

Good judgments do not require complex cognition

Julian N. Marewski · Wolfgang Gaissmaier ·
Gerd Gigerenzer

Received: 15 November 2008 / Accepted: 9 September 2009 / Published online: 27 September 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract What cognitive capabilities allow *Homo sapiens* to successfully bet on the stock market, to catch balls in baseball games, to accurately predict the outcomes of political elections, or to correctly decide whether a patient needs to be allocated to the coronary care unit? It is a widespread belief in psychology and beyond that complex judgment tasks require complex solutions. Countering this common intuition, in this article, we argue that in an uncertain world actually the opposite is true: Humans do not need complex cognitive strategies to make good inferences, estimations, and other judgments; rather, it is the very simplicity and robustness of our cognitive repertoire that makes *Homo sapiens* a capable decision maker.

Keywords Heuristics · Ecological rationality · Cognitive capacities

Introduction

Imagine a baseball player in an extremely important game who wants to catch a ball, which is hit high-up in the air in

his direction. How could the player master this rather complex task? One approach would entail solving “a set of differential equations [to predict] the trajectory of the ball” (Dawkins 1989, p. 96). Note that these differential equations need to incorporate all relevant parameters, such as initial speed of the ball, wind, spin, and so forth. Then, the player could calculate exactly where the ball is most likely to come down, and run directly there. But would anyone really solve the task like this? Some believe so, and they would probably subscribe to Dawkins’s (p. 96) speculation that “[at] some subconscious level, something functionally equivalent to the mathematical calculations is going on.”

A very different approach to catching the ball involves using a *heuristic*. A heuristic is a simple strategy that ignores information.¹ For catching a ball that is high-up in the air, players could use the *gaze heuristic*, which is easier than solving a set of differential equations. This heuristic is one from a set that has been observed in experienced players. It works if the ball is already high-up in the air. The heuristic does not allow a player to predict where the ball will come down, but it will help him to be exactly there. To catch the ball, one simply has to fixate it, start running, and adjust the speed of running such that the angle

Julian N. Marewski and Wolfgang Gaissmaier contributed equally to this article; the order of authorship was determined by a coin flip.

J. N. Marewski (✉) · G. Gigerenzer
Center for Adaptive Behavior and Cognition,
Max Planck Institute for Human Development,
Lentzeallee 94, 14195 Berlin, Germany
e-mail: marewski@mpib-berlin.mpg.de

W. Gaissmaier (✉)
Harding Center for Risk Literacy,
Max Planck Institute for Human Development,
Lentzeallee 94, 14195 Berlin, Germany
e-mail: gaissmaier@mpib-berlin.mpg.de

¹ The term “heuristic” originates from the Greek word “heuriskein”, which means “to find” or “to discover”. In the problem solving literature and other areas of psychology (cf. Groner et al. 1983), it is employed to refer to methods for finding solutions, rather than to simple decision strategies, which is how we use the term. However, somewhat similar to the classic meaning of the term, the computational models of heuristics that we discuss have explicit rules for search that, for instance, prescribe in which order information should be looked up and considered when making decisions. Our usage of the term heuristics also resembles what biologists call *rules of thumb* (Hutchinson and Gigerenzer 2005) and Simon (1955) the *satisficing principle*, which attempts to find solutions that are good enough for achieving a given goal.

of gaze remains constant. By using this heuristic, a player can ignore all the aforementioned parameters involved in the computation of the trajectory (Fig. 1). Variants of this heuristic can also be observed in dogs catching Frisbees (Shaffer et al. 2004). And even in times of highly advanced technology, airplane pilots can rely on it to avoid collisions with other aircrafts: When another vehicle is flying toward them on a potential collision course, they simply need to fixate a small scratch on their windshield and observe whether the approaching aircraft is moving away from this scratch. If the other aircraft holds on the scratch, then a collision is imminent, and the pilots better change their course; if the other aircraft moves away from the scratch, then the current course can be continued (Fig. 2).

The gaze heuristic is only one illustration of the major point we will make in this paper: Complex judgment tasks often do not need complex cognitive strategies to be solved successfully. Quite to the contrary, there are even situations in which complex cognition may hurt performance compared to simpler cognition. In what follows, we will review

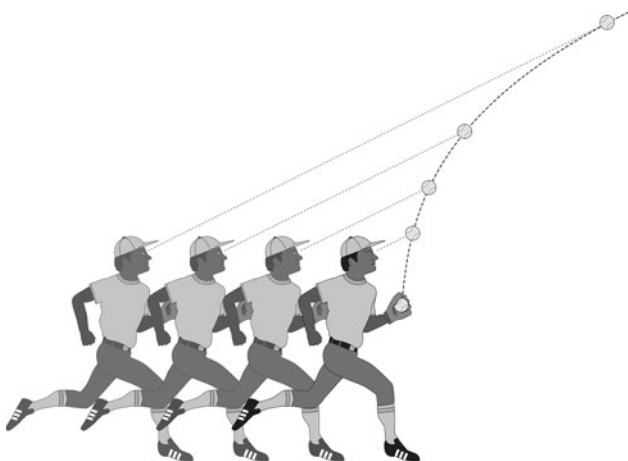


Fig. 1 Illustration of the gaze heuristic for catching baseballs. Figure adapted from Gigerenzer (2007)

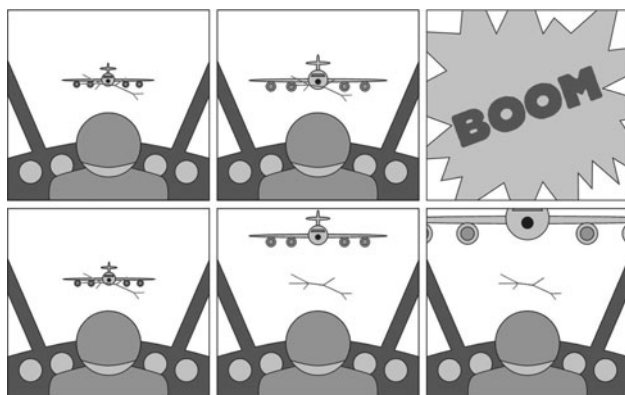


Fig. 2 Illustration of a variant of the gaze heuristic that can help pilots avoid collisions

research of the *fast and frugal heuristics framework* (e.g., Gigerenzer et al. 1999), which is an approach to judgment and decision making that focuses on spelling out when, how, and why simple cognitive strategies can help people make good judgments and decisions. In doing so, we are mostly concerned with how people make inferences, estimations, and other judgments about unknown or uncertain criteria, such as about tomorrow's weather or the prices of stocks. First, setting the historical preliminaries, we will give an overview of two competing visions of the rationale of human cognition, each of which implies a different approach to catching balls in baseball games. Second, we will explain when simple strategies are superior to more complex ones. Third, we will present a selection of very simple cognitive heuristics, explaining when and how they outperform more complex ones in solving real-world problems. Fourth, we will give a series of examples showing that the limitations of our *core capacities* such as vision or memory are not accidental but may actually be beneficial.

Visions of rationality

What cognitive capabilities allow *Homo sapiens* to be able to both catch baseballs and avoid collisions with other airplanes? The answer to this question depends on one's view of human rationality, because this view determines what kind of models of cognition one believes constitute humans' cognitive machinery. There are three major approaches.

Unbounded rationality = rational behavior?

Unbounded rationality assumes that a person knows all the relevant information (all different alternatives, their consequences, and the probabilities of their consequences), has unlimited time, unfailing memory and is endowed with large computational power (i.e., information-processing capacity), needed to run complex calculations and compute mathematically optimal solutions. The maximization of subjective expected utility is one example (e.g., Edwards 1954); the laws of logic constitute another. Models of unbounded rationality are common in economics, optimal foraging theory, and cognitive science. The idea of catching a ball by solving complex equations would also fit in this vision. From this perspective, the human being should ideally be omniscient: the more information, and the more processing capacity, the better. Omniscience and optimization go hand in hand with a third ideal: universality (as opposed to modularity). Universality is best exemplified in Leibniz's (1677/1951) dream of a universal calculus that can solve all problems.

Bounded rationality = unfortunate but unavoidable, irrational behavior?

As we have illustrated earlier, simple rules of thumb such as the gaze heuristic represent an alternative view. This view acknowledges that unbounded rationality may be a convenient modeling assumption, but it is an unrealistic description of how people make decisions. Our resources—time, knowledge, and computational power—are limited. Simon (1956, 1990), the father of this *bounded rationality* view, argued that people rely on simple strategies to deal with situations of sparse resources. Many researchers used Simon's ideas (or misused them, in his view) to argue that bounded rationality is the study of cognitive fallacies, maintaining that rational behavior should be still defined in terms of the beautiful ideals of universal, optimal, and logic solutions. Similarly, the premise of limited cognitive capacities has often been directly linked to its supposed negative consequences, such as reasoning errors or poor cognitive performance (e.g., Johnson-Laird 1983). From this perspective, cognitive (and other) limitations force people to abandon what would be optimal decision strategies. Instead, people need to rely on shortcuts or on heuristics, which, in a pessimistic appraisal of human cognition, make people vulnerable to systematic and predictable reasoning errors. This view is most prominently represented by the *heuristics-and-biases program* (e.g., Kahneman et al. 1982; Tversky and Kahneman 1974). According to this framework, behavior deviating from the laws of logic or the maximization of subjective expected utility can be explicated by assuming that people's heuristics are error prone and subject to systematic cognitive biases. Conversely, people's use of heuristics explains why decisions can be suboptimal, irrational, or illogical when compared to the normative yardstick of unbounded rationality.

The adaptive toolbox of heuristics

However, Simon (e.g., 1990) not only stressed the cognitive limitations of humans and proposed simple strategies people may rely on, but also emphasized that behavior is a function of both cognition and the environment: "Human rational behavior...is shaped by a scissors whose two blades are the structure of task environments and the computational capabilities of the actor" (1990, p. 7). The fast and frugal heuristics research program (e.g., Gigerenzer et al. 1999) has taken up this emphasis. It assumes that rationality is not only bounded, but also ecological. A heuristic is *ecologically rational* to the degree it fits to the structure of the environment. Researchers following this perspective believe that the beautiful vision of a universal calculus is a mere dream (although their hearts would be

overjoyed if someone were to finally show them the calculus). Instead of searching for the universal tool that can solve all tasks—simple and complex ones alike—they take humans to possess a repertoire of specialized heuristics that can solve specific tasks in specific environments. Gigerenzer et al. called this collection of cognitive strategies and the core capacities they exploit the *adaptive toolbox*. The toolbox contains heuristics that allow people to make inferences (e.g., about movie quality), develop preferences (e.g., for brands), plan interactions with others (e.g., decide which negotiation style to adopt when discussing), or make other judgments and decisions in social and nonsocial contexts. By drawing on our core capacities such as memory or vision and by exploiting regularities in the structure of the human physical and social environment, the heuristics in the toolbox can yield accurate decisions in the face of limited time, knowledge, and computational power.

For instance, as we will discuss in more detail later, the inability of human memory to store large amounts of records facilitates remembering important, up-to-date information, sorting out irrelevant and outdated information by forgetting it. Memory does this by updating its records as a function of their environmental occurrence: The more often a piece of information is encountered in the environment, and the less time has elapsed since it has last occurred, the more likely it is that a person will retrieve a memory of that piece of information, and the more likely it is that this piece of information is actually the most relevant for the system's current processing goals (Anderson and Schooler 1991). Library search engines, word processors, and other machines work in a similar way, pulling up documents that have been very frequently or very recently used, essentially betting that these documents will most likely be the ones the user is looking for.

