

What Is Bias? And Why Are We Biased?

Gerd Gigerenzer

▶ **To cite this version:**

Gerd Gigerenzer. What Is Bias? And Why Are We Biased?. Proceedings of the Paris Institute for Advanced Study, 2025, 21, ⟨10.5281/zenodo.17901966⟩. ⟨hal-05412045⟩

HAL Id: hal-05412045

https://hal.science/hal-05412045v1

Submitted on 11 Dec 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

PPIAS

What Is Bias? And Why Are We Biased?

Gerd Gigerenzer[1]

[1]Max Planck Institute for Human Development, Berlin, Germany

TO CITE

Gerd, G. (2025). What Is Bias? And Why Are We Biased?. *In Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.17901966

PUBLICATION DATE

11/12/2025

ABSTRACT

I distinguish two conceptions of bias. First, bias can be understood as unwanted error that hampers cognition and should be eliminated. Second, it can be understood as functional; it is necessary for cognition. The dominant interpretation in the social sciences is the former, which links biases to prejudice or cognitive errors. It is based on norms such as consistency, coherence, and social norms. I provide examples for functional biases that protect individuals and enable perception. Finally, I provide a general perspective on the apparent contradiction between bias as error and functional bias, based on the distinction between small and large worlds. My hypothesis is that bias is detrimental and without any function in small worlds. In all other situations that involve a substantial degree of uncertainty, bias is likely to be necessary and functional, or even rational.

There is something irritating about the use of the term *bias* in the social sciences. First, the term is highly invasive. Although prejudices against women, people of color, and minorities have been its classical topic, beginning in the 1960s and 70s, the term was extended beyond emotional or motivational prejudice to include cognitive defects. Bias spreads from the guts to the intellect. Today, we have an avalanche of cognitive biases: the hot-hand fallacy, overconfidence bias, framing bias, base rate fallacy, loss aversion, conjunction fallacy, and dozens more mental quirks. Given this proliferation, a colleague of mine used to ask: What is the bias of the day?

My second observation is that the increasing use of the term *bias* is itself often based on a bias: the tendency to spot biases even if there are none (see next section). For instance, the experience of basketball coaches and fans that players can have a "hot streak" has been said to be a systematic bias in coaches and fans' intuitions – the hot-hand fallacy (Gilovich, Vallone & Tversky, 1985). But ultimately, there turned out to be no bias: coaches and players had correctly recognized that the hot hand indeed exists. This flaw originated in the "hallucinations" of researchers who were overly eager to detect a new bias (Miller & Sanjuro, 2018).

My third observation is that in some research programs, bias is considered something that should be reduced to zero, while in others, it is deemed necessary for cognitive functioning. That

This is an open access article published under the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0)

apparent contradiction leads to the surprising fact that the same phenomenon is interpreted by one group of researchers as a bias, and by others as rational. For instance, according to social psychologists, the reliance on base rates is dismissed as stereotype and prejudice, while for cognitive psychologists, attention to base rates is rational, as prescribed by Bayes' rule. How can we reconcile these antagonistic evaluations of biases? What can biases do for us, and are they more than prejudice and error?

In this essay, I will try to bring some clarity into the existing confusion and will focus on theories of human cognition, such as perception, learning, and judgment.

What is bias?

The term *bias* has multiple meanings, some of which can be traced back to the 16th century and to the French term *biais*, meaning oblique direction or slant. Bias originally described the tendency of a bowl to swerve in lawn bowling, the strategy of cutting textiles against the grain of the fabric, and both right-handedness and left-handedness. Current definitions, in contrast, associate the term with predominantly negative meanings.

The Merriam-Webster Dictionary lists several other definitions of the noun *bias* besides these historical ones. The first is "an inclination of temperament or outlook, especially: a personal and sometimes unreasoned judgment: PREJUDICE."[1] Examples given are teachers' biases about their pupils, and economists' bias that only work resulting in a physical output is productive. A second definition is statistical, such as biased sampling techniques. Similarly, the Cambridge Dictionary defines bias as "the action of supporting or opposing a particular person or thing in an unfair way, because of allowing personal opinions to influence your judgment."[2] An example given is unconscious bias. A third meaning is "preferring a particular subject or thing," such as showing "a scientific bias at an early age." The fourth definition refers to flaws in data collection and reporting, such as selection bias and publication bias.

