

# Can psychology learn from the natural sciences?

Theory & Psychology  
2024, Vol. 34(3) 295–310  
© The Author(s) 2023



Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: 10.1177/09593543231209342  
[journals.sagepub.com/home/tap](https://journals.sagepub.com/home/tap)



**Gerd Gigerenzer**

Max Planck Institute for Human Development

## Abstract

I argue that psychology can learn from the natural sciences and focus on the weight that physics attributes to precise theories. Much of psychology can be increasingly characterized by theory aversion—yet not the kind motivated by positivism. Theory aversion in psychology arises from a conflict between two desires: to come up with a theory, and to avoid the necessary mental effort and time as well as the risk of refutation. The results are ersatz theories, or surrogates. I outline three common, but independent, research practices that avoid building precise theories of psychological processes: *the null ritual*, which allows researchers to get away with not specifying their research hypothesis; *as-if theories*, which refrain from modeling psychological processes; and *lists of binary oppositions*, as in dual-system theories, which consist of vague dichotomies. Psychologists could learn from physics to walk forward on two feet—theory and experiment—rather than hobble on one.

## Keywords

as-if theories, dual-system theories, ersatz theories, psychology and physics, theory aversion

Science walks forward on two feet, namely theory and experiment . . .  
Sometimes it is one foot that is put forward first, sometimes the  
other, but continuous progress is only made by the use of both.

(Millikan, 1924, p. 54)

When Kieran O’Doherty kindly invited me to join a debate on “Should psychology follow the methods and principles of the natural sciences?” I was puzzled. The natural sciences have namely no single method or set of methods. Newton’s prism experiments have little in common with Einstein’s thought experiments. Neither relied on statistical

---

## Corresponding author:

Gerd Gigerenzer, Max Planck Institute for Human Development, Lentzeallee 94, Berlin, 14195, Germany.  
Email: [gigerenzer@mpib-berlin.mpg.de](mailto:gigerenzer@mpib-berlin.mpg.de)

methods, nor do they share much with, say, the large-particle-collider experiments on the Higgs boson or with the CRISPR-Cas9 genome editing technique.

So I told Kieran that question he had asked me to answer was wrong. He responded that he posed the question because of the prevailing sentiment among psychologists of their field being a “science” without reflecting what exactly that means. That is a revealing observation. Moreover, the concern about being recognized as a “science” is in part a consequence of the English language, where “science” includes, for instance, physics but not psychology. For German-speaking psychologists, “Wissenschaft” (science) includes psychology, as well as empirical fields such as sociology and history, meaning that no such anxiety about being excluded from the science club and longing to be accepted arises. Similarly, there is a reverse sentiment that psychology is unlike a “science,” which is equally puzzling—psychology and the natural sciences are not unrelated endeavors. In fact, psychology and physics have long been involved in a fruitful exchange. Let me begin with a few examples.

### *Psychology inspires physics, physics inspires psychology*

The physicist Gustav Theodor Fechner, known as the father of psychophysics, held that the internal and external world were two sides of the same coin, a philosophical position called monism (Heidelberger, 1987). By demonstrating empirically that an exact relation between the psychological and the physical exists—the psychophysical function—Fechner intended to prove monism. Less well-known is that Fechner used the psychological concept of “free will” to argue that the physical world cannot be deterministic as Newton and, later, Einstein thought, but must be of a similarly indeterministic nature (Heidelberger, 1987). Fechner thereby established an ontic view of probability, where probability exists in nature, unlike the reigning epistemic view in physics, where probability exists only in the mind. In this way, Fechner became the first indeterminist, foreshadowing the Copenhagen interpretation of quantum physics (Heidelberger, 1987). Similarly, the physicist Ernst Mach ranked psychology above the physical by recognizing only sensations as real (B. F. Skinner is reported to have carried a copy of Mach’s [1883/2013] *Science of Mechanics* in the back pocket of his trousers). Thomas Kuhn (1970) borrowed generously from Gestalt psychology, particularly the perceptual Gestalt shifts, which provided him the analogy to understand paradigm shifts in physics and astronomy.

The effect of a psychological concept on the natural sciences can spread over several disciplines. For instance, 19th-century astronomers noted that when judging the position of a star, human observational errors form a normal distribution around the true position of the star. The Belgian astronomer Adolphe Quetelet, who founded social physics (later called sociology), observed that the behavior of individuals (from crime to suicide to marriage) was erratic and largely unpredictable, but viewed as a collective, their means and variances were stable and predictable (Porter, 1986). He used the normal distribution for observational error as a theory of society: the true position of a star (the mean) translated into *l’homme moyen*, the ideal human, and observational errors into individuals’ deviation from the ideal. Physicists Ludwig Boltzmann and James Clerk Maxwell read Quetelet’s social physics while pondering the erratic behavior of gas molecules and

reasoned that molecules might behave like humans do, unpredictable as individuals but predictable as a collective (Porter, 1986). The result was the discovery of statistical mechanics, relying on the same normal distribution. This route of discovery—from astronomers' model of human observational errors, to the moral statistics of social systems, to the mechanics of gas molecules—led to a statistical view of nature that finally overthrew the deterministic Newtonian world view (Gigerenzer et al., 1989). As a final example, the physicist Richard Feynman (1967) emphasized the importance of trying different framings of the same physical law because what is mathematically the same can be psychologically different and inspire new insights (p. 53). The case of Feynman illustrates that physicists may value psychology more than do some contemporary social scientists who consider people who pay attention to framing as “somewhat mindless” (Thaler & Sunstein, 2008, p. 39).