The human vision system with its ability to track objects against a noisy background, such as when a baseball is up in the air, is another example of a core capacity. It allows people to rely almost effortlessly on the gaze heuristic, something computers cannot do as well as humans yet. Thanks to the vision system, the gaze heuristic can essentially operate on light as environmental information, and in adjusting the speed of running while keeping the angle to the ball constant, it can transform complex trajectories into linear ones.

In short, the success of the human cognitive machinery is anchored in three aspects: The environment, the core capacities of the human mind, and the way highly specialized heuristics exploit both environmental structure and core capacities (e.g., Gaissmaier et al. 2008; Marewski and Schooler 2009).

Importantly, none of the heuristics in the adaptive toolbox are all-purpose tools that can and should be applied invariantly in all kinds of situations. Nor is any other

strategy—contrary to the misleading opposition between the fast and frugal heuristics framework and allegedly single-strategy models such as put forward recently by Newell (2005) and Glöckner et al. (2009; see Marewski 2009). Rather, each heuristic is tuned to specific environmental regularities and designed for specific decision problems. Just as a screwdriver is of little use to hammer a nail into a wall but works well to attach a screw, each heuristic is a specialized tool. From this perspective, the right question to ask thus is not whether a cognitive strategy is universally successful with respect to unboundedly rational yardsticks, such as the maximization of subjective expected utility or traditional logic. Rather, if one assumes that organisms behave ecologically rationally, what really matters is to identify the tasks that a heuristic can solve, and in doing so, to find out when this heuristic can help an organism to reach a goal, for instance, by enabling it to make accurate and fast inferences. Acting fast and predicting accurately illustrate benchmarks organisms must live up to in order to survive.² As we will demonstrate next, heuristics need to be simple in order to meet these benchmarks, particularly in a world that is so fundamentally uncertain that, as Franklin (1987) pointed out in 1789 at the dawn of the French Revolution, “there is nothing certain but death and taxes.”

Why can simple strategies outperform complex ones?

To illustrate the following points, we would like to invite you to consider a thought experiment. Imagine there are two species, operating with different cognitive systems. The species *simplissimus* is boundedly rational, and its cognition is based on a repertoire of rules of thumb. The species *complexicus*, in turn, can rely on complex, highly sophisticated strategies. To survive, both species need to make accurate predictions about future events and unknown quantities, such as where in the vicinity food can be found, or which sleeping sites are most likely to offer protection from predators.

How could one assess which species comes equipped with the better cognitive machinery for making such predictions? From a methodologist’s point of view, this is akin to asking the question which of two competing *models*, or formal theories, provides a better account of data. The two cognitive machineries represent the two models, and the future events and unknown quantities are the data the machineries need to forecast in order to ensure the species’ survival. For instance, think of two of the species’

cognitive strategies as two regression models. One of *complexicus*’s strategies might read: $y = w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + w_5x_5$. A simpler strategy of *simplissimus* could throw away both the regression weights, w_i , and some pieces of information (and, in doing so, eliminate a bunch of free parameters) and might look like this: $y = x_1 + x_2$. The criterion to be predicted could be the amount of food, y , likely to be available in a certain area, and the predictor variables, x_i , the characteristics of the area, such as texture of the soil, or the type and size of nearby plants. In methodology, a standard way of answering which of these two models is better would entail computing R^2 or other standard goodness-of-fit indices. Such measures are based on the distance between each model’s estimate and the criterion y . And indeed, paying attention to more variables (x_3, x_4, x_5) and weighting them in an optimal way (i.e., minimizing least squares) can never lead to a smaller R^2 than the simpler strategy. It seems that *simplissimus* can never predict food better than *complexicus*.

Yet there is a problem with this approach. Goodness-of-fit measures alone cannot disentangle the variation in data due to noise from the variation due to the variables of interest. As a result, a model can end up *overfitting* the data, that is, it can capture not only the variance due to the variables of interest but also that from random error, which organisms are likely to encounter in an uncertain world.

The problem of overfitting

Figure 3 illustrates a situation in which one model, call it Model A (thin line), overfits existing data by chasing after idiosyncrasies in that data. This model fits the existing, already observed data (filled squares) perfectly but does a relatively poor job of predicting new, thus far unseen data (filled circles). Model B (thick line), while not fitting the existing data as well as Model A, captures the main tendencies in that data and ignores the idiosyncrasies. This makes it better equipped to predict new observations, as can be seen from the deviations between the model’s predictions and the new data, which are indeed smaller than the deviations for Model A. Formally, Model A overfits data if there is an alternative Model B and A performs better than B in fitting existing data (e.g., in a learning sample) but worse in predicting new data (e.g., in a test sample). In this case, B is called the more *robust* model (Gigerenzer 2004a).

The ability of a model to predict new data is called its *generalizability*, that is, the degree to which it is capable of predicting *all* potential samples generated by the same process, rather than to fit only a *particular* sample of existing data. The degree to which a model is susceptible to overfitting, in turn, is related to the model’s *complexity*, which, following Pitt et al. (2002), refers to a model’s

² Hammond (1996) called outside criteria for rationality such as speed and accuracy of prediction *correspondence criteria* as opposed to *coherence criteria*, which take the laws of logic or probability theory as normative yardstick for rationality.

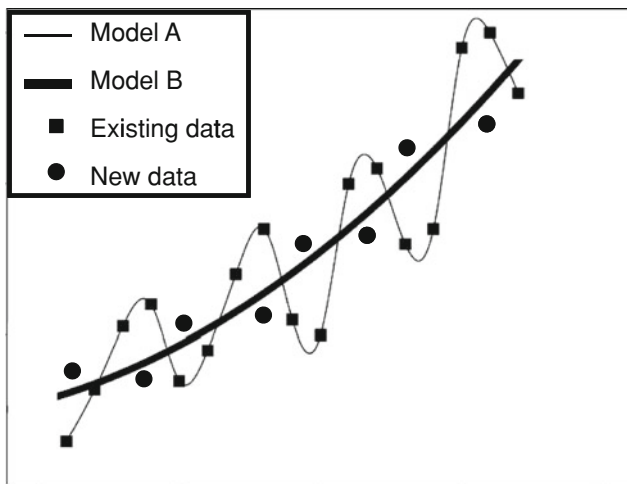


Fig. 3 Illustration of how two models fit existing data (*squares*) and how they predict new data (*filled circles*; see Pitt et al. 2002). Model A (*thin line*) overfits the data and is not as accurate in predicting new data as Model B (*thick line*)

inherent flexibility that enables it to fit diverse patterns of data.

Among the factors that contribute to a model's complexity are (1) the number of free parameters it has and (2) how the parameters are combined in it—in other words, its *functional form*. The impact of many free parameters is illustrated in Fig. 3, where the highly flexible Model A that overfits the data has more free parameters than the less flexible Model B that captures the main tendencies in the data. The impact of the functional form on the flexibility of a model's prediction can be illustrated by comparing Fechner's (1860/1966) and Stevens's (1957) famous models of the relation between psychological dimensions (e.g., brightness, called y here) and their physical counterparts (e.g., the intensity of light, called x here). In both models, there are two free parameters, a and b , but they are combined differently (Stevens's model: $y = ax^b$; Fechner's model: $y = a \ln[x + b]$). As noted by Townsend (1975), Stevens's model is more complex than Fechner's model. Since it assumes that a power function relates the psychological and physical dimensions, Stevens's model can fit data that have negative, positive, and zero curvature. Fechner's model, in turn, can only fit data with a negative curvature because it assumes a logarithmic relationship.

The dilemma can be summarized in the following way. When data is not completely free of random error, increased complexity makes a model more likely to end up overfitting data while its generalizability to new data decreases. At the same time, decreasing a model's complexity can eventually lead to underfitting; thus, in an uncertain world, there is often an inversely U-shaped function between complexity and predictive power (see Pitt

et al. 2002). In short, a good fit to existing data, achieved by a high model complexity, does not necessarily imply good generalizability to new data. As we will argue in the remainder of this article, there is, in fact, growing evidence that human repertoire of cognitive strategies has evolved to be simple—to the right degree of robustness to cope with the uncertainty of the world. Next, we will demonstrate that there are a large number of tasks in which simpler heuristics outperform more complex cognitive mechanisms in making accurate forecasts of the future, regardless of whether it comes to predicting rainfall, the performance of stocks, or the outcomes of sports events.

A case study of four heuristics

Research in the fast and frugal heuristics program focuses on three interrelated questions (see Gigerenzer et al. 2008). The first is descriptive and concerns the adaptive toolbox: What heuristics do organisms use to make decisions and when do people rely on which heuristic from the toolbox? The second question is prescriptive and deals with ecological rationality: To what environmental structures is a given heuristic adapted—that is, in what situations does it perform well, say, by being able to yield accurate, fast, and effortless decisions? In contrast to these two questions, the third one focuses on practical applications: How can the study of people's repertoires of heuristics and their fit to environmental structure aid decision making in the applied world?

Ecologically rational heuristics have been studied in diverse areas. These include applied ones such as medicine (Wegwarth et al. 2009), where heuristics can improve coronary care unit allocations (Green and Mehr 1997) and aid first-line antibiotic prescriptions in children (Fischer et al. 2002), as well as risk communication for lawyers, patients, and doctors (Gigerenzer 2002; Gigerenzer et al. 2007; Hoffrage et al. 2000), and the library sciences (Cokely et al. 2009; Marewski et al. 2009b). At the same time, the fast and frugal heuristics approach is discussed in several fields, including philosophy (e.g., Bishop 2006), economics (Gigerenzer and Selten 2001), the law (e.g., Gigerenzer and Engel 2006), and biology (e.g., Hutchinson and Gigerenzer 2005). In particular, this program has proposed and tested a range of heuristics for different tasks—mate search (Todd and Miller 1999), parental investment (Davis and Todd 1999), inferential judgments (e.g., Gigerenzer and Goldstein 1996; Goldstein and Gigerenzer 2002), estimation (Hertwig et al. 1999; von Helversen and Rieskamp 2008), categorization (Berretty et al. 1999), moral judgment (Coenen and Marewski 2009), and choices between risky alternatives (Brandstätter et al. 2006), to name a few. Moreover, the program has produced

a large amount of research investigating whether and when people, both young and old, rely on given heuristics (Bröder and Schiffer 2003; Cokely and Kelley 2009; Mata et al. 2007; Pachur et al. 2008; Pachur and Hertwig 2006; Pohl 2006; Rieskamp and Hoffrage 1999, 2008; Rieskamp and Otto 2006), under what environmental structures the heuristics perform well (e.g., Gigerenzer and Goldstein 1996; Hogarth and Karelaia 2007; Katsikopoulos and Martignon 2006a, b; Martignon and Hoffrage 1999), and how accurate they are for predicting events in the real, uncertain world such as the performance of stocks on the stock market (Ortmann et al. 2008), the outcomes of sports events (e.g., Pachur and Biele 2007; Scheibehenne and Bröder 2007; Serwe and Frings 2006), or how much time various mammals spend sleeping (Czerlinski et al. 1999). In what follows, we will discuss how the three questions of the fast and frugal heuristics program—the descriptive, the prescriptive, and the question about applications—have been answered for four simple heuristics.³ These are the *recognition heuristic*, the *fluency heuristic*, the *take-the-best heuristic*, and *tallying*. Each of these heuristics consists of three simple building blocks, one rule for searching information, one rule for stopping this search, and one for making the decisions.