As these definitions illustrate, only a few of the modern meanings of bias are neutral, such as preferring a particular subject to study. Notably, functional or positive valuations are virtually absent.

In this essay, I will focus on cognition and argue that there are two conceptions of biases (Gigerenzer, 2025):

Bias is error: Biases hamper cognition and should be eliminated.
Bias is functional: Biases are necessary for cognition.

Both views share a common definition: Bias is systematic deviation of a judgment, conscious or unconscious, from a norm, such as a societal value or the true state of the world. Moreover, the same belief can be an error or functional. For instance, to believe that the world is flat is a bias in the sense of an error. To treat the world as if it were flat when measuring short distances is a bias that is functional.

Bias is error

The most widely used meaning of the term *bias* is that it amounts to an error or violation of a norm. Stereotypes of people are one example. In the field of management, thousands of articles on managerial bias have been published, virtually all of which treat bias as potentially harmful and something to be eradicated. Consulting firms offer debiasing workshops that companies make part of their corporate training, such as on biases in executives and strategic decision making, or on how to mitigate biases in organizational decisions.

Algorithms can inherit these biases from human data. When OpenAI tested its large language models, it found such stereotypes. GPT described females by their appearance, while the words associated with males were both more positive and more diverse (Briggs & Briggs, 2024).

Similarly, in behavioral economics, the term *bias* is virtually always used in the sense of a fallacy or systematic cognitive illusion that violates logic, probability, or some other norm. The most frequently used norms are internal consistency, external correspondence, and social norms.

Consistency norms

Deviations of judgment from logical and probabilistic rules have been interpreted as cognitive biases. Examples include the base rate fallacy, conjunction fallacy, and framing bias. For instance, when people react differently to how a message is framed ("if you choose to undergo surgery, you have a 90% chance of surviving" versus "you have a 10% chance of dying"), this has been interpreted as a bias in the sense of an error (Kahneman, 2011).

Correspondence norms

Deviations of judgment from states of the world have also been interpreted as cognitive biases. Examples are the overconfidence bias, anchoring bias, and planning fallacy. For instance, large-scale capital investment projects, such as building airports or opera houses, typically cost more than projected and are finished later than planned. This has been attributed to cognitive biases, in particular to the optimism bias and the planning fallacy (Lovallo & Kahneman, 2003).

Social norms

Deviations of judgments from social norms are a third class of biases in the sense of error. Examples include stereotypes about social groups and gender. For instance, males and females in Germany and Spain believed that women have better intuitions about choosing the right romantic partner and understanding intentions, while men were believed to have better intuitions in scientific discovery and investments in stocks (Gigerenzer et al., 2014).

Bias Bias

The bias bias is the tendency to see biases even when there are none. That can happen because researchers locate the cause of societal problems solely in individual minds while ignoring strategic and systematic deception, or they overlook the ambiguities of language, which require people to make intelligent inferences. For instance, doctors tend to use positive frames to

recommend surgery and negative frames to warn against it; intelligent patients understand these implicit messages conveyed by framing and read between the lines (McKenzie & Nelson, 2003). Making intelligent inferences is not an error; in this case, it could benefit one's health. Thus, much framing bias may be little more than researchers' bias bias. Or take the planning fallacy. As Flyvbjerg (2025) argued, underestimating costs and/or the duration of megaprojects is a strategic misrepresentation by powerful managers and planners. Thus, the cause is likely not managers' unconscious overconfidence, but deliberate strategy: to acquire and justify projects to keep funds flowing.