Physics has also influenced psychological theories. For instance, the first principle of thermodynamics, the “conservation of force,” heavily shaped Freud's theorizing—thermodynamic energy turned into his idea of nervous energy—and, a century later, thermodynamics again inspired Karl Friston's neuroscientific notion of “free energy” (The et al., 2018). Similarly, quantum theory inspired several psychological theories of memory and decision making, including quantum decision theory (Busemeyer et al., 2015).

This mutual exchange illustrates how fruitful crossing disciplinary borders can be, including for physics. A similar point can be made for the exchange between the life sciences and psychology. In sum, the key question is not whether psychology is a natural science or not, which is mainly a matter of language and definition or of political agenda. Instead, let me pose a more productive question: What can psychologists learn from the natural sciences?

I will focus solely on one aspect: the weight that physics attributes to precise theories. Note that psychological theories may not be as general as those in physics but instead tend to resemble those in chemistry and biology, that is, consist of a “toolbox” of adaptive, middle-range processes. In place of the universal laws that Shepard (2004) envisions in psychology, my topic is more modest, simply precision—that is, theories that specify psychological processes in detail and make clear and bold predictions.

## Surrogates for theory

When the first editorial team for *Theory & Psychology* met, we discussed what name to give the new journal. Some of us proposed “Journal of Theoretical Psychology.” After all, there is a *Journal of Theoretical Biology* and an *International Journal of Theoretical Physics*, while economic journals do not even bother with this qualifier because articles without formal theory are mostly desk rejected. As I recall, however, the proposed title encountered resistance, including from the publisher, who feared that such a precarious name would not sell (Gigerenzer, 2010). Few psychologists, it was postulated, would want to read a journal with anything “theoretical” in the title. The sad truth is that these objections were not off the mark. The final compromise was *Theory & Psychology*, which both distanced psychology from the problematic term and connected the two.<sup>1</sup>

What is wrong with theory in psychology? Not all areas in psychology have a problem with it. Behaviorists, although often looked down upon by cognitive psychologists,

have strong theories of operant and respondent learning, including reinforcement schedules, avoidance and discrimination learning, matching law, chaining, and their cognitive extensions (Staddon, 2021). Skinner's concept of intermittent reinforcement has gained new prominence after tech companies began to exploit it to glue users to their social media platforms (Gigerenzer, 2022). Among the more cognitively oriented theorists, Herbert Simon (1979) developed theories of heuristic search such as satisficing, and Roger Shepard (2004) and Amos Tversky (1977) proposed precise theories of similarity. Many more examples could be added.

At the same time, theory has become a foreign word for many researchers, who fill the void with surrogates for theory. There is nothing wrong with having no explanation for a finding and admitting it. The problem arises when surrogates are used in order to pretend that an explanation exists. *Redescription* is a case in point: attributing an observed behavior  $X$  to an essence or trait  $X$ . Moliere ridiculed this practice long ago: Why does opium make us sleepy? Explanation: because of its dormative properties (Gigerenzer, 1996). Redescription abounds in published research, a fact noted long ago (e.g., Gigerenzer, 2009, 2010; Katzko, 2006; Wallach & Wallach, 1994). A choice of  $X$  is explained by a "preference" for  $X$ , the fact that a decision is influenced by affect is explained by an "affect heuristic," availability bias is explained by an "availability heuristic," or the differential effect of representations of information is attributed to their different "salience."

In what follows, I contribute to the study of *theory aversion*. This is neither a new topic nor a new diagnosis. Meehl (1967) took notice long ago when he observed a "free reliance upon ad hoc explanations to avoid refutation" (p. 103) and Kruglanski (2001) pondered on social psychologists' risk aversion and their lack of courage as possible causes for "our theoretical aversions" (p. 871).

By theory aversion, I do not mean those variants of positivism that insist that researchers should record only what they observe and refrain from speculation about unobservable causes. Mach, for instance, rejected the atomic theory in physics because the atom could not be seen. Influenced by Mach, Skinner rejected theorizing about cognitive processes because these were unobservable. The theory aversion I have in mind is different. Instead of embracing positivism, it results from an approach-avoidance conflict. Theory aversion arises from a conflict between two desires: the desire to come up with a theory and the desire to avoid the necessary mental time and effort as well as the risk of refutation. The results are ersatz theories, or surrogates.

I will outline three common, but independent, research practices that avoid building precise theories of psychological processes: (a) *the null ritual*, which specifies a precise null hypothesis and allows researchers to get away with not specifying their research hypothesis; (b) *as-if theories*, which do not model psychological processes in the first place; and (c) *lists of binary oppositions*, which avoid specifying psychological processes.

As Robert Millikan (1924) reminds us (see epigraph), science walks on two feet: theory and experiment. Theory aversion means that many psychologists hobble along on one leg and a crutch: the leg is experimentation, the crutch is a surrogate for theory.

## How to avoid theory: The null ritual

In his 1967 article “Theory-Testing in Psychology and Physics,” Paul Meehl pointed out a methodological paradox. In physics, improvements in experimental method and amount of data make it harder for a theory to pass the test whereas in psychology it is the reverse. Theories in physics typically make a point prediction, and thus higher precision in measurement makes it easier to detect a difference between prediction and reality. Theories in psychology, by contrast, rarely make point predictions and sometimes not even a directional prediction. Instead, an unspecified prediction is tested against the point prediction of a null hypothesis (e.g., “no difference”). Hence, more data, more participants, and more statistical power make it easier to detect a difference between prediction (of the null) and reality, and to accept the unspecified theory (or hypothesis).