Recognition heuristic

Which car brand is of better quality, a German-engineered BMW or a KIA? A person who has heard of the carmaker BMW before reading this article, but has never heard of KIA Motors, a fairly large Asian car company could use the recognition heuristic (Goldstein and Gigerenzer 1999, 2002) to respond: This heuristic would bet on BMW, which is the recognized car brand.

In its simplest form, the recognition heuristic is designed for inferring which of two alternatives, one recognized and the other not, has a larger value on a quantitative criterion. It simply searches for recognition information and stops information search once an alternative is judged as recognized. When recognition correlates strongly with the criterion on which alternatives are evaluated, the heuristic is ecologically rational and can be defined as follows.

Search rule: In a comparison of two alternatives, search in memory which alternative is recognized and which is not.

Stopping rule: Stop once both alternatives are classified as recognized or unrecognized.

Decision rule: If one alternative is recognized but not the other, infer the recognized alternative to have a larger value on the criterion.

Even more so than in the case of two alternatives, recognition is particularly useful when winnowing down many alternatives, for instance, when ranking them. Many theories of choice assume a two-stage process: When evaluating multiple alternatives, first a smaller set of relevant alternatives is formed, and then a choice is made after more detailed examination of the alternatives in this *consideration set* (e.g., Alba and Chattopadhyay 1985; Hauser and Wernerfelt 1990; Howard and Sheth 1969). When recognition correlates strongly with the criterion on which alternatives are evaluated, the recognition heuristic is useful to generate “consideration sets” consisting of recognized alternatives (Marewski et al. 2009a):

Search rule: If there are N alternatives, search in memory which n alternatives are recognized and which $N-n$ alternatives are not recognized.

Stopping rule: Stop once all alternatives are classified as recognized or unrecognized.

Decision rule: Infer that the n recognized alternatives rank higher on the criterion than the $N-n$ unrecognized ones.

Consideration sets facilitate decisions by reducing the number of alternatives. To illustrate, imagine consumers wanted to rank eight car companies according to the quality of their cars. If they considered all the companies, they would face a total of $8!$ (40,320) possible rank orders. In contrast, if the recognition heuristic is used, and, say, four companies are unrecognized and four recognized, then there are only $4!$ (24) possible rank orders, namely, those of the recognized companies that constitute the consideration set of top ones. In a second stage, the final rank order of these companies can be determined with heuristics that use other information, such as knowledge about whether the company is operating in many countries. The four unrecognized alternatives can be put aside because they are likely to score low on the criterion.

When is it ecologically rational to rely on the recognition heuristic?

The recognition heuristic is a specialized tool: It is only applicable when at least one alternative is recognized while others are unrecognized. Using the heuristic will result in accurate decisions in environments in which the probability of recognizing alternatives is correlated with the criterion to be inferred. This is, for example, the case in many geographical domains such as city or mountain size (Goldstein and Gigerenzer 2002), and in many competitive

³ For a discussion of principles that provide guidance for how to study people’s use of heuristics and the adaptive toolbox, see Gigerenzer and Brighton (2009), Marewski et al. (2009c), and Marewski and Olsson (2009).

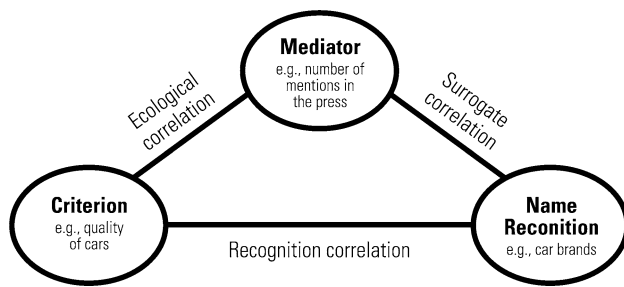


Fig. 4 How does the recognition heuristic work? An unknown criterion (e.g., the quality of car brands) is reflected by a mediator (e.g., the press). The mediator makes it more likely for a person to encounter alternatives with larger criterion values than those with smaller ones (e.g., the press mentions quality brands more frequently). As a result, the person will be more likely to recognize alternatives with larger criterion values than those with smaller ones, and ultimately, recognition can be relied upon to infer the criterion (e.g., to infer which of two cars is of a higher quality). The relations between the criterion, the mediator, and recognition can be measured in terms of correlations

domains such as predicting which university is better (Hertwig and Todd 2003), or which political party will win more votes in an election (Marewski et al. 2009a). One reason why alternatives with larger criterion values are more often recognized is that they are more often mentioned in the environment: The BBC, CNN, The New York Times, and other *environmental mediators* make it probable that many people will encounter and recognize alternatives with large criterion values. Figure 4 illustrates the ecological rationality of the recognition heuristic in terms of three correlations. There is a criterion, an environmental mediator, and a person who infers the criterion. Using the recognition heuristic is ecologically rational when there is both a substantial *ecological correlation* between the mediator and the criterion and a substantial *surrogate correlation* between the mediator and recognition. This combination can yield a substantial *recognition correlation*; that is, recognized alternatives tend to have higher criterion values than unrecognized ones. If either or both the ecological and surrogate correlations are zero, the use of the recognition heuristic is not ecologically rational.

When do people rely on the recognition heuristic?

When the correlation between one's recognition of alternatives and the criterion is substantial, people tend to make inferences in accordance with the recognition heuristic (e.g., Goldstein and Gigerenzer 2002; Hertwig et al. 2008). In contrast, when they are less pronounced, people tend not to do so. For instance, Pohl (2006) asked people to infer which of two cities is situated farther away from the Swiss city of Interlaken, and which of the two cities is larger. Most people may have intuitively known that their recognition of city names is not indicative of the cities' spatial

distance to Interlaken but is indicative of their size, and indeed, for the very same cities, people tended not to make inferences in accordance with the recognition heuristic when inferring spatial distance but seemed to rely on it when inferring size. There is also evidence for a range of other determinants of people's reliance on the recognition heuristic (e.g., Newell and Fernandez 2006; Pachur et al. 2008; Pachur and Hertwig 2006). In particular, it seems that the recognition heuristic is relied upon by default (Volz et al. 2006). In a series of studies, Marewski et al. (2009a) provided evidence that this default can be overruled in a number of situations in which recognition is unlikely to be predictive of the criterion to be inferred, for instance, when the strength of the recognition signal is weak. In these studies, they also showed that the simple recognition heuristic outperforms a number of more complex alternative models in predicting people's behavior.

The recognition heuristic in the wild

What is a good strategy for betting money on sports events? Take Wimbledon, for example. For an upcoming match, one could check the rankings of the Association of Tennis Professionals (ATP), which are based on a large amount of information about the players' previous performance. One then could predict that the player with the higher ranking would win the match. The recognition heuristic offers an alternative. Serwe and Frings (2006) pitted the recognition heuristic against the ATP rankings to predict the results of Wimbledon 2003. They speculated that this may be a difficult environment for the recognition heuristic, as it is dynamic. People would still recognize players who used to be successful some time ago but are no longer quite as good. To their surprise, the recognition of amateur tennis players predicted the actual winner better than did the ATP rankings. This result was replicated for Wimbledon 2005 by Scheibehenne and Bröder (2007) who showed that recognition can also outperform the seedings of the Wimbledon experts. Moreover, the recognition heuristic was also shown to successfully predict other sports events, such as the European Soccer Championship 2004 (Pachur and Biele 2007), or which of two Canadian hockey players has more career points (Snook and Cullen 2006). The recognition heuristic is also of help in marketing—where billions of dollars are spent every year in the United States alone. Daniel Goldstein, professor at London Business School, suggested how marketing strategists could exploit its principle to attract attention to unknown brands; for instance, an unknown product could be placed among well-known ones. In a reversal of the heuristic, a lack of recognition of that one product among many could draw consumers' attention to it (Goldstein 2007). For instance, an unknown brand name such as

“Dr. Gaissmaier’s Incredible Chocolate” may call a consumer’s attention when placed together with more famous products such as Milka, Hershey’s, and Cadbury in the shelves. In short, using the recognition heuristic where it is ecologically rational can help a person make better decisions than when relying on more complex, information-intensive strategies.

Fluency heuristic

The recognition heuristic operates on a binary representation of recognition: An alternative is simply either recognized or unrecognized. But this heuristic essentially discards information that could be useful when two alternatives are both recognized but one is recognized more strongly than the other—a difference that could be exploited by another strategy. The strategy that fills this gap is the fluency heuristic. This heuristic has been defined in different ways (e.g., Jacoby and Brooks 1984; Whittlesea 1993). Here, we use the term to refer to Schooler and Hertwig’s (2005) model, which builds on these earlier definitions, and a long research tradition on fluency (e.g., Jacoby and Dallas 1981), and related notions such as *accessibility* (e.g., Bruner 1957; Tulving and Pearlstone 1966), *availability* (Tversky and Kahneman 1973), or *familiarity* (e.g., Dougherty et al. 1999; Gillund and Shiffrin 1984; Hintzman 1988; Mandler 1980; for a discussion of the similarities between different notions of recognition, availability, and fluency, see Hertwig et al. 2008; Hertwig et al. 2005; Sedlmeier et al. 1998; Schooler and Hertwig 2005). The fluency heuristic is defined as follows:

Search rule: If two alternatives are recognized, determine which one is retrieved faster from memory, that is, which one is recognized more quickly.

Stopping rule: Stop once one alternative is classified as having been recognized more quickly.

Decision rule: Infer that this alternative has the higher value with respect to the criterion.

When is it ecologically rational to rely on the fluency heuristic?

Like the recognition heuristic, the fluency heuristic is a specialized tool. First, it can only be relied on when both alternatives are recognized and when one alternative is more quickly retrieved than the other. An alternative’s retrieval time largely depends on a person’s history of past encounters with the alternative. Roughly speaking, the more often and the more recently an alternative, say, the name of a car brand, is encountered, the more quickly it will be retrieved. Second, using the fluency heuristic is only ecologically rational when the frequency and recency

of encounters with alternatives, and consequently, their retrieval time, correlates with the alternatives’ values on a given criterion. Again, environmental mediators can create such correlations by making it more likely to encounter alternatives that have larger values on the criterion. Thus, the names of, say, popular cars tend to be more quickly retrieved than the names of less popular ones. Ultimately, a person can rely on retrieval time to correctly infer which of two alternatives scores a higher value on the criterion, such as which of two cars is more popular. In short, the ecological rationale of the fluency heuristic resembles very closely that of the recognition heuristic, which is illustrated in Fig. 4. And just like the recognition heuristic, the fluency heuristic has been shown to yield accurate inferences for a range of criteria, including inferences about record sales of music artists (Hertwig et al. 2008), countries’ gross domestic product (Marewski and Schooler 2009), and the size of cities (Schooler and Hertwig 2005), among others.

When do people rely on the fluency heuristic?

