy.1047.
- Gigerenzer, G. 2014. *Risk savvy: How to make good decisions*. New York: Viking.
- Gigerenzer, G. 2015. "On the supposed evidence for libertarian paternalism". *Review of Philosophy and Psychology*. 6: 363–383. DOI: 10.1007/s13164-015-0248-1.
- Gigerenzer, G. 2016a. "Introduction: Taking heuristics seriously". In: *The Behavioral Economics Guide 2016*. Ed. by A. Samson. London: Behavioral Science Solutions. V–XI. URL: www.behavioraleconomics.com/the-behavioral-economics-guide-2016/.
- Gigerenzer, G. 2016b. "Towards a rational theory of heuristics". In: *Minds, models, and milieu: Commemorating the centennial of the birth of Herbert Simon*. Ed. by R. Frantz and L. Marsh. New York: Palgrave Macmillan. 34–59.
- Gigerenzer, G., K. Fiedler, and H. Olsson. 2012. "Rethinking cognitive biases as environmental consequences". In: P. M. Todd, G. Gigerenzer, and the ABC Research Group. *Ecological rationality: Intelligence in the world*. New York: Oxford University Press. 80–110.
- Gigerenzer, G., W. Gaissmaier, E. Kurz-Milcke, L. M. Schwartz, and S. Woloshin. 2007. "Helping doctors and patients to make sense of health statistics". *Psychological Science in the Public Interest*. 8: 53–96. DOI: 10.1111/j.1539-6053.2008.00033.x.
- Gigerenzer, G., W. Hell, and H. Blank. 1988. "Presentation and content: The use of base rates as a continuous variable". *Journal of Experimental Psychology: Human Perception and Performance*. 14: 513–525.
- Gigerenzer, G., R. Hertwig, and T. Pachur, eds. 2011. *Heuristics: The foundations of adaptive behavior*. New York: Oxford University Press.
- Gigerenzer, G. and U. Hoffrage. 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. and U. Hoffrage. 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., U. Hoffrage, and H. Kleinbölting. 1991. "Probabilistic mental models: A Brunswikian theory of confidence". *Psychological Review*. 98: 506–528.
- Gigerenzer, G. and R. Selten. 2001. *Bounded rationality: The adaptive toolbox*. Cambridge, MA: MIT Press.
- Gilovich, T., R. Vallone, and A. Tversky. 1985. "The hot hand in basketball: On the misperception of random sequences". *Cognitive Psychology*. 17: 295–314.
- Grice, H. P. 1989. *Studies in the way of words*. Cambridge, MA: Harvard University Press.
- Hahn, U. 2014. "Experiential limitation in judgment and decision". *Topics in Cognitive Science*. 6: 229–244.
- Hahn, U. and P. A. Warren. 2009. "Perceptions of randomness: Why three heads are better than four". *Psychological Review*. 116: 454–461.
- Hahn, U. and P. A. Warren. 2010. "Why three heads are better than four: A reply to Sun, Tweney, and Wang (2010)". *Psychological Review*. 117: 706–711.
- Hasher, L. and R. T. Zacks. 1984. "Automatic processing of fundamental information: The case of frequency of occurrence". *American Psychologist*. 39: 1372–1388.
- Helbing, D., B. S. Frey, G. Gigerenzer, E. Hafen, M. Hagner, Y. Hofstetter, J. van der Hoven, R. V. Zicari, and A. Zwitter. 2017. "Will democracy survive big data and artificial intelligence?" *Scientific American*. URL: <https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence/>.
- Hertwig, R. and G. Gigerenzer. 1999. "The 'conjunction fallacy' revisited: How intelligent inferences look like reasoning errors". *Journal of Behavioral Decision Making*. 12: 275–305. DOI: 10.1002/(SICI)1099-0771(199912)12:4<275::AID-BDM323>3.0.CO;2-M.
- Hertwig, R. and T. Grüne-Yanoff. 2017. "Nudging and boosting: steering or empowering good decisions". *Perspectives in Psychological Science*. 12: 973–986. DOI: 10.1177/1745691617702496.
- Hertwig, R., T. Pachur, and S. Kurzenhäuser. 2005. "Judgments of risk frequencies: Tests of possible mechanisms". *Journal of Experimental Psychology: Learning, Memory & Cognition*. 31: 621–642. DOI: 10.1037/0278-7393.31.4.621.
- Hoffman, J. R. and H. K. Kanzaria. 2014. "Intolerance of error and culture of blame drive medical excess". *British Medical Journal*. 349: g57–02. DOI: 10.1136/bmj.g5702.

- Hoffrage, U., S. Lindsey, R. Hertwig, and G. Gigerenzer. 2000. "Communicating statistical information". *Science*. 290: 2261–2262.
- Jenny, M. A., N. Keller, and G. Gigerenzer. 2018. "Assessing minimal medical literacy using the Quick Risk Test: A prospective observational study in Germany". *BMJ Open*. 8(8): e020847. DOI: 10.1136/bmjopen-2017-020847.