Since the mid-1950s, a majority of psychologists follow a method of hypothesis testing that I have baptized the “null ritual” (Gigerenzer, 2004, p. 588):

1. Set up a statistical null hypothesis of “no mean difference” or “zero correlation.” Do not specify the predictions of your research hypothesis or of any alternative substantive hypotheses.
2. Use 5% as a convention for rejecting the null. If significant, accept your research hypothesis. Report the result as  $p < .05$ ,  $p < .01$ , or  $p < .001$  (whichever comes next to the obtained  $p$ -value).
3. Always perform this procedure.

The null ritual is what Meehl had in mind. It has sophisticated aspects I will not cover here, such as alpha adjustment and ANOVA procedures. But these do not change its essence. Some psychologists mistakenly perform the ritual because they think it follows the scientific method. Yet, as Meehl (1967) clarified, it is not practiced in the natural sciences and, if anything, is contrary to scientific method. Nor does the ritual exist in statistics proper (Gigerenzer, 1993). Rather, it was invented by psychologists who wrote textbooks on statistics and is enforced by journal editors who, confusing significance with quality of research, use significance as a screening tool for accepting papers. The great advantage of the null ritual is that one does not have to specify one’s prediction and, thus, any theory in the first place. Instead, one rejects a null and claims victory.

The null ritual is sometimes confused with Ronald A. Fisher’s statistical theory of significance testing. Fisher (1955, 1956), however, did not mean that the hypothesis tested postulates a nil difference (Step 1); the term null hypothesis signifies a hypothesis to be nullified. Rather, the null hypothesis could be a treatment effect of 100 milliseconds, a percentage correct of 64, or a loss of 10 IQ points, whatever a theory predicts. In this way, the research hypothesis becomes the null hypothesis, as Meehl (1967) pointed out for physics. Unlike Step 2, Fisher insisted that researchers should publish the exact level of significance, such as  $p = .03$ . In fact, Fisher rejected the use of a fixed level of significance and the practice of classifying results as “significant” and “not significant” (Step 2), comparing it to Soviet 5-year plans (Gigerenzer et al., 1989). Step 3 would have been absolutely unacceptable to Fisher, and also to statisticians Jerzy Neyman and Egon

Pearson, who otherwise disagreed with Fisher on almost everything else. Statistical inference should not be automatic; it requires judgment (Gigerenzer, 2004).

The null ritual is strikingly similar to social rites that include the following elements (Dulaney & Fiske, 1994): (a) sacred numbers or colors, (b) repetition of the same actions, (c) fear about being punished if one stops performing the actions, and (d) wishful thinking and delusions.

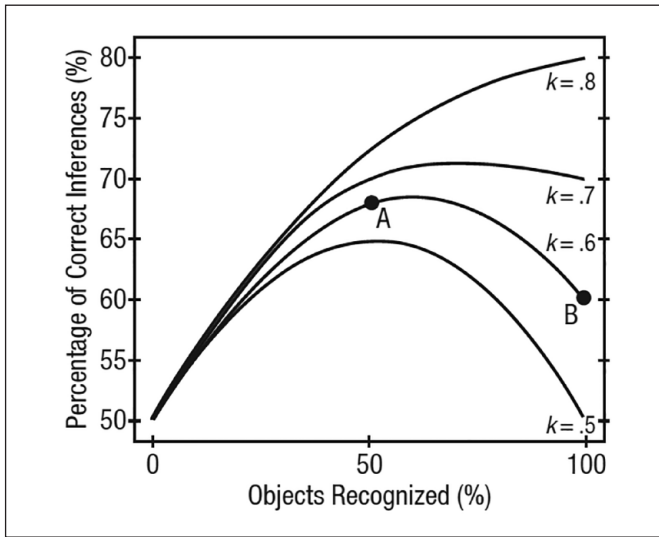
The null ritual contains all of these features: a fixation on the sacred 5% number (or on colors, as in functional MRI images), repetitive behavior akin to compulsive hand washing (“always perform this procedure”), fear of sanctions by editors and advisors, and wishful thinking and delusions about the meaning of the  $p$ -value. For instance, a review of studies with 839 academic psychologists in Chile, Germany, Italy, Netherlands, Spain, and the UK showed that the far majority had delusions about what a significant  $p$ -value means, such as that  $1 - p$  specifies the probability of a successful replication or the probability that the (unspecified) research hypothesis is true (Gigerenzer, 2018). These delusions are not accidental; they are necessary to uphold the ritual.

In defense of the null ritual, I have heard the argument that psychology is a premature science, which does not allow for making point predictions or predictions of functions, meaning that all we can do is test for null differences. That is a curious argument. For one, considering a field premature that has existed for far more than a century! And, as mentioned above, there are plenty of psychological theories that do make precise predictions. In psychophysics, exponential and logarithmic and other functional forms make competitive predictions; in developmental psychology, hypotheses of additive versus multiplicative information integration make testable point predictions; and in judgment and decision making, models of heuristics make testable point predictions. None of this requires a grand theory such as quantum theory. Even a theory of the processes people use to solve a knowledge quiz can lead to precise and surprising predictions. Here is an example from my own research group.