As for the recognition heuristic, a number of mechanisms might be responsible for people’s use of the fluency heuristic (Hertwig et al. 2008). Most recently, in a series of experimental and computer simulation studies, Marewski and Schooler (2009) showed that the fluency heuristic is most likely applicable when a person has little or no knowledge about the alternatives in question. In such situations of limited knowledge, differences in retrieval times tend to be large and easier to detect, favoring the applicability of the fluency heuristic. A person is less likely to be able to apply the fluency heuristic when knowledge is abundant, because in this case differences in retrieval times tend to be small. Knowledge-based strategies, in turn, can only be relied upon when knowledge is available. Correspondingly, people will most likely be able to rely on the fluency heuristic when they cannot rely on knowledge instead, and vice versa. At the same time, Marewski and Schooler’s data suggest that the least effort and time is involved in applying the fluency heuristic when using this strategy is also most likely to result in accurate decisions, illustrating how this heuristic can be more easily relied upon when it is, in fact, ecologically rational to use.

The fluency heuristic in the wild

Fluent processing stemming from previous exposure can increase the perceived truth of repeated assertions (e.g., Begg et al. 1992; Hertwig et al. 1997), the perceived fame of names (Jacoby et al. 1989), and the perceived geographical distance of cities (Alter and Oppenheimer 2008). Moreover, there is evidence that people’s sense of fluency

can be used to predict the performance of stocks (Alter and Oppenheimer 2006; Marewski and Schooler 2009), and the fortunes of billionaires (Hertwig et al. 2008). Another situation where the fluency heuristic may play a role is consumer choice: People increasingly like objects when they are repeatedly exposed to them, which is called the *mere exposure effect* (Zajonc 1968). For instance, experimental research suggests that priming a familiar brand increases the probability that it will be considered for purchase (e.g., Coates et al. 2004), while at the same time, only a single exposure can lead people to consider buying a novel brand (e.g., Coates et al. 2006). Just as for the recognition heuristic, a sense of fluency might thus drive consumers to choose certain products over others—a mechanism that corporations exploit when marketing their brands. In sum, the fluency heuristic is a simple cognitive tool that can yield accurate decisions when the recognition heuristic cannot be used.

Betting on one good cue: the take-the-best heuristic

While the fluency heuristic and the recognition heuristic rely on retrieval fluency and recognition, other heuristics use knowledge about alternatives' attributes as *cues* to help making judgments. For instance, when judging which of two newspapers is of better quality one could consider whether the newspapers are nationally distributed. Being a national newspaper might be a positive cue to quality; being a local newspaper, in turn, might be a negative cue, indicating poorer quality. Another attribute to consider might be whether the newspapers are published in a capital city. One can also think of such positive and negative cues as being coded with numbers, such as “1” (positive) and “−1” (negative). Sometimes, a person might not know an alternative's attribute, for instance whether a particular newspaper is published in a capital city or not. In this case, the cue for this particular newspaper can be coded with “0” (unknown).

A prominent representative of such knowledge-based heuristics is Gigerenzer and Goldstein's (1996) take-the-best heuristic, which belongs to the family of lexicographic heuristics that have been developed for choice by Payne et al. (1993). It considers cues sequentially (i.e., one at a time) in the order of their *validity*. The validity of a cue is the probability that an alternative A (e.g., a newspaper) has a higher value on a criterion (e.g., quality) than another alternative B, given that Alternative A has a positive value on that cue and Alternative B a negative or unknown value. Take-the-best bases an inference on the first cue that discriminates between alternatives, that is, on the first cue for which one alternative has a positive value and the other a negative or unknown one. Take-the-best is defined as follows:

Search rule: Look up cues in the order of validity.

Stopping rule: Stop search when the first cue is found that discriminates between alternatives.

Decision rule: Decide for the alternative that this cue favors.

When is it ecologically rational to rely on take-the-best?

Take-the-best ignores available information by looking up cues in the order of their validity and basing an inference on the first discriminating cue. In doing so, it ignores dependencies between cues. Many complex rational models, in contrast, integrate all available information into a judgment, for instance, by weighting and adding it. Now, if a decision maker has unlimited access to information and enough computational power and time to weight and add, say, by computing a multiple regression in his head, should he rely on take-the-best or on strategies that integrate all information? Czerlinski et al. (1999) tried to answer this question by predicting a range of diverse phenomena in 20 different real-world environments, ranging from rainfall to house prices. As Fig. 5 shows, take-the-best outperformed multiple regression in making more accurate predictions in the majority of those 20 environments. At the same time, take-the-best only considered 2.4 cues on average, while multiple regression used 7.7 cues on average, demonstrating how less information can be more. Note that just as Fig. 3, Fig. 5 also illustrates the difference between fitting

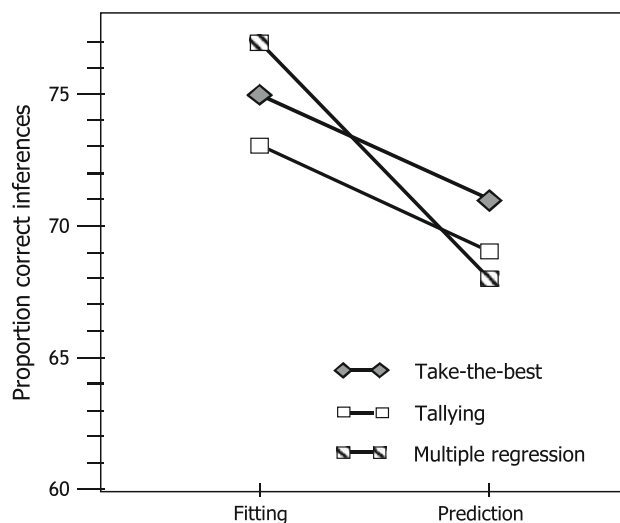


Fig. 5 Accuracy of three different models: multiple regression, tallying, and take-the-best (Czerlinski et al. 1999). The proportion of correct inferences of the models is depicted for fitting existing data (*left*) and for predicting new data (*right*). Naturally, all models do better in fitting than in prediction. More complex models, such as multiple regression, however, forfeit much more accuracy when predicting new data than do more simple models such as tallying and take-the-best

existing data and predicting new data: The regression model, which is the model that fits the existing data best, is the least robust model—it does the poorest job in generalizing to new data. Brighton (2006; see also Gigerenzer and Brighton 2009) pushed these findings to the limit in a competition between take-the-best and heavy-weight computation machineries such as the classification and regression trees (CART, Breiman et al. 1984) or the decision tree induction algorithm C4.5 (Quinlan 1993). He showed that, across the same 20 data sets, it was the rule rather than the exception that take-the-best even outperformed these extreme manifestations of *complexicus* in predicting new data. In short, even if a decision maker could integrate all information and engage in heavy computations, in many environments he would be better off not doing so but using take-the-best instead.

A few general principles have been identified that allow predicting in which environments take-the-best will be more successful than more complex alternative models, and where it will be inferior (see Gigerenzer and Brighton 2009, for a recent overview). In environments consisting of binary cues, take-the-best will be as successful in fitting existing data as any linear model if the cue validities and the cue weights in the linear model have the same order *and* if the weight of each cue cannot be exceeded by the sum of the weights of the subsequent cues (Martignon and Hoffrage 1999). If this is the case, one calls the cue weights *noncompensatory*. To illustrate, it is impossible to compensate for a cue weight of 1 if the weight of each subsequent cue will always be half of the previous cue, such as would be the case with weights 1/2, 1/4, 1/8, and so forth. If the cue weights are skewed rather than strictly non-compensatory, this combination of cues on average still matches or beats tallying (a model incorporating all cues, but weighting them equally, described in detail later) and multiple regression (Hogarth and Karelaia 2007; see also Hogarth and Karelaia 2005, for similar results for a generalization of take-the-best to tasks involving more than two alternatives).

In prediction, take-the-best is often as good as or better than models that take into account more information if the cues are highly intercorrelated with each other, that is, if there is a lot of redundancy in the environment (Dieckmann and Rieskamp 2007). To illustrate, imagine the extreme case of cue intercorrelations of 1. Here, all cues carry identical information, such that there is no difference between only looking at one of them and looking at all of them.

When do people use take-the-best?

Numerous experiments have been conducted that investigate people's use of this simple heuristic (e.g., Bergert and Nosofsky 2007; Bröder and Gaissmaier 2007; Glöckner and

Betsch 2008; Bröder and Schiffer 2003, 2006; Newell and Shanks 2003; Rieskamp 2006; Rieskamp and Otto 2006). In general, people tend to make inferences consistent with take-the-best when applying it is ecologically rational. However, there is also a range of other determinants of strategy selection (see Bröder 2009; Bröder and Newell 2008; for overviews). Bröder and Gaissmaier (2007) summarized the results from several experiments as showing that the need to retrieve cue information from memory (as opposed to reading information on a computer screen) led people to rely on take-the-best, especially when cues were represented verbally and when working memory load was high. Another task feature boosting people's use of take-the-best is time pressure (Rieskamp and Hoffrage 2008).

Take-the-best in the wild

Imagine a patient is rushed to the hospital with serious chest pain. The doctor suspects acute ischemic heart disease and needs to make a decision quickly: Should the patient be assigned to the coronary care unit or to a regular nursing bed for monitoring?

Many doctors in this situation prefer to err on what they believe is the safe side. As a consequence, some 90% of the patients were sent to the coronary care unit in a rural Michigan hospital, of which only about 25% actually had a myocardial infarction (Green and Mehr 1997). However, there is a price to incorrectly sending people to the coronary care unit, too: An overly crowded unit is expensive and can decrease the quality of care for those who need it, while those who do not require it run the risk of getting a severe infection. Something had to be done. The first approach to dealing with this problem built on the classical assumption that more information must be better. An expert system (the *Heart Disease Predictive Instrument, HDPI*) was developed with which doctors needed to check the presence and absence of combinations of seven symptoms and insert the relevant probabilities into a pocket calculator. This expert system made better allocation decisions than did the physicians, but the physicians did not like using it. It was cumbersome, complicated, and nontransparent.

As an alternative, Green and Mehr (1997) developed a decision tree based on the building blocks of take-the-best (Fig. 6). It ignores all probabilities and asks only a few yes-or-no questions. If a patient has a certain anomaly in his electrocardiogram (i.e., an *ST segment change*), he is immediately admitted to the coronary care unit. No other information is searched for. If that is not the case, a second variable is considered: whether the patient's primary complaint is chest pain. If this is not the case, he is immediately classified as low risk and assigned to a regular nursing bed. No further information is considered. If the

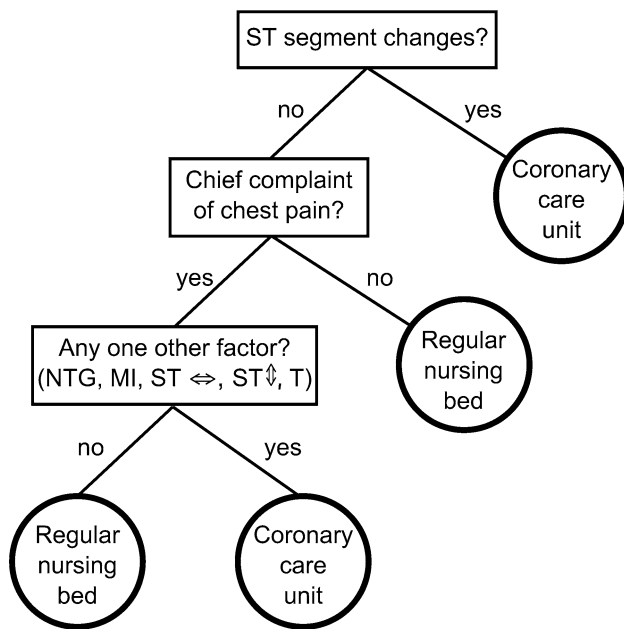


Fig. 6 Should a patient be sent to the coronary care unit or to a regular nursing bed? A decision tree based on variety of symptoms to help doctors make these decisions (Source: Green and Mehr 1997)

answer is yes, then the third and final question is asked to classify the patient.