- Johnson, E. J. and D. G. Goldstein. 2003. "Do defaults save lives?" *Science*. 302: 1338–1339. DOI: 10.1126/science.1091721.
- Jonides, J. and C. M. Jones. 1992. "Direct coding for frequency of occurrence". *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 18: 368–378.
- Juslin, P., A. Winman, and H. Olsson. 2000. "Naive empiricism and dogmatism in confidence research: A critical examination of the hard–easy effect". *Psychological Review*. 107: 384–396.
- Kahneman, D. 2011. *Thinking, fast and slow*. London: Allen Lane.
- Kahneman, D. and A. Tversky. 1972. "Subjective probability: A judgment of representativeness". *Cognitive Psychology*. 3: 430–454.
- Kahneman, D. and A. Tversky. 1984. "Choices, values, and frames". *American Psychologist*. 39: 341–350.
- Kandel, E. 2012. *The quest to understand the unconscious in art, mind, and brain, From Vienna 1900 to the present*. New York: Random House.
- Kane, J. S. and D. J. Woehr. 2006. "Performance measurement reconsidered: An examination of frequency estimation as a basis for assessment". In: *Performance measurement: Current perspectives and future challenges*. Ed. by W. J. Bennett, C. E. Lance, and D. J. Woehr. Mahwah, NJ: Lawrence Erlbaum. 77–109.
- Kant, I. 1784. "Beantwortung der Frage: Was ist Aufklärung?" *Berlinische Monatsschrift, Dezember-Heft*. 2: 481–494.
- Kareev, Y. 1992. "Not that bad after all: Generation of random sequences". *Journal of Experimental Psychology: Human Perception and Performance*. 18: 1189–1194.
- Kühberger, A. 1995. "The framing of decisions: A new look at old problems". *Organizational Behavior and Human Decision Processes*. 6: 230–240. DOI: 10.1006/obhd.1995.1046.
- Kühberger, A. and C. Tanner. 2010. "Risky choice framing: Task versions and a comparison of prospect theory and fuzzy-trace theory". *Journal of Behavioral Decision Making*. 23(3): 314–329.
- Lewis, M. 2017. *The undoing project*. New York: Norton.
- Lichtenstein, S., B. Fischhoff, and L. D. Phillips. 1982. "Calibration of subjective probabilities: The state of the art up to 1980". In: *Judgment under uncertainty: Heuristics and biases*. Ed. by D. Kahneman, P. Slovic, and A. Tversky. New York: Cambridge University Press. 306–334.

- Lichtenstein, S., P. Slovic, B. Fischhoff, M. Layman, and B. Combs. 1978. "Judged frequency of lethal events". *Journal of Experimental Psychology: Human Learning and Memory*. 4: 551–578.
- Mandel, D. R. 2001. "Gain-loss framing and choice: Separating outcome formulations from descriptor formulations". *Organizational Behavior and Human Decision Processes*. 85(1): 56–76. DOI: 10.1006/obhd.2000.2932.
- Mandel, D. R. 2014. "Do framing effects reveal irrational choice?" *Journal of Experimental Psychology: General*. 143: 1185–1198. DOI: 10.1037/a0034207.
- McDowell, M. and P. Jacobs. 2017. "Meta-analysis of the effect of natural frequencies on Bayesian reasoning". *Psychological Bulletin*. 143: 1273–1312. DOI: 10.1037/bul0000126.
- McKenzie, C. R. M. and V. Chase. 2012. "Why rare things are precious: How rarity benefits inference". In: P. Todd, G. Gigerenzer, and the ABC Research Group. *Ecological rationality: Intelligence in the world*. New York: Oxford University Press. 309–334.
- McKenzie, C. R. M. and J. D. Nelson. 2003. "What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects". *Psychonomic Bulletin and Review*. 10: 596–602. DOI: 10.3758/BF03196520.
- Miller, J. B. and A. Sanjuro. Forthcoming. "Surprised by the gambler's and hot hand fallacies?" *Econometrica*.
- Miller, J. B. and A. Sanjuro. 2016. *Selection bias, sample-size neglect, and the persistence of gambler's fallacy beliefs*. Working Paper.
- Mousavi, S. and G. Gigerenzer. 2011. "Revisiting the "error" in studies of cognitive errors". In: *Error in organizations*. Ed. by D. A. Hofmann and M. Frese. New York: Taylor & Francis. 97–112.
- Moxey, A., D. O'Connell, P. McGettigan, and D. Henry. 2003. "Describing treatment effects to patients: How they are expressed makes a difference". *Journal of General Internal Medicine*. 18: 948–959. DOI: 10.1046/j.1525-497.2003.20928.x.
- Olsson, H. 2014. "Measuring overconfidence: Methodological problems and statistical artifacts". *Journal of Business Research*. 67: 1766–1770.
- Parikh, M. 2009. *Value investing and behavioral finance: Insights into Indian stock market realities*. New Delhi: Tata McGraw Hill.