### *Less is more*

One of my colleagues made a puzzling observation. When he quizzed German students on the population of American cities (such as, which city has the larger population: Detroit or Milwaukee?) and German cities (Bielefeld or Hanover?), they unexpectedly scored more correct answers for American than German cities (see Hoffrage, 2011). The reason was not that they knew more about American cities; they knew less. Yet they relied on a smart heuristic that exploited their lack of knowledge—the *recognition heuristic*: If one of two objects is recognized and the other is not, then infer that the recognized object has the higher value with respect to the criterion.

For instance, many had not heard of Milwaukee, and concluded that Detroit must have the larger population, which is correct. But with Bielefeld and Hanover, they were unsure; they had heard of both. The next step is to analyze the *ecological rationality* of this psychological process, that is, the conditions under which the recognition heuristic performs well (Goldstein & Gigerenzer, 2002). These include: the proportion of objects (such as cities) a person recognizes; the recognition validity, measured by the



**Figure 1.** Recognition heuristic and less-is-more. Predictions of correct inferences as a function of objects recognized, for different levels of knowledge  $k$ . The recognition validity is  $.8$ . For the derivation of the curves, see Goldstein and Gigerenzer (2002).

performance in all tasks where exactly one of the two objects is recognized; and the knowledge validity  $k$ , measured by the performance in all tasks where both objects are recognized.

Figure 1 shows the prediction of correct inferences. Consider the lowest curve for a knowledge validity  $k=.5$ , which means that knowledge is at chance level. Here, performance first improves with an increasing number of objects recognized, from 50% to 64%, but then declines back to chance level when a person recognizes all objects (and can no longer use the recognition heuristic). The right side of the curve shows a *less-is-more effect*: recognizing fewer cities increases performance. Similarly, for  $k=.6$ , person A, who recognizes only half of the objects, is expected to score better than person B, who knows as much as A and also recognizes all objects. In general, less-is-more effects occur if the recognition validity ( $.8$  in Figure 1) is equal to or larger than the knowledge validity. If both are equal (top curve), no less-is-more effect occurs.

Less-is-more effects are counterintuitive. Figure 1 predicts when they occur and how large the effect size is. Less-is-more has been shown in contexts where these conditions hold, including the surprising finding that amateur tennis players (who recognized only half of the names of the 128 Gentleman Wimbledon players) could predict the outcomes of all 127 games as well as or better than the Association of Tennis Professionals rankings and the Wimbledon experts (Serwe & Frings, 2006; see Gigerenzer & Goldstein, 2011).

In general, an analysis of psychological processes such as the recognition heuristic makes it possible to derive precise predictions. The theory describes a causal mechanism that exploits semi-ignorance in order to make good choices.

### *The null ritual supports theory avoidance*

Let me return to the claim that psychology is a premature science and is ill equipped to make precise predictions. The recognition heuristic is a simple counter-example, and the adaptive toolbox of heuristics provides a general theoretical framework (Gigerenzer et al., 2011). In contrast, the earlier and related notion “availability” has never been clearly specified and continues to be used in multiple, unrelated senses, which hinders testable predictions.

Prematurity of psychological science is not the reason for the null ritual; rather, the null ritual promotes prematurity. One can publish papers with significant results that appear to prove a theory without even specifying it. The ritual cultivates what Kieran O’Doherty calls the “variable psychology,” where the researcher tests whether one variable has a significant impact on another one, and then uses meta-analyses to estimate the effect size (Personal communication, April 17, 2023). As Fisher (1955, 1956) noted many decades ago, this type of research has its place only when one knows little about a subject matter. Yet, constantly testing null hypotheses solidifies the state of knowing little.

A theory of a psychological process, such as the recognition heuristic, predicts the change of the size of effect under various conditions, which is an improvement over meta-analyses that try to measure an average effect size. In a meta-analysis without theory, a bigger effect is always better; if there is a theory of the psychological process, that is not necessarily the case. For instance, the prediction for person A is 68 % (Figure 1) and that for Person B 60 %, meaning that the less-is-more effect is eightpercentage points. Bigger effects are not always better once a theory predicts the size of the effect.

In summary, psychological research can learn two principles from the natural sciences: (a) derive precise predictions from a theory; the theory need not be universal, it can be small range, such as in the case of the recognition heuristic and (b) test these predictions, not a null: forget the null ritual.

### **How to avoid theory: As-if theories**

Milton Friedman, winner of the Nobel Memorial Prize for Economic Sciences, had a clear opinion about theories of psychological processes. He famously argued that psychological realism is irrelevant; what counts is solely the predictive power of a theory (M. Friedman, 1953). A theory that deliberately does not reflect reality is called an *as-if* theory. The core assumptions of an as-if theory are decided a priori by modeling preferences that are considered aesthetically pleasing or mathematically convenient, such as the optimization calculus. In this way, the theory of maximizing expected utility and its variants can be defended and maintained even in situations where its assumptions are obviously false. For instance, as-if theories assume that business firms behave as if they were fully informed about all possible options, their consequences, and their probabilities, and would calculate the option that maximizes profit. Here, the fact that omniscient firms and managers do not exist simply does not matter.

As-if theories are best known from astronomy, with the Ptolemaic model a case in point (Gingerich, 1973). Similar to in economics, Ptolemy’s as-if theory was not



arbitrary but rather committed to a priori assumptions, including that celestial bodies move only in circles and that they revolve around the earth at the center of the universe. Because both assumptions were false, Ptolemy had to introduce a host of extra circles, called epicycles, to improve the predictions of a false model. Eventually, Copernicus put the sun in the center while maintaining the doctrine of circles, and Kepler finally replaced the circles with ellipses (Gigerich, 1973). In this way, the as-if model was replaced by a process model of the actual movement of planets. Interestingly, the improved realism resulted in a much simpler theory.