Not only was the decision tree simpler than the logistic regression and thereby more easily accepted by the physicians, but it was also more successful. Figure 7 shows the performance of the decision tree in comparison with the HDPI and with the physicians' initial performance: The decision tree reached a higher sensitivity (proportion of patients correctly assigned to the coronary care unit) and a lower false-positive rate (proportion of patients incorrectly assigned to the coronary care unit) than both the HDPI and the physicians. It did so by considering only a fraction of the information that the HDPI used.

On a side note, a model closely related to take-the-best has been applied to the social scientists' world, namely, to prioritizing literature searches from the PsycINFO database: Lee et al. (2002; see also Van Maanen and Marewski 2009) examined the performance of a take-the-best variant in identifying articles that are relevant to a given topic of interest (e.g., eyewitness testimony), in which they simply ordered the cues in terms of the evidence they provide in favor of an article being relevant. If an unread article is found based on this cue, the search is terminated without considering any further cues. A researcher going by this take-the-best variant would have had to read fewer articles in order to find the relevant ones than a person behaving in accordance with an alternative Bayesian model. In contrast to the take-the-best variant, the Bayesian model integrated all available cues such as the authors of the article, the

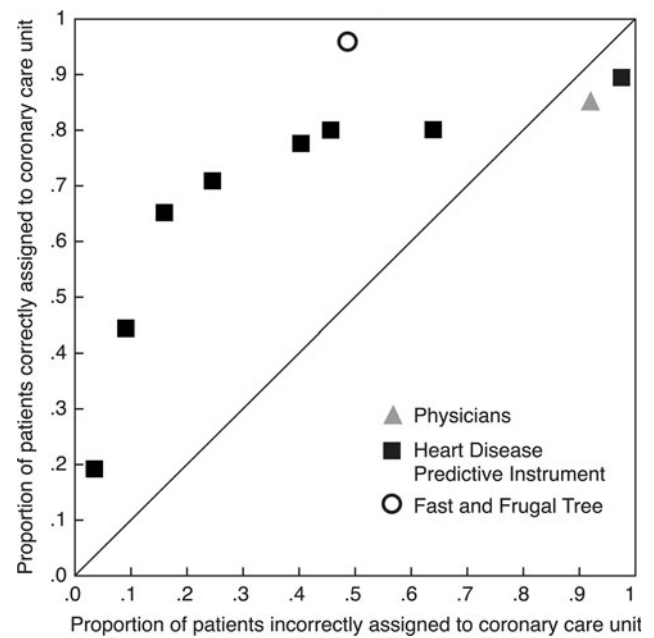


Fig. 7 The performance of a decision tree for coronary care unit allocations in comparison with a logistic regression tool, the Heart Disease Predictive Instrument (HDPI), and physicians. Performance is measured on two dimensions. The y-axis represents the proportion of patients who have been correctly assigned to the coronary care unit (sensitivity), and the x-axis represents the proportion of patients who have been incorrectly assigned to the coronary care unit (false-positive rate). Note that the HDPI classification depends on how one wants to trade sensitivity off against the false-positive rate, which is why there are several data points for the HDPI. Figure adapted from Green and Mehr (1997)

journal it appeared in, or the keywords in the abstract to rate the articles' relevance. In short, take-the-best is a powerful yet simple heuristic that helps people make inferences when knowledge is available beyond a sense of recognition or fluency.

Simply counting cues: tallying

Homo sapiens can simplify in more than one way. While take-the-best ignores cues (but includes a simple form of weighting cues by ordering them), tallying ignores weights, weighting all cues equally. Robin Dawes already analyzed such a heuristic in the 1970s, which we today call tallying. It entails simply counting the number of cues favoring one alternative in comparison with other alternatives.

Search rule: Look up cues in any order.

Stopping rule: Stop search when m out of a total of M cues (with $1 < m \leq M$) have been looked up, and determine which alternative is favored by more cues; if the number of positive cues is the same for both alternatives, search for another cue. If no more cues are found, guess.

Decision rule: Decide for the alternative that is favored by more cues.

When is it ecologically rational to rely on tallying?

Dispensing weights obviously simplifies the task, but can it also be successful? In the original demonstrations (Dawes 1979; Dawes and Corrigan 1974), tallying proved to be almost as successful as multiple regression, and sometimes even better. In a more extensive test across a wide variety of environments, Czerlinski et al. (1999) replicated this result (see Fig. 5). Since multiple regression tended to overfit the data, tallying had, on average, a higher predictive accuracy. The point, however, is not that tallying is always superior to multiple regression. The interesting question is to figure out *when* this is the case. Einhorn and Hogarth (1975) found that unit weight models were successful in comparison with multiple regression when the ratio of alternatives to cues was 10 or smaller, when the linear predictability of the criterion was small ($R^2 \leq .5$), and when cues were highly intercorrelated.

When do people rely on tallying strategies?

So far, relatively few studies have identified conditions under which people would predominantly use a tallying strategy. Interestingly, it seems that people prefer to dispense with particular cues (such as take-the-best does) than with cue weights. A reanalysis of a total of 5 experiments including 415 participants could only identify 83 participants who seemed to have relied on tallying, compared to 198 participants relying on take-the-best (Bröder and Gaissmaier 2007). Moreover, there were even slightly more participants who seemed to rely on a weighted additive strategy than on tallying (90 vs. 83). This is consistent with results by Rieskamp and Hoffrage (2008) who also found that the majority of participants relied on either a general form of take-the-best (called LEX) or a weighted additive strategy, but that a unit weight tallying strategy (called ADD in that paper) only captured very few participants. We can only speculate why this is the case.

One reason might be methodological: The aforementioned studies have investigated rather small worlds, in which participants could easily consider the cue weights for all cues, as there are usually not more than 4 or 5 of them. In larger worlds, where several cues are available, one might expect that people dispense with the cue weights and—as the number of cues increases—with the cues themselves, similar to how people seem to form consideration sets of alternatives in consumer choice.

Another reason might be ecological: In the wild, people often face a trade-off between exploiting available

information and exploring it. Unit weight strategies such as tallying seem to aid exploring available information. By considering all cues, and weighing them equally, a person can learn which cues work best in a given task. Strategies such as take-the-best, in turn, help exploit information, basing decisions on the cues that are known to work best and ignoring the rest. In fact, Rieskamp and Otto (2006) provided evidence to suggest that people will often start out at an unknown task with exploring all cues, which looks like a weighted additive strategy. As their knowledge about the cues increases, they tend to consider fewer. This also matches the observation that ignoring information is part of expertise in the domain of medicine. For instance, contrary to the assumption that more information is always better, more knowledgeable medical professionals have been found to reach better decisions by using less information than less knowledgeable medical professionals (Reyna and Lloyd 2006). Similarly, in an experimental study, both policemen and burglars relied on much fewer pieces of information to predict which of two residential properties was more likely to be burgled than did graduate students (García-Retamero and Dhami 2009). More generally, what has been labeled analytic, or deliberate thinking often decreases with the acquisition of expertise, particularly in routine environments (e.g., Ericsson et al. 2007; Shanteau 1992).

Tallying in the wild

Despite the global financial crisis that is happening while we write this article, which is yet another demonstration that stocks are notoriously uncertain, decision makers might want to invest their money in N funds. They need to decide how to allocate their financial resources and could consider it a good idea to look into what people do who have more expertise. Harry Markowitz is such an expert: He developed a strategy to optimally allocate money, the mean–variance portfolio, for which he received the Nobel Prize. What would be more reasonable than to assume that he would invest his capital according to this portfolio? Well, he did not. Instead he relied on a variant of the tallying (or unit weight) model, $1/N$, which simply allocates financial resources equally across all alternatives. Such a simple strategy cannot be successful, can it? It can, as was demonstrated with a comparison of the $1/N$ rule with 14 optimizing models in seven investment problems (DeMiguel et al. 2009). To estimate the models' parameters, each optimizing strategy received 10 years of stock data and then had to predict the next month's performance on this basis. The same procedure was then done, with a moving window, for the next month, and so forth, until no data was left. Note that $1/N$ does not have any free parameters that need to be estimated. Nevertheless, it was quite successful on several financial criteria. It came out first on certainty equivalent

returns, second on turnover, and fifth on the Sharpe ratio, respectively. None of the far more complex optimizing models could consistently beat $1/N$.

Tallying strategies are also used in decision aids for avoiding avalanche accidents when traveling in avalanche terrain: The *obvious clues method* checks how many out of seven cues have been observed en route or at the slope that is evaluated (McCammon and Hägeli 2007). These cues include whether there has been an avalanche in the last 48 h and whether there is liquid water present on the snow surface as a result of recent sudden warming. When more than three of these cues are present on a given slope, it should be considered dangerous according to the *obvious clues method*. With this simple tallying strategy, 92% of the historical accidents (where the method would have been applicable) could have been prevented.

In short, tallying and its variants for capital allocation and avalanche forecasting are successful rules of thumb in an uncertain world: Throwing away weights makes these strategies at the same time simple and robust enough for predicting uncertain events.

How cognitive limitations can be beneficial

“Everyone blames his memory, no one blames his judgment,” de La Rochefoucauld (1898, p. 13) proposed in “Reflections; or Sentences and Moral Maxims” in 1678. Actually, if it were not for our imperfect memory, we probably would need to blame our judgment more. Not that we have a choice, but if we did, it seems that we would need to choose to either have a perfect memory, but worse judgment as a result of it, or to live with our perforated memory, but be able to make good judgments.

Of course it is not as simple as that, but it is not only heuristics for making judgments and decisions that benefit from being simple. Also, the core capacities that the heuristics exploit—such as memory—filter information, separating out the irrelevant and highlighting the relevant. As we will demonstrate next, there is growing evidence from several domains that cognitive limits, biases, and other constraints in our core capacities can be beneficial (for an overview, see Hertwig and Todd 2003).

A limited memory aids heuristic inference

An important function of memory is not simply to store all information that is encountered, but rather to provide the cognitive system with important, relevant information when it is needed. According to Anderson and colleagues’ *rational analysis* of memory (e.g., Anderson and Milson 1989; Anderson and Schooler 1991, 2000; Schooler and Anderson 1997), the cognitive system retrieves memories

as a function of how likely they will be needed to achieve some processing goal.⁴ In doing so, human memory capitalizes on a person’s history of past encounters with objects (e.g., stock names), which, in turn, can be indicative of how likely objects are to reoccur in the environment and be needed in the future. In their view, human memory essentially makes a bet, namely, that as the recency and frequency with which a piece of information has been encountered increases, so too does the probability that this information will be needed to achieve a given processing goal in the future. Conversely, the more time that has passed since an object has been encountered, the less is the likelihood that memories of the object will need to be retrieved in the future and, ultimately, memories of such objects can be forgotten. This way, memory drops outdated, largely irrelevant information and gives a retrieval advantage to recently and frequently encountered, most likely more relevant information.