Bias is functional

The error view of bias assumes a correct state of the world, or a correct value, against which a deviating judgment is identified as error. In contrast, the functional interpretation of bias does not need to know the truth. There is no need to assume that consistency, coherence, or a given social norm is a general norm. Rather, the functional view is focused on getting things done even without being sure of the best path. Biases, in this view, help decision-makers in a world of uncertainty, where the true state is not known or not knowable, yet where actions are nevertheless imperative. The functional interpretation aligns with a pragmatic view of the world, best known from the philosophical pragmatism of William James, John Dewey, and Charles Sanders Peirce.

Here are two examples.

Biological preparedness

The idea that the laws of learning are domain-general flourished in Skinner's theories. Reinforcement learning, so the assumption, uniformly applies to all stimuli and all behavior. This assumption was challenged and rejected following experimental research by John Garcia and others. For instance, when the taste of flavored water is repeatedly paired with an electric shock immediately after the tasting, rats have great difficulty in learning to avoid the flavored water. In contrast, rats can learn to avoid flavored water after just one trial when it is followed by experimentally induced nausea, even if the nausea occurs two hours later. "From an evolutionary view, the rat is a biased learning machine designed by natural selection to form certain CS–US [conditional stimulus – unconditional stimulus] associations rapidly but not others." (Garcia y Robertson & Garcia, 1985). In contrast, from the point of traditional learning theory, the rat was seen as an unbiased learner.

Rhesus monkeys reared in the laboratory show no fear toward poisonous snakes. If, however, the monkey observes another monkey exhibiting fear of snakes, it can acquire the fear reaction in a single trial. Note that this fast learning does not happen with all stimuli. If the monkey observes the same fear reaction toward a flower, it will not acquire fear of flowers. The general conclusion from these experiments is that fear of *certain* objects, such as snakes and spiders, is genetically prepared to be acquired immediately if a conspecific shows fear of them. These prepared objects were dangerous in the evolutionary past.

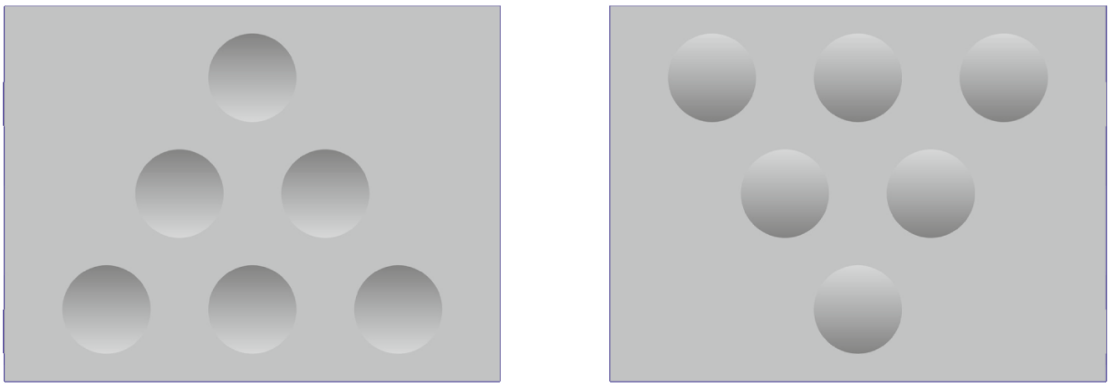
Thus, biological preparedness enables certain relations or objects to be learned in a single trial. This can be through individual experience, as in the cause of nausea, or even less dangerously, through social imitation of a conspecific (Mineka & Cook, 1988).

Biological preparedness also applies to humans. Humans learn fear and avoidance behavior faster for prepared objects, such as snakes and spiders, than for other objects – a learning mechanism that explains the sudden emergence of phobias. Preparedness relates to objects that were dangerous in the past, not to dangerous objects per se. If a parent shows fear of an electric outlet or a knife, the child does not learn to fear these objects as rapidly. But for the prepared objects, social learning can avoid individual learning from bitter experience, a bias that protects the organism.

Unconscious inferences

Without bias, humans would perceive little but chaotic sensory input; with bias, we see what the it allows us to see. Consider Figure 1. In the left-hand image in Figure 1, we see six large dots that appear to be concave; on the right, the six dots appear to be convex (Kleffner & Ramachandran, 1992). After rotating both images 180 degrees, the concave dots suddenly appear to be convex, and the convex dots concave. The illusion remains stable, independent of how often one repeats the action.