In the natural sciences, moving from as-if to process theories is considered progress. That is not necessarily so in economics and psychology. For instance, when Reinhard Selten, a Nobel laureate in economics, submitted a paper with both an as-if utility maximization analysis and a psychological model of the actual decision processes to the *American Economic Review*, the editor asked him to discard the psychological part, which he refused to do, instead submitting the paper to a second-tier journal (R. Selten, personal communication, October 7, 2010). Surprisingly, quite a few psychologists became fascinated with expected utility maximization and followed, deliberately or unknowingly, M. Friedman's (1953) as-if doctrine. Examples are normative utilitarian theories of moral reasoning, but also descriptive theories of decision making such as prospect theory, or utility theories of consumer choice such as conjoint analysis. By maintaining the ideal of optimization (similar to the ideal of circles), these models have to make unrealistic assumptions about what humans can know about the future, and when doing so, avoid building theories about how people actually make decisions.

### *The free parameter game*

There is a striking difference between the Ptolemaic model and as-if models in psychology. The Ptolemaic model was complex with all its epicycles—necessarily so because it was wrong—but Ptolemy fixed its parameters, the epicycles. Similarly, Kepler's process model fixed its parameters, such as the parameters of the ellipses. Expected utility models in psychology, however, rarely fix their parameters and mostly do not commit to their values but fit these anew for each new set of data. I call this practice the "free-parameter game." By not committing to the parameter values, one always gains a better fit, but not necessarily better predictions.

To illustrate, in their review of some 50 years of empirical research, D. Friedman et al. (2014) analyzed how well utility functions—such as utility of income functions, utility of wealth functions, and the value function in prospect theory—actually predict behavior. They concluded: "Their power to predict out-of-sample is in the poor-to-nonexistent range, and we have seen no convincing victories over naïve alternatives" (p. 3).<sup>2</sup> The observation that theories with many free parameters achieve excellent fit but fail to predict is due to overfitting.

The practice of keeping parameters of a theory open and fitting them to each new data set is virtually unknown in the natural sciences. Physics builds theories with physical constants, rarely with free parameters. For instance, in  $e = mc^2$ , the speed of light  $c$  is a constant, not a parameter to be fitted to each data set anew in order to increase the fit.

Physical constants may be adjusted over years in the light of improved measurement, but that should not be confused with data fitting.

Thus, psychological research can learn from the natural sciences two principles: (a) replace as-if-models with process models and (b) fix parameters; don't fit them from experiment to experiment.

One surprising consequence of these two principles is that psychological theories typically become simpler, just as Kepler's theory is simpler than Ptolemy's. For instance, when one moves from cumulative prospect theory (an as-if theory with five free parameters that assumes complex calculations few people could perform) to the priority heuristic (a process theory with zero free parameters based on actual decision processes), much of the complexity can be avoided and the process becomes transparent (Brandstätter et al., 2006). A good process model can make better predictions of actual behavior than as-if models. For instance, the priority heuristic can predict difficult choices between monetary gambles (with similar expected value) better than cumulative prospect theory (Brandstätter et al., 2006).

In sum, the reliance on as-if-models of utility maximization, including their modifications, is a second technique to avoid theories of psychological processes. Despite their lack of realism, the free parameter game makes these theories appear to be successful by delivering a good fit for every new data set.

## How to avoid theory: Lists of binary oppositions

Physics started with binary oppositions and moved from there to precise laws. The Hippocratic theory of physics was based on the oppositions of hot versus cold and dry versus wet—today it is history (Lloyd, 1964). Much of psychology, however, remains obsessed with oppositions, such as nature versus nurture and intuition versus reason. Binary oppositions are tools to understand the world by creating dichotomies, where one pole is sometimes identified as the superior one. The opposition between male versus female is a case in point. Immanuel Kant was convinced that women's nature is sense and man's nature is reason. Similarly, the founder and first president of the American Psychological Association, G. Stanley Hall (1904/1976), held that women are intuitive and emotional, slow in logical thought, better at mental reproduction than production, and too impatient for analysis and science:

She works by intuition and feeling; fear, anger, pity, love, and most of the emotions have a wider range and greater intensity. If she abandons her natural naiveté and takes up the burden of guiding and accounting for her life by consciousness, she is likely to lose more than she gains, according to the old saw that she who deliberates is lost. (p. 561)

For centuries, the difference between the prototypical male and female was understood in terms of binary oppositions. Women's thinking was seen to be intuitive, fast, associative, inconsistent, concrete, and illogical. In contrast, men's thinking was seen to be rational, slow, reflective, consistent, abstract, and logical (Gigerenzer, 2023). Because women's supposed fast and intuitive thinking prevented them from grasping abstract moral principles and logical reasoning, they needed men's guidance to prevent them from making errors.

This very polarity has returned in the 21st century, now cleansed from its association with gender. Similar lists of binary oppositions are now presented as dual-system theories of thinking. System 1 is said to be fast and unconscious, associative and inconsistent, to work by intuition and heuristics, to lack rationality and to be the source of error. System 2, in contrast, is said to work by logic and statistics, to be slow and conscious, rule-based and consistent, rational and apparently always right.<sup>3</sup> These two systems resemble closely what once was believed to be the characteristics of women (System 1) and men (System 2). Moreover, human errors are blamed on the intuitive System 1, and the failure of the slow, rational System 2 to monitor it and correct what it gets wrong. As proponents of this view explain: “system 1 quickly proposes intuitive answers to judgment problems as they arise, and system 2 monitors the quality of these proposals, which it may endorse, correct, or override” (Kahneman & Frederick, 2005, p. 267). Just as men were once held responsible for preventing females from committing errors, the logical system is now assigned a similar paternalistic task.