To illustrate their point, Anderson and Schooler (1991) analyzed environments consisting of text and word utterances. For instance, in an analysis of recorded conversations of children’s speech, they observed that the probability of a particular word utterance decreased as a function of the amount of time that passed since the word was last uttered. Similarly, the likelihood of recalling a memory of a given object drops as a function of the amount of time since the object was last encountered. In fact, in various environments, they found strong correspondences between regularities in the patterns of occurrence of information (e.g., a word’s recency and frequency of occurrence) and the classic forgetting and learning functions (e.g., as described by Ebbinghaus 1885/1964).

The interplay between cognitive limitations such as the forgetting of information and the workings of heuristics can be illustrated with both the recognition and the fluency heuristic. In computer simulations and mathematical analyses, Goldstein and Gigerenzer (2002) showed how knowing less can help a person make more accurate inferences when using the recognition heuristic. In later computer simulations, Schooler and Hertwig (2005) then provided evidence that some forgetting could additionally fuel the accuracy of both the recognition and the fluency heuristic, providing these heuristics with the most relevant information, rather than with all the information a decision

⁴ In contrast to the fast and frugal heuristics framework, the rational analysis makes the assumption that the goal of an organism is to respond optimally to a given stimulus. The Bayesian analysis, which has been used to define the optimal response, has been criticized (e.g., Brighton and Gigerenzer 2008), and in fact, a key tenet of the fast and frugal heuristic framework is that organisms do not need to optimize in order to behave adaptively. Quite the contrary: often optimization is not an option, for instance, because a problem is intractable or because it is ill defined (e.g., Gigerenzer 2004b).

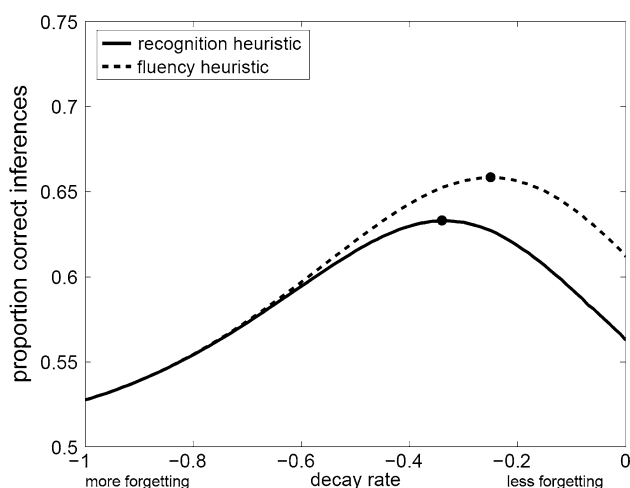


Fig. 8 Performances of the recognition and fluency heuristics vary with decay rate, that is, the amount of forgetting in declarative memory: Both too much forgetting (on the very left end of the *x*-axis) and too little forgetting (on the very right end of the *x*-axis) hurt the proportion of correct inferences that can be made by using the heuristics. Intermediate levels of forgetting are best (Reprinted with permission from Schooler and Hertwig 2005)

maker would have been able to accumulate throughout her lifetime if she had a perfect memory (Fig. 8).

In a related vein, Marewski and Schooler (2009) argued that limits in the ability of the human memory system to detect certain characteristics of memories can help a person choose between different heuristics when facing a task. In particular, their series of computer simulation studies and experiments suggest that the interplay between the human memory system and the structure of information in the environment constrains the choice set of heuristics that can be applied to solve a given task, giving rise to what they called different *cognitive niches* of the heuristics in the adaptive toolbox—niches that can make it easier and faster to rely on a given heuristic when using it is also most likely to result in accurate inference.

Less capacity can be more

Contrary to the view that many cognitive biases result from people being not smart enough, there is evidence that some biases actually *decrease* with limited cognitive capacities (or are related by a U-shaped function to them; see Weir 1964). Let us illustrate this point with a classical choice anomaly, *probability matching* (see Vulkan 2000 for a review of the literature). In the typical task, people repeatedly have to predict which of two events would occur on the next trial. Assume that event E1 occurs with a probability of $p(E1) = 0.75$, while event E2 only occurs with $p(E2) = 1 - p(E1) = 0.25$. Given that the succession of events is serially independent, the best people could do is always predict the more frequent event E1. This strategy

is called *maximizing* and would yield an average accuracy of 75%. However, the modal strategy is *probability matching*, that is, to predict the events in proportion to their probability of occurrence, with an expected accuracy of only 62.5% on average ($0.75 \times 0.75 + 0.25 \times .25$).

Why do people fail to see the optimal solution in such a simple task? The typical assumption is that people are not smart enough (e.g., West and Stanovich 2003). In contrast with that view, however, there is convergent evidence showing that beings with lower cognitive capacities, such as children, pigeons, people with a lower short-term memory capacity or people who are distracted by a secondary task, are more likely to maximize than human adults (e.g., Weir 1964; Kareev et al. 1997; Wolford et al. 2004). The finding that lower cognitive capacities actually foster maximizing instead of preventing it suggests that probability matching could be the result of a more complex strategy—people explore the space of hypotheses on how to improve performance on the task. One hypothesis participants typically entertain in probability learning tasks is that there are patterns in the sequence, and any reasonable pattern tends to match the probabilities. Since there are no patterns, however, searching for patterns is counterproductive. Therefore, people who do not search for patterns, for example, because of capacity limitations, are more likely to settle on maximizing and will be more successful. In fact, Gaissmaier and Schooler (2008) showed that probability matchers who would typically be assumed to be irrational are actually better in finding patterns in sequences. Likewise, exploratory behavior resulting in probability matching can help people detect changes in the environment (Gaissmaier et al. 2006). What works poorly in one environment, namely, searching for patterns if there are none and ending up with suboptimal probability matching, may work well in another where it can help to actually find patterns if they exist, in this way adapting to changes in the environment.

Similarly, DeCaro et al. (2008) demonstrated that lower working memory capacity actually helped people learn category structures that require implicit procedural learning, which is largely outside of conscious control. For category structures that can be more successfully learned based on explicit hypothesis testing, however, the opposite was true: Here, participants with a higher working memory capacity were more successful. Congruent with the probability matching case, this example illustrates Simon's (1990) point that one always needs to consider both the structure of the task and the cognitive capacities of the actor.

The importance of starting small

Many of our international colleagues who have joined the ABC Research Group in Berlin take advantage of their stay in Germany and try to learn German. Quite often, these

colleagues find it frustrating how hard it is to master the language. And some of them suffer from the even more frustrating experience of having their children confront them with fluent German they have picked up in the local kindergarten or school. Why do children achieve a high level of mastery of this difficult language, in which even substantives have one of three genders? Part of their success seems to result from their cognitive limitations.

Newport (1990) formulated this less-is-more hypothesis with regard to language acquisition. She proposed the following: Adults have greater capacities than children. They are thus able to learn language strings as a whole and do not have to analyze them further. In this manner, they are learning faster, but they are not able to reach the final level of mastery. Children, in contrast, are not able to learn language strings as a whole due to their smaller capacities, so they have to break them down into fragments. This enables them to learn the internal structure of the language and to successfully recombine parts of the strings in novel forms. Thus, they are learning more slowly, but they are better in the end. Building on this idea, Elman (1993) stressed the importance of “starting small.” In a series of neural network simulations, he showed that a system with unlimited processing capacity failed to master the grammar of a complex language if confronted with the entire language. Instead, mastery was possible for a limited but gradually increasing system. Cochran et al. (1999) took an experimental approach to study this phenomenon. Two groups of adults had to learn a sign language. One of these groups was given a concurrent task during the learning phase to decrease working memory capacity. In the other group, there was no concurrent task. Adults learning the sign language without a concurrent task showed faster learning of studied materials. But the concurrent task group did better in generalizing what they had learned to new contexts, what can be seen as a higher level of mastery. As an alternative to limiting the learning capacities, Kersten and Earles (2001) reduced the complexity of the learning environment, with the same effect: Adults learned a miniature artificial language better if presented with small bits of language than when confronted with the full complexity at once.

Thinking less can be more

“When you ask the centipede how it manages to coordinate its thousand feet, it will stumble,” an almost prototypical example in lessons about skill acquisition goes. It nicely illustrates that once one has completely internalized a motor skill, it is good not to think about how exactly one does it anymore. Demonstrating how too much thinking can hurt, Beilock et al. (2002) found that expert golfers performed better when they were distracted by a secondary task than when they were asked to pay attention to their

performance. For novices, in turn, distraction did not pay—they played better when focusing on their skills.

Thinking less is also beneficial for experienced handball players. Johnson and Raab (2003) showed players videos of situations from a game. After a while, the video was stopped and held in a freeze frame. Players were instructed to imagine they were the player with the ball in the video and to spontaneously say what they would do next. Would they pass the ball to another player, shoot on the goal, or do something else? After their spontaneous reaction, they were given more time and asked to generate as many further options as possible. Finally, they were asked to choose one of all the options they generated. Their choices and all the other options they generated were later on evaluated by four qualified professional-league handball coaches. The first, spontaneous option was the best option on average. The more players thought about it, the more alternative options they generated, the more they started to distrust their initial intuition and the more they ended up deciding for an option that was actually worse than their first, spontaneous idea.

In short, these examples illustrate that complex tasks do not always need extensive thinking. Rather, sometimes even the opposite seems to be true: Less thinking can outperform more thinking in certain situations, and the limitations of their cognitive machinery can help people think less.

Conclusion

Do good judgments need complex cognition? A glance into the literature, which is populated with complex Bayesian and other models, suggests that the answer is yes. Countering this view, here we have reviewed evidence to suggest that actually the opposite may be true: Simple cognitive mechanisms can outperform more complex cognitive machinery, which is prone to overfit irrelevant, noisy, and meaningless information in a fundamentally uncertain world. As human cultural evolution continues at a rapid pace, the environment in which we humans live changes dramatically. While the world continues to have its dice and random number generators built in, today humans not only have to forage for food and mates, but also need to cope with vast environmental pollution, massive information overload, colliding cultures, ever more destructive weaponry, and extremely sophisticated technology. One can only hope that *Homo sapiens'* cognition has the right degree of simplicity and robustness to pass this ongoing test of generalizing to new situations.