Figure 1 Perception is inference



Note: On the left-hand side, we see large dots that are curved inward (concave); on the right-hand side, they pop out like bubbles (convex). If the images are turned upside down, the brain will make the opposite inferences (based on Kleffner & Ramachandran, 1992).

It is important to recall that the brain cannot perceive the world with certainty; it can only make intelligent inferences. Here, it faces a daunting task: to generate a three-dimensional object from a two-dimensional retinal image. In doing so, the brain draws on three assumptions: the world is three-dimensional, light comes from above, and there is only one source of light.

These assumptions are biases. They are not necessarily true. We do not know whether the real world – if it exists – is truly three-dimensional; there could be more dimensions. But the bias reflects human experience: We move in space from south to north, west to east, and up to down. Similarly, through most of mammalian history, light came from above, and there was only one source of light – the sun or the moon. Even in the contemporary world, light generally comes from above in homes, offices, and other buildings, and light coming at us horizontally, such as by the headlights of another vehicle while driving, can cause confusion. At issue is not objective precision but rather that, by means of these biases, brains can make intelligent inferences. Our brains use a heuristic to infer the third dimension from the position of shading:

Shade Heuristic:

If the shade is in the upper part, then infer that the dots are concave;
if the shade is in the lower part, then infer that the dots are convex.

In Figure 1 (left), the upper parts of the dots are darker and less distinct. Thus, the brain infers that the dots are concave. In the same figure (right), the shade is in the lower part, and the dots are seen as convex.

Without these three biases, along with the shade heuristic, the brain would not see a third dimension. We would live in Flatland. In fact, sight-recovery patients with brain damage do not see the third dimension in similar visual illusions (e.g., Fine et al., 2003). All they perceive is the two-dimensional picture, nothing more.

Mistaking functional bias for irrationality

Visual illusions have been misinterpreted as biases that should be eliminated. For instance, bestsellers featuring cognitive biases regularly compare these to visual illusions (e.g., Ariely, 2008; Thaler & Sunstein, 2008). The message is that humans would fare better by seeing exactly what is on the paper rather than by making intelligent inferences. As mentioned above, however, doing so would put us in the situation of patients with brain damage.

The heuristics-and-biases program is best known for the bias-as-error view. However, we know by now that not only visual illusions but also many cognitive illusions have been mistaken for errors in need of elimination (for an overview, see Gigerenzer, 2008).

In the final part, I provide a general perspective on the apparent contradiction between bias as error and functional bias, based on the distinction between small and large worlds by Savage (1954).

Small and Large Worlds

My hypothesis is that this apparent contradiction can be resolved by using the distinction made by Savage between small (closed) worlds and large (uncertain) worlds. Specifically, the hypothesis is that bias is detrimental and without function in small worlds. In all other situations

that involve a substantial degree of uncertainty, bias might be necessary and functional, or even rational.

Savage became famous for laying the foundations of *subjective* expected utility theory, which he restricted to small worlds. He defined a small world by two features:

1. *Perfect Foresight of Future States*: The agent knows the exhaustive and mutually exclusive set *S* of future states of the world.
2. *Perfect Foresight of Consequences*: The agent knows the exhaustive and mutually exclusive set *C* of consequences (outcomes) of each of their actions, given a state.

In plain words, a small world is a situation where one knows for certain *all* possible future states of the world and *all* possible consequences of one's actions. The game of roulette is an example of a small world. Here, every possible outcome is known: the next ball will land on a number between 0 and 36, as will all those that follow.

Biases in Small Worlds

In a small world such as roulette, a bias is the deviation between a judgment and a known true state. For instance, believing that after the first two rounds showed red, the next roll will result in a black number, is known as the gambler's fallacy. It is a bias in the sense of an error because we know the probabilities for sure (assuming the table is not rigged). It is hard to see how this bias could be functional. In general, in small worlds, biases amount to errors.