I have no reason to assume that these similarities were by any means intentional. Given that two-system theories provide only a list of general dichotomies without specifying testable models of the underlying processes, they appear unfalsifiable. But, in fact, they do make two theoretical claims. First, intuition is opposed to deliberate reasoning and, moreover, considered inferior; second, the poles in each binary opposition are aligned with each other. Consider the first claim. If intuition and reasoning were opposites, they should be negatively correlated (either intuition or reason). However, a meta-analysis of 75 studies showed that measures of intuition and reasoning are not negatively correlated, but independent (Wang et al., 2017). Nor does research on expert intuition support the claim that intuition is opposed to deliberate reasoning. Instead, the majority of 17 Nobel laureates explained that their big leap had occurred by switching back and forth between intuition and reasoning (Dörfler & Eden, 2019); such a close interaction is also reported by chess experts and firefighters (Klein, 2015, 2017). There is also little evidence for the second testable claim, that the binary poles are aligned. On the contrary, every heuristic can be used consciously or unconsciously, and among experts, fast, unconscious decisions are often better than slow, deliberate reasoning (Beilock et al., 2004; Johnson & Raab, 2003).

In response to criticism (e.g., Gigerenzer & Regier, 1996; Kruglanski & Gigerenzer, 2011), Evans and Stanovich (2013) replaced their earlier terminology of two systems with two processes, called Type-1 and Type-2 process. That change avoided the unsupported implication of two brain systems. Yet binary oppositions are not processes.

Once again, psychological research can learn from the natural sciences two principles: (a) move from binary oppositions to theories of psychological processes and (b) beware of recycling value-laden dichotomies such as that between men and women.

In his brilliant 1973 paper “You Can’t Play 20 Questions with Nature and Win,” Allen Newell begins by acknowledging “I am a man who is half and half” (p. 283). One half of him was impressed by the beauty of the experiments he was asked to comment on, the other half depressed by the theoretical explanations put forward, mostly consisting of binary oppositions. Newell lists 24 of these, such as serial versus parallel processing and uniprocess versus duoprocess learning. By theorizing about experimental results in terms

of these general oppositions, Newell felt that “clarity is never achieved. Matters simply become muddier and muddier as we go down through time” (pp. 288–289).

## Can psychology learn from the natural sciences?

My answer is a clear “yes.” Psychologists can learn to walk forward on two feet—theory and experiment—rather than hobble on one. In this essay, I described three widespread techniques that maintain this hobbling: the null ritual, as-if-theories, and lists of binary oppositions. What worries me is not hobbling per se. It has been pointed out for decades, and by some of the most prominent psychologists. For instance, not only Paul Meehl explained that the logic of the null ritual runs contra to scientific method. B. F. Skinner (1972) was equally aware that psychologists have “taught statistics in lieu of scientific method” (p. 319). The mathematical psychologist R. Duncan Luce (1988) called null hypothesis testing “a wrongheaded view about what constituted scientific progress” (p. 582). The Skinnerians founded a new journal, the *Journal of the Experimental Analysis of Behavior*, in order to escape the ritual (Skinner, 1984, p. 138), and one of the reasons for launching the *Journal of Mathematical Psychology* was to escape the editors’ pressure to perform the mindless ritual (see Gigerenzer, 2004). What troubles me is that in spite of these warnings, the three techniques continue to be in fashion.

To learn from the natural sciences, researchers need to stop engaging in these three practices. That requires a co-ordinated effort of journal editors, academic societies, and grant agencies in the first place. A second measure is to teach graduate students how to construct theories about psychological processes with the help of instructive examples and also how to recognize surrogate theories. A third measure is to integrate theories, which requires sufficiently precise ones. Unlike Popper’s program of successively eliminating theories, the theory integration program promotes their integration. I have described this program in detail elsewhere (Gigerenzer, 2017).

Finally, why does much of psychology suffer from theory aversion? Earlier on I mentioned a form of risk aversion related to the fact that the more precisely an idea is formulated, the higher the risk that it can be refuted, even though that is exactly what science is about. A second reason is that most psychology departments do not teach the art of theory construction and integration. A third obvious reason is that thinking is hard whereas running an experiment is comparatively easy. A final reason is that quantity has become a surrogate for quality when committees decide on funding and promotion. The time and effort necessary for developing theories is not always valued when research quality is measured by the number of published articles multiplied by the impact factor of the journal. The European Research Council has taken measures against the increasing fixation on quantity by asking principal investigators to remove all impact factors from their grant applications, and submit only their 5 or 10 best articles, depending on seniority. This is an example that funding agencies can reduce the proliferation of *just publishable units*, and help refocus researchers’ minds on quality of theory and experiment.

## Funding

The author received no financial support for the research, authorship, and/or publication of this article.