Acknowledgments We would like to thank Ulrich Hoffrage, Markus Knauff, Ann Wolf, and one anonymous reviewer for many helpful comments on an earlier draft of this work. We would like to thank Anita Todd for editing the manuscript. This research was funded by the Max Planck Institute for Human Development, Berlin, Germany.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Alba JW, Chattopadhyay A (1985) Effects of context and part-category cues on recall of competing brands. *J Mark Res* 22:340–349
- Alter AL, Oppenheimer DM (2006) Predicting short-term stock fluctuations by using processing fluency. *Proc Natl Acad Sci USA* 103:9369–9372
- Alter AL, Oppenheimer DM (2008) Effects of fluency on psychological distance and mental construal (or why New York is a large city, but New York is a civilized jungle). *Psychol Sci* 19:161–167
- Anderson JR, Milson R (1989) Human memory: an adaptive perspective. *Psychol Rev* 96:703–719
- Anderson JR, Schooler LJ (1991) Reflections of the environment in memory. *Psychol Sci* 2:396–408
- Anderson JR, Schooler LJ (2000) The adaptive nature of memory. In: Tulving E, Craik FIM (eds) *Handbook of memory*. Oxford University Press, New York, pp 557–570
- Begg IM, Anas A, Farinacci S (1992) Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *J Exp Psychol Gen* 121:446–458
- Beilock SL, Carr TH, MacMahon C, Starkes JL (2002) When paying attention becomes counterproductive: impact on divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *J Exp Psychol Appl* 8:6–16
- Bergert FB, Nosofsky RM (2007) A response-time approach to comparing generalized rational and take-the-best models of decision making. *J Exp Psychol Learn Mem Cogn* 33:107–129
- Berretty PM, Todd PM, Martignon L (1999) Categorization by elimination. Using few cues to choose. In: Gigerenzer G, Todd PM, the ABC Research Group (eds) *Simple heuristics that make us smart*. Oxford University Press, New York, pp 235–254
- Bishop M (2006) Fast and frugal heuristics. *Philos Compass* 1:201–223
- Brandstätter E, Gigerenzer G, Hertwig R (2006) The priority heuristic: making choices without trade-offs. *Psychol Rev* 113:409–432
- Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) *Classification and regression trees*. Wadsworth International Group, Belmont
- Brighton H (2006) Robust inference with simple cognitive models. In: Lebiere C, Wray B (eds) *Between a rock and a hard place: cognitive science principles meet AI-hard problems*. Papers from the AAAI spring symposium (AAAI Tech. Rep. No. SS-06-03). AAAI Press, Menlo Park, pp 17–22
- Brighton H, Gigerenzer G (2008) Bayesian brains and cognitive mechanisms: harmony or dissonance? In: Chater N, Oaksford M (eds) *The probabilistic mind: prospects for Bayesian cognitive science*. Oxford University Press, New York, pp 189–208
- Bröder A (2009) The quest for take the best—insights and outlooks from experimental research. In: Todd P, Gigerenzer G, the ABC Research Group (eds) *Ecological rationality: intelligence in the world*. Oxford University Press, New York (in press)
- Bröder A, Gaissmaier W (2007) Sequential processing of cues in memory-based multi-attribute decisions. *Psychon Bull Rev* 14:895–900
- Bröder A, Newell BR (2008) Challenging some common beliefs: empirical work within the adaptive toolbox metaphor. *Judgm Decis Mak* 3:205–214
- Bröder A, Schiffer S (2003) Take the best versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *J Exp Psychol Gen* 132:277–293
- Bröder A, Schiffer S (2006) Stimulus format and working memory in fast and frugal strategy selection. *J Behav Decis Mak* 19:361–380
- Bruner JS (1957) On perceptual readiness. *Psychol Rev* 64:123–152
- Coates SL, Butler LT, Berry DC (2004) Implicit memory: a prime example for brand consideration and choice. *Appl Cogn Psychol* 18:1195–1211
- Coates SL, Butler LT, Berry DC (2006) Implicit memory and consumer choice: the mediating role of brand familiarity. *Appl Cogn Psychol* 20:1101–1116
- Cochran BP, McDonald JL, Parault SJ (1999) Too smart for their own good: the disadvantage of a superior processing capacity for adult language learners. *J Mem Lang* 41:30–58
- Coenen A, Marewski JN (2009) Predicting moral judgments of corporate responsibility with formal decision heuristics. In: Taatgen NA, van Rijn H (eds) *Proceedings of the 31st annual conference of the cognitive science society*. Cognitive Science Society, Austin, pp 1524–1528
- Cokely ET, Kelley CM (2009) Cognitive abilities and superior decision making under risk: a protocol analysis and process model evaluation. *Judgm Decis Mak* 4:20–33
- Cokely ET, Schooler LJ, Gigerenzer G (2009) Information use for decision making. In: Maack MN, Bates MJ (eds) *Encyclopedia of library and information sciences* (in press)
- Czerlinski J, Gigerenzer G, Goldstein DG (1999) How good are simple heuristics? In: Gigerenzer G, Todd PM, the ABC Research Group (eds) *Simple heuristics that make us smart*. Oxford University Press, New York, pp 97–118
- Davis JN, Todd PM (1999) Parental investment by simple decision rules. In: Gigerenzer G, Todd PM, the ABC Research Group (eds) *Simple heuristics that make us smart*. Oxford University Press, New York, pp 309–324
- Dawes RM (1979) The robust beauty of improper linear models in decision making. *Am Psychol* 34:571–582
- Dawes RM, Corrigan B (1974) Linear models in decision making. *Psychol Bull* 81:95–106
- Dawkins R (1989) *The selfish gene*, 2nd edn. Oxford University Press, Oxford
- de La Rochefoucauld F (1898) *Reflections; or sentences and moral maxims*. Sampson Low, Marston & Company, London (original work published 1678 and 1827)
- DeCaro MS, Thomas RD, Beilock SL (2008) Individual differences in category learning: sometimes less working memory capacity is better than more. *Cognition* 107:284–294
- DeMiguel V, Garlappi L, Uppal R (2009) Optimal versus naive diversification: how inefficient is the 1/N portfolio strategy? *Rev Fin Stud* 22:1915–1953
- Dieckmann A, Rieskamp J (2007) The influence of information redundancy on probabilistic inferences. *Mem Cogn* 35:1801–1813
- Dougherty MRP, Gettys CF, Ogden EE (1999) Minerva-DM: a memory processes model for judgments of likelihood. *Psychol Rev* 106:180–209
- Ebbinghaus H (1964) *Memory: a contribution to experimental psychology*. Dover, Mineola (original work published 1885)
- Edwards W (1954) The theory of decision making. *Psychol Bull* 51:380–417
- Einhorn HJ, Hogarth RM (1975) Unit weighting schemes for decision making. *Organ Behav Hum Perform* 13:171–192
- Elman JL (1993) Learning and development in neural networks: the importance of starting small. *Cognition* 48:71–99
- Ericsson KA, Prietula MJ, Cokely ET (2007) The making of an expert. *Harv Bus Rev* 85:114–121

- Fechner GT (1966) Elements of psychophysics (trans: Adler HE). Holt, Rinehart and Winston, New York (original work published 1860)
- Fischer JE, Steiner F, Zucol F, Berger C, Martignon L, Bossart W et al (2002) Use of simple heuristics to target macrolide prescription in children with community-acquired pneumonia. *Arch Pediatr Adolesc Med* 156:1005–1008
- Franklin B (1987) Writings. The Library of America, New York
- Gaissmaier W, Schooler LJ (2008) The smart potential behind probability matching. *Cognition* 109:416–422
- Gaissmaier W, Schooler LJ, Rieskamp J (2006) Simple predictions fueled by capacity limitations: when are they successful? *J Exp Psychol Learn Mem Cogn* 32:966–982
- Gaissmaier W, Schooler LJ, Mata R (2008) An ecological perspective to cognitive limits: modeling environment-mind interactions with ACT-R. *Judgm Decis Mak* 3:278–291
- García-Retamero R, Dhami MK (2009) Take-the-best in expert-novice decision strategies for residential burglary. *Psychon Bull Rev* 16:163–169
- Gigerenzer G (2002) Calculated risks: how to know when numbers deceive you. Simon & Schuster, New York
- Gigerenzer G (2004a) Fast and frugal heuristics: the tools of bounded rationality. In: Koehler D, Harvey N (eds) *Handbook of judgement and decision making*. Blackwell, Oxford, pp 62–88
- Gigerenzer G (2004b) Striking a blow for sanity in theories of rationality. In: Augier M, March JG (eds) *Models of a man: Essays in memory of Herbert A. Simon*. MIT Press, Cambridge, pp 389–409
- Gigerenzer G (2007) Gut feelings: the intelligence of the unconscious. Viking, New York
- Gigerenzer G, Brighton H (2009) Homo heuristicus: why biased minds make better inferences. *Top Cogn Sci* 1:107–143
- Gigerenzer G, Engel C (eds) (2006) *Heuristics and the law*. MIT Press, Cambridge
- Gigerenzer G, Goldstein DG (1996) Reasoning the fast and frugal way: models of bounded rationality. *Psychol Rev* 104:650–669
- Gigerenzer G, Selten R (2001) Rethinking rationality. In: Gigerenzer G, Selten R (eds) *Bounded rationality. The adaptive toolbox*. MIT Press, Cambridge, pp 1–12
- Gigerenzer G, Todd PM, the ABC Research Group (1999) *Simple heuristics that make us smart*. Oxford University Press, New York
- Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S (2007) Helping doctors and patients make sense of health statistics. *Psychol Sci Public Interest* 8:53–96
- Gigerenzer G, Hoffrage U, Goldstein DG (2008) Fast and frugal heuristics are plausible models of cognition: reply to Dougherty, Franco-Watkins, & Thomas (2008). *Psychol Rev* 115:230–239
- Gillund G, Shiffrin RM (1984) A retrieval model for both recognition and recall. *Psychol Rev* 91:1–67
- Glöckner A, Betsch T (2008) Multiple-reason decision making based on automatic processing. *J Exp Psychol Learn Mem Cogn* 34:1055–1075
- Glöckner A, Betsch T, Schindler N (2009) Coherence shifts in probabilistic inference tasks. *J Behav Decis Mak* (in press)
- Goldstein DG (2007) Getting attention for unrecognized brands. *Harv Busi Rev* 85:24–28
- Goldstein DG, Gigerenzer G (1999) The recognition heuristic: how ignorance makes us smart. In: Gigerenzer G, Todd PM, the ABC Research Group (eds) *Simple heuristics that make us smart*. Oxford University Press, New York, pp 37–48
- Goldstein DG, Gigerenzer G (2002) Models of ecological rationality: the recognition heuristic. *Psychol Rev* 109:75–90
- Green L, Mehr DR (1997) What alters physicians' decisions to admit to the coronary care unit? *J Fam Pract* 45:219–226
- Groner M, Groner R, Bischof WF (1983) Approaches to heuristics: a historical review. In: Groner R (ed) *Methods of heuristics*. Erlbaum, Hillsdale, NJ, pp 1–18
- Hammond KR (1996) *Human judgment and social policy: irreducible uncertainty, inevitable error, unavoidable injustice*. Oxford University Press, Oxford
- Hauser JR, Wernerfelt B (1990) An evaluation cost model of consideration sets. *J Consum Res* 16:393–408
- Hertwig R, Todd PM (2003) More is not always better: the benefits of cognitive limits. In: Hardman D, Macchi L (eds) *Thinking: psychological perspectives on reasoning, judgment and decision making*. Wiley, Chichester, pp 213–231
- Hertwig R, Gigerenzer G, Hoffrage U (1997) The reiteration effect in hindsight bias. *Psychol Rev* 104:194–202
- Hertwig R, Hoffrage U, Martignon L (1999) Quick estimation: letting the environment do the work. In: Gigerenzer G, Todd PM, the ABC Research Group (eds) *Simple heuristics that make us smart*. Oxford University Press, New York, pp 209–234
- Hertwig R, Pachur T, Kurzenhäuser S (2005) Judgments of risk frequencies: tests of possible cognitive mechanisms. *J Exp Psychol Learn Mem Cogn* 31:621–642
- Hertwig R, Herzog SM, Schooler LJ, Reimer T (2008) Fluency heuristic: a model of how the mind exploits a by-product of information retrieval. *J Exp Psychol Learn Mem Cogn* 34:1191–1206
- Hintzman DL (1988) Judgment of frequency and recognition memory in a multiple trace memory model. *Psychol Rev* 95:528–551
- Hoffrage U, Lindsey S, Hertwig R, Gigerenzer G (2000) Communication of statistical information. *Science* 290:2261–2262
- Hogarth RM, Karelaia N (2005) Simple models for multi-attribute choice with many alternatives: when it does and does not pay to face tradeoffs with binary attributes. *Manag Sci* 51:1860–1872
- Hogarth RM, Karelaia N (2007) Heuristic and linear models of judgment: matching rules and environments. *Psychol Rev* 114:733–758
- Howard JA, Sheth JN (1969) *The theory of buyer behavior*. Wiley, New York
- Hutchinson JMC, Gigerenzer G (2005) Simple heuristics and rules of thumb: where psychologists and behavioural biologists might meet. *Behav Process* 69:87–124
- Jacoby LL, Brooks LR (1984) Nonanalytic cognition: memory, perception and concept learning. In: Bower GH (ed) *Psychology of learning and motivation*, vol 18. Academic Press, New York, pp 1–47
- Jacoby LL, Dallas M (1981) On the relationship between autobiographical memory and perceptual learning. *J Exp Psychol Gen* 110:306–340
- Jacoby LL, Woloshyn V, Kelly CM (1989) Becoming famous without being recognized: unconscious influences of memory produced by dividing attention. *J Exp Psychol Gen* 118:115–125
- Johnson J, Raab M (2003) Take the first: option generation and resulting choices. *Organ Behav Hum Decis Process* 91:215–229
- Johnson-Laird PN (1983) *Mental models*. Cambridge University Press, Cambridge
- Kahneman D, Slovic P, Tversky A (eds) (1982) *Judgment under uncertainty: heuristics and biases*. Cambridge University Press, Cambridge
- Kareev Y, Lieberman I, Lev M (1997) Through a narrow window: sample size and the perception of correlation. *J Exp Psychol Gen* 126:278–287
- Katsikopoulos K, Martignon L (2006a) Naïve heuristics for paired comparisons: some results on their relative accuracy. *J Math Psychol* 50:488–494
- Katsikopoulos K, Martignon L (2006b) On the accuracy of lexicographic strategies for pair comparison. *J Math Psychol* 50:116–122