- Peterson, C. R. and L. R. Beach. 1967. "Man as an intuitive statistician". *Psychological Bulletin*. 68: 29–46.
- Pfeifer, P. E. 1994. "Are we overconfident in the belief that probability forecasters are overconfident?" *Organizational Behavior and Human Decision Processes*. 58: 203–213.
- Piaget, J. and B. Inhelder. 1975. *The origin of the idea of chance in children*. New York: Norton.
- Raab, M., B. Gula, and G. Gigerenzer. 2012. "The hot hand exists in volleyball and is used for allocation decisions". *Journal of Experimental Psychology: Applied*. 18: 81–94. DOI: 10.1037/a0025951.

- Rabin, M. 1998. "Psychology and economics". *Journal of Economic Literature*. 36: 11–46.
- Rabin, M. 2002. "Inference by believers in the law of small numbers". *Quarterly Journal of Economics*. 117: 775–816.
- Rabin, M. and D. Vayanos. 2010. "The gambler's and hot-hand fallacies: Theory and applications". *Review of Economic Studies*. 77: 730–778.
- Rebonato, R. 2012. *Taking liberties: A critical examination of libertarian paternalism*. Basingstoke, United Kingdom: Palgrave Macmillan.
- Reyna, V. F., C. F. Chick, J. C. Corbin, and A. N. Hsia. 2014. "Developmental reversals in risky decision making: Intelligence agents show larger decision biases than college students". *Psychological Science*. 25(1): 76–84.
- Saffran, J. R., R. N. Aslin, and E. L. Newport. 1996. "Statistical learning by 8-month-old infants". *Science*. 274: 1926–1928.
- Samson, A. 2014. *Behavioral Economics Guide 2014*. URL: <https://www.behavioraleconomics.com/BEGuide2014.pdf>.
- Schneider, S. 2010. *Homo economicus – or more like Homer Simpson?* (Report). Deutsche Bank Research.
- Sedlmeier, P. and G. Gigerenzer. 2001. "Teaching Bayesian reasoning in less than two hours". *Journal of Experimental Psychology: General*. 130: 380–400.
- Sedlmeier, P., R. Hertwig, and G. Gigerenzer. 1998. "Are judgments of the positional frequencies of letters systematically biased due to availability?" *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 24: 754–770.
- Shepard, R. N. 1990. *Mind sights*. New York: Freeman.
- Sher, S. and C. R. M. McKenzie. 2006. "Information leakage from logically equivalent frames". *Cognition*. 101: 467–494.
- Silver, N. 2012. *The signal and the noise*. New York: Penguin.
- Simon, H. A. 1979. "Information processing models of cognition". *Annual Review of Psychology*. 30: 363–396.
- Simon, H. A. 1989. "The scientist as problem solver". In: *Complex information processing: The impact of Herbert A. Simon*. Ed. by D. Klahr and K. Kotovsky. Hillsdale, NJ: Erlbaum. 375–398.
- Slovic, P., B. Fischhoff, and S. Lichtenstein. 1982. "Facts versus fears: Understanding perceived risk". In: *Judgment under uncertainty: Heuristics and biases*. Ed. by D. Kahneman, P. Slovic, and A. Tversky. Cambridge, United Kingdom. 463–489.
- Smith, V. L. 2007. *Rationality in economics*. New York: Cambridge University Press.
- Stigler, S. M. 1999. *Statistics on the table*. Cambridge, MA: Harvard University Press.
- Thaler, R. H. 1991. *Quasi rational economics*. New York: Russell Sage Foundation.

- Thaler, R. H. 2015. *Misbehaving: The making of behavioral economics*. New York: Norton.
- Thaler, R. H. and C. R. Sunstein. 2003. "Libertarian paternalism". *American Economic Review*. 93(2): 175–179.
- Thaler, R. H. and C. R. Sunstein. 2008. *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.
- Tombu, M. and D. R. Mandel. 2015. "When does framing influence preferences, risk perceptions, and risk attitudes? The explicated valence account". *Journal of Behavioral Decision Making*. 28: 464–476. DOI: 10.1002/bdm.1863.
- Tversky, A. and D. Kahneman. 1973. "Availability: A heuristic for judging frequency and probability". *Cognitive Psychology*. 4: 207–232.
- Tversky, A. and D. Kahneman. 1974. "Judgment under uncertainty: Heuristics and biases". *Science*. 185: 1124–1131.
- Tversky, A. and D. Kahneman. 1981. "The framing of decisions and the psychology of choice". *Science*. 211: 453–458.
- Tversky, A. and D. Kahneman. 1983. "Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment". *Psychological Review*. 90: 293–315. DOI: 10.1037/0033-295X.90.4.293.
- Tversky, A. and D. Kahneman. 1986. "Rational choice and the framing of decisions". *The Journal of Business*. 59(4): S251–S278.
- Zarembo, A. 2003. "Funding studies to suit need". *Los Angeles Times*. URL: <http://www.latimes.com/nation/la-na-valdez3-2003dec03-story.html>.