## Notes

1. An exception is the *International Society for Theoretical Psychology*, home to *Theory & Psychology*. However, this society is only loosely connected to journals such as *Psychological Review*, the theoretical flagship of psychology.
2. Out-of-sample prediction means that the parameter values are fitted to one sample of data, and tests of the fitted model are performed on a new, independent sample from the same population. In contrast, when Milton Friedman speaks of prediction, he has in mind old-fashioned statistical hypothesis testing, which pertains to data fit, not real prediction (Minford, 2016, p. 563).
3. Dual-process theories come in many kinds, creating a scattered, dynamic, and blurred framework. The first versions included the idea that deliberate reasoning is typically used to rationalize an intuitive conclusion (see Evans & Stanovich, 2013) whereas Sloman (1996) did not emphasize that one system is superior to another. Here, I refer to dual system theories that assume (a) an opposition between the two systems, and (b) the superiority of System 2, as popularized by Kahneman (2011; Kahneman & Frederick, 2005) and Thaler and Sunstein (2008). For a history, see Mercier and Sperber (2018).

## References

- Beilock, S. L., Bertenthal, B. I., McCoy, A. M., & Carr, T. H. (2004). Haste does not always make waste: Expertise, direction of attention, and speed versus accuracy in performing sensorimotor skills. *Psychonomic Bulletin & Review*, *11*, 373–379. <https://doi.org/10.3758/BF03196585>
- Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, *113*(2), 409–432. <https://doi.org/10.1037/0033-295X.113.2.409>
- Busemeyer, J. R., Wang, Z., & Shiffrin, R. M. (2015). Bayesian model comparison favors quantum over standard decision theory account of dynamic inconsistency. *Decision*, *2*(1), 1–12. <https://doi.org/10.1037/dec0000017>
- Dörfler, V., & Eden, C. (2019). Understanding “expert” scientists: Implications for management and organizational research. *Management Learning*, *50*(5), 534–555. <https://doi.org/10.1177/1350507619866652>
- Dulancy, S., & Fiske, A. P. (1994). Cultural rituals and obsessive-compulsive disorder: Is there a common psychological mechanism? *Ethos*, *22*(3), 243–283. <https://doi.org/10.1525/eth.1994.22.3.02a00010>
- Evans, J., St. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223–241. <https://doi.org/10.1177/1745691612460685>
- Feynman, R. (1967). *The character of physical law*. MIT Press.
- Fisher, R. A. (1955). Statistical methods and scientific induction. *Journal of the Royal Statistical Society: Series B (Methodological)*, *17*(1), 69–77. <https://doi.org/10.1111/j.2517-6161.1955.tb00180.x>
- Fisher, R. A. (1956). *Statistical methods and scientific inference*. Oliver & Boyd.
- Friedman, D., Isaac, R. M., James, D., & Sunder, S. (2014). *Risky curves. On the empirical failure of expected utility*. Routledge.

- Friedman, M. (1953). *Essays in positive economics*. University of Chicago Press.
- Gigerenzer, G. (1993). The superego, the ego, and the id in statistical reasoning. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 311–339). Erlbaum.
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review*, 103(3), 592–596. <https://doi.org/10.1037/0033-295X.103.3.592>
- Gigerenzer, G. (2004). Mindless statistics. *Journal of Socio-Economics*, 33, 587–606. <https://doi.org/10.1016/j.socec.2004.09.033>
- Gigerenzer, G. (2009). Surrogates for theory. *Association for Psychological Science Observer*, 22(2), 21–23. <https://www.psychologicalscience.org/observer/surrogates-for-theory>
- Gigerenzer, G. (2010). Personal reflections on theory and psychology. *Theory & Psychology*, 20(6), 733–743. <https://doi.org/10.1177/0959354310378184>
- Gigerenzer, G. (2017). A theory integration program. *Decision*, 4(3), 133–145. <https://doi.org/10.1037/dec0000082>
- Gigerenzer, G. (2018). Statistical rituals: The replication delusion and how we got there. *Advances in Methods and Practices in Psychological Science*, 1(2), 198–218. <https://doi.org/10.1177/2515245918771329>
- Gigerenzer, G. (2022). *How to stay smart in a smart world*. Penguin.
- Gigerenzer, G. (2023). *The intelligence of intuition*. Cambridge University Press.
- Gigerenzer, G., & Goldstein, D. G. (2011). The recognition heuristic: A decade of research. *Judgment and Decision Making*, 6(1), 100–121. <https://doi.org/10.1017/S1930297500002126>
- Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). *Heuristics: The foundations of adaptive behavior*. Oxford University Press.
- Gigerenzer, G., & Regier, T. (1996). How do we tell an association from a rule? Comment on Slovic (1996). *Psychological Bulletin*, 119(1), 23–26. <https://doi.org/10.1037/0033-2909.119.1.23>
- Gigerenzer, G., Switjink, Z., Porter, T., Daston, L., Beatty, J., & Krüger, L. (1989). *The empire of chance: How probability changed science and everyday life*. Cambridge University Press.
- Gingerich, O. (1973). From Copernicus to Kepler: Heliocentrism as model and as reality. *Proceedings of the American Philosophical Society*, 117(6), 513–522. <https://www.jstor.org/stable/986462>
- Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109(1), 75–90. <https://doi.org/10.1037/0033-295X.109.1.75>
- Hall, G. S. (1976). Biological and anthropological differences between the sexes. In P. C. Lee & R. S. Stewart (Eds.), *Sex differences: Cultural and developmental dimensions* (pp. 371–379). Urizen Books. (Original work published 1904)
- Heidelberger, M. (1987). Fechner's indeterminism: From freedom to laws of chance. In L. Krüger, L. J. Daston, & M. Heidelberger (Eds.), *The probabilistic revolution: Vol. 1. Ideas in history* (pp. 117–156). MIT Press.
- Hoffrage, U. (2011). Recognition judgments and the performance of the recognition heuristic depend on the size of the reference class. *Judgment and Decision Making*, 6(1), 43–57. <https://doi.org/10.1017/S1930297500002084>
- Johnson, J. G., & Raab, M. (2003). Take the first: Option-generation and resulting choices. *Organizational Behavior and Human Decision Processes*, 91(2), 215–229. [https://doi.org/10.1016/S0749-5978\(03\)00027-X](https://doi.org/10.1016/S0749-5978(03)00027-X)
- Kahneman, D. (2011). *Thinking, fast and slow*. Allen Lane.
- Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 267–293). Cambridge University Press.