- Kersten AW, Earles JL (2001) Less really is more for adults learning a miniature artificial language. *J Mem Lang* 44:250–273
- Lee MD, Loughlin N, Lundberg IB (2002) Applying one reason decision-making: the prioritisation of literature searches. *Aust J Psychol* 54:137–143
- Leibniz GW (1951) Toward a universal characteristic. In: Wiener PP (ed) *Leibniz: selections*. Scribner's Sons, New York (original work published 1677)
- Mandler G (1980) Recognizing: the judgment of previous occurrence. *Psychol Rev* 87:252–271
- Marewski JN (2009) On the theoretical precision, and strategy selection problem of a single-strategy approach: a comment on Glöckner, Betsch, and Schindler. *J Behav Decis Mak* 22:1–5 (in press)
- Marewski JN, Olsson H (2009) Beyond the null ritual: formal modeling of psychological processes. *J Psychol* 217:49–60
- Marewski JN, Schooler LJ (2009) How memory aids strategy selection. Manuscript submitted for publication
- Marewski JN, Gaissmaier W, Schooler LJ, Goldstein DG, Gigerenzer G (2009a) From recognition to decisions: a competition of formal models for multi-alternative probabilistic inference. Manuscript submitted for publication
- Marewski JN, Galesic M, Gigerenzer G (2009b) Fast and frugal media choices. In: Hartmann T (ed) *Media choice: a theoretical and empirical overview*. Routledge, New York, pp 102–128
- Marewski JN, Schooler LJ, Gigerenzer G (2009c) Five principles for studying people's use of heuristics. *Acta Psychol Sinica* (in press)
- Martignon L, Hoffrage U (1999) Why does one-reason decision making work? A case study in ecological rationality. In: Gigerenzer G, Todd PM, the ABC Research Group (eds) *Simple heuristics that make us smart*. Oxford University Press, New York, pp 119–140
- Mata R, Schooler LJ, Rieskamp J (2007) The aging decision maker: cognitive aging and the adaptive selection of decision strategies. *Psychol Aging* 22:796–810
- McCammon I, Hägeli P (2007) An evaluation of rule-based decision tools for travel in avalanche terrain. *Cold Reg Sci Technol* 47:193–206
- Newell BR (2005) Re-visions of rationality. *Trends Cogn Sci* 9:11–15
- Newell BR, Fernandez D (2006) On the binary quality of recognition and the inconsequentiality of further knowledge: two critical tests of the recognition heuristic. *J Behav Decis Mak* 19:333–346
- Newell BR, Shanks DR (2003) Take the best or look at the rest? Factors influencing 'one-reason' decision making. *J Exp Psychol Learn Mem Cogn* 29:53–65
- Newport EL (1990) Maturation constraints on language learning. *Cogn Sci* 14:11–28
- Ortmann A, Gigerenzer G, Borges B, Goldstein DG (2008) The recognition heuristic: a fast and frugal way to investment choice? In: Plott CR, Smith VL (eds) *Handbook of experimental economics results: vol 1 (Handbooks in Economics No. 28)*. North-Holland, Amsterdam, pp 993–1003
- Pachur T, Biele G (2007) Forecasting from ignorance: the use and usefulness of recognition in lay predictions of sports events. *Acta Psychol* 125:99–116
- Pachur T, Hertwig R (2006) On the psychology of the recognition heuristic: retrieval primacy as a key determinant of its use. *J Exp Psychol* 32:983–1002
- Pachur T, Bröder A, Marewski JN (2008) The recognition heuristic in memory-based inference: is recognition a non-compensatory cue? *J Behav Decis Mak* 21:183–210
- Payne JW, Bettman JR, Johnson EJ (1993) *The adaptive decision maker*. Cambridge University Press, New York
- Pitt MA, Myung IJ, Zhang S (2002) Toward a method for selecting among computational models for cognition. *Psychol Rev* 109:472–491
- Pohl R (2006) Empirical tests of the recognition heuristic. *J Behav Decis Mak* 19:251–271
- Quinlan JR (1993) *C4.5: programs for machine learning*. Morgan Kaufmann, San Mateo
- Reyna VF, Lloyd FJ (2006) Physician decision making and cardiac risk: effects of knowledge, risk perception, risk tolerance, and fuzzy processing. *J Exp Psychol Appl* 12:179–195
- Rieskamp J (2006) Perspectives on probabilistic inferences: reinforcement learning and an adaptive network compared. *J Exp Psychol Learn Mem Cogn* 32:1355–1370
- Rieskamp J, Hoffrage U (1999) When do people use simple heuristics and how can we tell? In: Gigerenzer G, Todd PM, the ABC Research Group (eds) *Simple heuristics that make us smart*. Oxford University Press, New York, pp 141–167
- Rieskamp J, Hoffrage U (2008) Inferences under time pressure: how opportunity costs affect strategy selection. *Acta Psychol* 127:258–276
- Rieskamp J, Otto P (2006) SSL: a theory of how people learn to select strategies. *J Exp Psychol Gen* 135:207–236
- Scheibehenne B, Bröder A (2007) Predicting Wimbledon 2005 tennis results by mere player name recognition. *Int J Forecast* 23:415–426
- Schooler LJ, Anderson JR (1997) The role of process in the rational analysis of memory. *Cogn Psychol* 32:219–250
- Schooler LJ, Hertwig R (2005) How forgetting aids heuristic inference. *Psychol Rev* 112:610–628
- Sedlmeier P, Hertwig R, Gigerenzer G (1998) Are judgments of the positional frequencies of letters systematically biased due to availability? *J Exp Psychol Learn Mem Cogn* 24:754–770
- Serwe S, Frings C (2006) Who will win Wimbledon? The recognition heuristic in predicting sports events. *J Behav Decis Mak* 19:321–332
- Shaffer DN, Krauchunas SM, Eddy M, McBeath MK (2004) How dogs navigate to catch Frisbees. *Psychol Sci* 15:437–441
- Shanteau J (1992) How much information does an expert use? Is it relevant? *Acta Psychol* 81:75–86
- Simon HA (1955) A behavioral model of rational choice. *Q J Econ* 69:99–118
- Simon HA (1956) Rational choice and the structure of the environment. *Psychol Rev* 63:129–138
- Simon HA (1990) Invariants of human behavior. *Annu Rev Psychol* 41:1–19
- Snook B, Cullen RM (2006) Recognizing national hockey league greatness with an ignorance-based heuristic. *Can J Exp Psychol* 60:33–43
- Stevens SS (1957) On the psychophysical law. *Psychol Rev* 64:153–181
- Todd PM, Miller GF (1999) From pride and prejudice to persuasion: realistic heuristics for mate search. In: Gigerenzer G, Todd PM, the ABC Research Group (eds) *Simple heuristics that make us smart*. Oxford University Press, New York, pp 287–308
- Townsend JT (1975) The mind–body problem revisited. In: Chang C (ed) *Philosophical aspects of the mind–body problem*. Honolulu University Press, Honolulu, pp 200–218
- Tulving E, Pearlstone Z (1966) Availability versus accessibility of information in memory for words. *J Verbal Learn Verbal Behav* 5:381–391
- Tversky A, Kahneman D (1973) Availability: a heuristic for judging frequency and probability. *Cogn Psychol* 5:207–232
- Tversky A, Kahneman D (1974) Judgment under uncertainty: heuristics and biases. *Science* 185:1124–1130
- Van Maanen L, Marewski JN (2009) Recommender systems for literature selection: a competition of decision making and memory models. In: Taatgen NA, van Rijn H (eds) *Proceedings of the 31st annual conference of the cognitive science society*. Cognitive Science Society, Austin, pp 2914–2919

- Volz KG, Schooler LJ, Schubotz RI, Raab M, Gigerenzer G, von Cramon DY (2006) Why you think Milan is larger than Modena: neural correlates of the recognition heuristic. *J Cogn Neurosci* 18:1924–1936
- von Helversen B, Rieskamp J (2008) The mapping model: a cognitive theory of quantitative estimation. *J Exp Psychol Gen* 137:73–96
- Vulkan N (2000) An economist's perspective on probability matching. *J Econ Surveys* 14:101–118
- Wegwarth O, Gaissmaier W, Gigerenzer G (2009) Smart strategies for doctors and doctors-in-training: heuristics in medicine. *Med Educ* 43:721–728
- Weir MW (1964) Developmental changes in problem-solving strategies. *Psychol Rev* 71:473–490
- West RF, Stanovich KE (2003) Is probability matching smart? Associations between probabilistic choices and cognitive ability. *Mem Cogn* 31:243–251
- Whittlesea BWA (1993) Illusions of familiarity. *J Exp Psychol Learn Mem Cogn* 19:1235–1253
- Wolford G, Newman S, Miller MB, Wig G (2004) Searching for patterns in random sequences. *Can J Exp Psychol* 58:221–228
- Zajonc RB (1968) Attitudinal effects of mere exposures. *J Person Soc Psychol* 9:1–27