Biases in Large Worlds

In a large world, there is no perfect foresight of all possible future states and their consequences. That characterizes the real world we live in, with exceptions such as well-defined games. In a large world, the brain needs biases for learning and perception; without them, it could not function. Thus, in large worlds, biases are needed in order to be prepared for learning rapidly what objects pose potentially life-threatening dangers or to perceive more than sensory chaos. Note that while biases are necessary, they can nevertheless mislead us. When individuals learn to fear spiders in a country where poisonous spiders no longer exist, there is no benefit. But without any bias at all, survival would be next to impossible. Even artificial neural networks, such as convolutional neural networks (CNN), implement systematic biases to learn faster and more efficiently. In general, models that are biased – from fast-and-frugal heuristics to CNNs – can enable better predictions (Gigerenzer & Brighton, 2009).

Why are we biased?

In this essay, I distinguished between two kinds of biases: bias as error and functional bias. Most frequently, biases are associated with errors such as prejudices or cognitive biases. I introduced Savage's distinction between small worlds and large worlds to argue that both interpretations have their place. In small worlds where all potential events and consequences, including their probabilities, are known, biases can be defined as a discrepancy from the true value.

Here, biases are errors and have no function. In large worlds, where the future or even the present state of the world is unknowable, biases can be functional. They help humans to reach goals.

References

Ariely, D. (2008). *Predictably irrational: the hidden forces that shape our decisions*. HarperCollins.

Fine, I., Wade, A. R., Brewer. A. A., May, M. G., Goodman, D. F., Boynton, G. M., Wandell, B. A., & MacLeod, D. A. (2003). Long-term deprivation affects visual perception, *Nature Neuroscience, 6*, 915–916.

Flyvbjerg, B. (2025). Why behavioral economics is only half right: An adversarial collaboration with Daniel Kahneman. Manuscript.

Garcia y Robertson, R. & Garcia, J. (1985). X-rays and learned taste aversions: Historical and psychological ramifications. In T. G. Burish, S. M. Levy, & B. E. Meyerowitz (Eds.), *Cancer, nutrition and eating behavior: A biobehavioral perspective* (pp. 11–41). Erlbaum.

Gigerenzer, G. (2018). The bias bias in behavioral economics. *Review of Behavioral Economics, 5,* 303–336.

Gigerenzer, G. (2025). Two kinds of bias. *Mind & Society, 24*, 193–207.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1,* 107–143.

Gigerenzer, G., Galesic, M., & Garcia-Retamero, R. (2014). Stereotypes about men's and women's intuitions: A study of two nations. *Journal of Cross-Cultural Psychology*, *45*, 62–81.

Gilovich, T., Vallone, R., & Tversky, A. 1985. "The hot hand in basketball: On the misperception of random sequences". Cognitive Psychology. 17: 295–314.

Kahneman, D. (2011). *Thinking, fast and slow*. Allen Lane.

Kleffner, D. A., & Ramachandran, V. S. (1992). On the perceptions of shape from shading. *Perception & Psychophysics*, *52*, 18–36.

Lovallo, D., & Kahneman, D. (2003). Delusions of success: How optimism undermines executives' decisions," *Harvard Business Review*, July Issue, pp. 56–63.

McKenzie, C. R. M., & Nelson, J. D. (2003) What a speaker's choice of frame reveals: reference points, frame selection, and framing effects. *Psychonomics Bulletin Review, 10*, 596–602

Miller J. B., & Sanjuro, A. (2018) Surprised by the hot hand fallacy? A truth in the law of small numbers. *Econometrica 86*, 2019–2047.

Mineka, S., & Cook, M. (1988). Social learning and the acquisition of snake fear in monkeys. In T. R. Zentall & B. G. Galef (Eds.), *Social learning: psychological and biological perspectives* (pp. 51–73). Erlbaum.

Savage, L. J. (1954/1972). *The foundations of statistics* (2nd ed.) Wiley

Thaler, R. H. , & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Endnotes

[1] https://www.merriam-webster.com/dictionary/bias
[2] https://dictionary.cambridge.org/de/worterbuch/englisch/bias