- Katzo, M. W. (2006). A study of the logic of empirical arguments in psychological research: “The automaticity of social behavior” as a case study. *Review of General Psychology, 10*(3), 210–228. <https://doi.org/10.1037/1089-2680.10.3.210>
- Klein, G. (2015). A naturalistic decision making perspective on studying intuitive decision making. *Journal of Applied Research in Memory and Cognition, 4*(3), 164–168. <https://doi.org/10.1016/j.jarmac.2015.07.001>
- Klein, G. (2017). *Sources of power: How people make decisions* (20th anniversary ed.). MIT Press.
- Kruglanski, A. W. (2001). That “vision thing”: The state of theory in social and personality psychology at the edge of the new millennium. *Journal of Personality and Social Psychology, 80*(6), 871–875. <https://doi.org/10.1037/0022-3514.80.6.871>
- Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review, 118*(1), 97–109. <https://doi.org/10.1037/a0020762>
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago University Press.
- Lloyd, G. E. R. (1964). The hot and the cold, the dry and the wet in Greek philosophy. *The Journal of Hellenic Studies, 84*, 92–106. <https://doi.org/10.2307/627697>
- Luce, R. D. (1988). The tools-to-theory hypothesis [Review of the book *Cognition as intuitive statistics*, by G. Gigerenzer, & D. J. Murray]. *Contemporary Psychology, 33*(7), 582–583. <https://doi.org/10.1037/030460>
- Mach, E. (2013). *The science of mechanics* (T. J. McCormack, Trans.). Cambridge University Press. (Original work published 1883)
- Meehl, P. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science, 34*(2), 103–115. <https://doi.org/10.1086/288135>
- Mercier, H., & Sperber, D. (2018). *The enigma of reason*. Penguin Books.
- Millikan, R. A. (1924, May 23). The electron and the light-quant from the experimental point of view. *The Nobel Prize*. <https://www.nobelprize.org/prizes/physics/1923/millikan/lecture/>
- Minford, P. (2016). Milton Friedman’s methodology, macroeconomics, and the great recession. In R. A. Cord & J. D. Hammond (Eds.), *Milton Friedman: Contributions to economics and public policy* (pp. 541–560). Oxford University Press.
- Newell, A. (1973). You can’t play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing: Proceedings of the eighth annual Carnegie symposium on cognition* (pp. 283–308). Academic Press.
- Porter, T. (1986). *The rise of statistical reasoning, 1820–1900*. Princeton University Press.
- Serwe, S., & Frings, C. (2006). Who will win Wimbledon? The recognition heuristic in predicting sports events. *Journal of Behavioral Decision Making, 19*(4), 321–322. <https://doi.org/10.1002/bdm.530>
- Shepard, R. N. (2004). How a cognitive psychologist came to seek universal laws. *Psychonomic Bulletin & Review, 11*, 1–23. <https://doi.org/10.3758/BF03206455>
- Simon, H. (1979). *Models of thought*. Yale University Press.
- Skinner, B. F. (1972). *Cumulative record*. Appleton-Century-Crofts.
- Skinner, B. F. (1984). *A matter of consequences*. New York University Press.
- Slooman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*(1), 3–22. <https://doi.org/10.1037/0033-2909.119.1.3>
- Staddon, J. (2021). *The new behaviorism: Foundations of behavioral science* (3rd ed.). Routledge.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.
- The, J. T., Magistretti, P., & Ansermet, F. (2018). The epistemological foundations of Freud’s energetics model. *Frontiers in Psychology, 9*, Article 1861. <https://doi.org/10.3389/fpsyg.2018.01861>

- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84(4), 327–352. <https://doi.org/10.1037/0033-295X.84.4.327>
- Wallach, L., & Wallach, M. A. (1994). Gergen versus the mainstream: Are hypotheses in social psychology subject to empirical test? *Journal of Personality and Social Psychology*, 67(2), 233–242. <https://doi.org/10.1037/0022-3514.67.2.233>
- Wang, Y., Highhouse, S., Lake, C. J., Petersen, N. L., & Rada, T. B. (2017). Meta-analytic investigations of the relation between intuition and analysis. *Journal of Behavioral Decision Making*, 30(1), 15–25. <https://doi.org/10.1002/bdm.1903>

### Author biography

Gerd Gigerenzer is director emeritus of the Max Planck Institute for Human Development, Member of the American Academy of Arts and Sciences, the American Philosophical Society, and the German Academy of Sciences. His many awards include the AAAS Prize for Behavioral Science Research, and his books have been translated into more than 20 languages. He has been distinguished as one of the top-100 Global Thought Leaders worldwide. Recent publications include: *The Intelligence of Intuition* (Cambridge University Press, 2023) and *How to Stay Smart in a Smart World: Why Human Intelligence Still Beats Algorithms* (MIT Press, 2022).