

Five Principles for Studying People's Use of Heuristics

Julian N. Marewski, Lael J. Schooler and Gerd Gigerenzer

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

Abstract: The fast and frugal heuristics framework assumes that people rely on an adaptive toolbox of simple decision strategies—called heuristics—to make inferences, choices, estimations, and other decisions. Each of these heuristics is tuned to regularities in the structure of the task environment and each is capable of exploiting the ways in which basic cognitive capacities work. In doing so, heuristics enable adaptive behavior. In this article, we give an overview of the framework and formulate five principles that should guide the study of people's adaptive toolbox. We emphasize that models of heuristics should be (i) precisely defined; (ii) tested comparatively; (iii) studied in line with theories of strategy selection; (iv) evaluated by how well they predict new data; and (v) tested in the real world in addition to the laboratory.

Key words: fast and frugal heuristics; experimental design; model testing

As we write this article, international financial markets are in turmoil. Large banks are going bankrupt almost daily. It is a difficult situation for financial decision makers — regardless of whether they are lay investors trying to make small-scale profits here and there or professionals employed by the finance industry. To safeguard their investments, these decision makers need to be able to foresee uncertain future economic developments, such as which investments are likely to be the safest and which companies are likely to crash next. In times of rapid waves of potentially devastating financial crashes, these informed bets must often be made quickly, with little time for extensive information search or computationally demanding calculations of likely future returns. Lay stock traders in particular have to trust the contents of their memories, relying on incomplete, imperfect

knowledge and readily accessible facts, for example, from a news ticker.

Humans are not omniscient. We do not come equipped with the ability to run computationally demanding calculations quickly in our heads. Rather, we make decisions under the constraints of limited information processing capacity, knowledge, and time — be they about the likely performance of stocks, which movie to see at the multiplex, whom to match up with in a speed-dating session, or whether to hospitalize a patient who has registered at the emergency room reception. According to the *fast and frugal heuristics research program* (for recent reviews, see Gigerenzer & Brighton, 2009; Marewski, Gaissmaier, & Gigerenzer, in press a), humans can nevertheless make such decisions successfully because they can rely on a repertoire of simple decision strategies, or *heuristics*. These simple rules of thumb can often perform well even under the constraints of limited knowledge, time, and information-processing capacity. They do so by exploiting the structure of information in the environment in which a decision maker acts and by building on the ways cognitive capacities work, such as the speed with which the human memory system retrieves information. Together, these simple rules of thumb form an *adaptive toolbox* of

Received date: 2008-12-18

We would like to thank Henry Brighton, Henrik Olsson for helpful comments. We would like to thank Julia Schooler and Rona Unrau for editing the manuscript.

Send correspondence to: Julian N. Marewski Center for Adaptive Behavior & Cognition Max Planck Institute for Human Development Lentzeallee 94, 14195 Berlin, Germany Tel: ++49-30-82406302 Fax: ++49-30-82406394 E-mail: marewski@mpib-berlin.mpg.de

the cognitive system, where the tools are heuristics a decision maker uses to respond adaptively to different decision situations, each one appropriate for a given task. In this article, we will first give a short overview of the fast and frugal heuristics framework and then formulate five principles that can help when investigating people's use of heuristics.

Ecologically Rational Heuristics

Which stocks to invest in, which movies to watch, whom to court, what to eat, which car to buy, which newspaper to read — our days are filled with decisions, yet how do we make them? The answer to this question depends on one's view of human rationality because this, in turn, determines what models of cognitive processes one believes to represent people's decision strategies¹. There are at least two major approaches.

Visions of Rationality

Unbounded rationality. The study of *unbounded rationality* asks how people would behave if they were omniscient and omnipotent, that is, if they could compute the future from what they know. The maximization of subjective expected utility is one proposal (e.g., Edwards, 1954). When judging, for instance, in which stocks to invest, such models assume that decision makers behave as if they collect and evaluate all information, weight each piece of it according to some criterion, and then combine the pieces to reach the mathematically optimal solution to

maximize the chance of attaining their goals (e.g., profit maximization). Typically, unbounded rationality models assume unlimited time to search for information, unlimited knowledge, and great computational power (i.e., information-processing capacity) to run complex calculations and compute mathematically optimal solutions. These models are common in economics and optimal foraging theory.

Bounded rationality. According to the second approach, unbounded rationality models provide unrealistic descriptions of how people make decisions, given our limited time, knowledge, and computational power. Herbert Simon (1956, 1990), the father of this *bounded rationality* view, argued that people rely on simple strategies to deal with situations where resources are sparse. One research program that is often associated with Simon's work is the *heuristics-and-biases framework* (e.g., Kahneman, Slovic, & Tversky, 1982), which proposes that humans rely on rules of thumb, or heuristics, as cognitive shortcuts to make decisions². It evaluates human decision making on the basis of principles of probability and logic. According to the heuristics-and-biases tradition, judgments deviating from these normative yardsticks are explained by people's reliance on heuristics that lead to systematic cognitive biases: "Our research attempted to obtain a map of bounded rationality, by exploring the systematic biases that separate the beliefs that people have and the choices they make from the optimal beliefs and choices assumed in rational-agent models" (Kahneman, 2003, p. 1449). In this tradition, the term *bounded rationality* refers to errors, biases, and judgmental fallacies (for a discussion of the "irrationality" rhetoric of the heuristics-and-biases tradition, see Lopes, 1991).

However, Simon (e.g., 1990) not only stressed the cognitive limitations of humans and proposed simple decision-making strategies but also emphasized how the strategies are adapted to our decision-making environment: "Human rational

¹ In the broadest sense, a model is a simplified representation of the world that aims to explain observed data, and/or to predict new data. The term model is typically used for a formal, as opposed to a verbal, instantiation of a theory that specifies the theory's predictions, for example, in mathematical equations or computer code. This category also includes statistical tools, such as structural equation or regression models. Unless one believes that the mind works like a regression analysis or other statistical procedure, such tools are not typically meant to model the workings of psychological mechanisms, say, those determining *how* a person processes information (but see Gigerenzer, 1991, for examples of theories inspired by statistical tools). Here, we mainly discuss *algorithmic-level models* (Marr, 1982), that is, formal instantiations of theories that model psychological processes.

² Kahneman et al. (1982) credited Simon in the preface to the anthology. Lopes (1992) points out that their major early papers, which appear in the anthology, do not cite Simon's work on bounded rationality and that this connection was possibly made in hindsight.

behavior ... is shaped by a scissors whose two blades are the structure of task environments and the computational capabilities of the actor" (p. 7). The fast and frugal heuristics research program has taken up this ecological emphasis. In its framework, *bounded rationality* stands for the notion that by exploiting the structure of information available in the environment, heuristics can lead to good decisions even in the face of limited knowledge, computational power, or time. This approach thus shares with the heuristics-and-biases program the idea that people rely on simple cognitive strategies to make decisions, but the yardstick for reasonable decisions is not logical but ecological. The study of ecological rationality asks the question, in which environment is a given strategy successful (with respect to a defined criterion such as accuracy or speed of decision), and where will it fail. Hammond (1996) distinguished these *correspondence criteria* from *coherence criteria*, which take the laws of logic or probability theory as general normative yardsticks for rationality.

Three Questions About Heuristics: What Heuristics Are Used, Where Should They Be Used, and How Can They Help?

Research in the fast and frugal heuristics program focuses on three interrelated questions. The first is descriptive and concerns the adaptive toolbox: What heuristics do people use to make decisions and when do they rely on a particular heuristic from the toolbox; that is, when and how are different decision strategies from the repertoire selected? 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, for example, by allowing us to make accurate, fast, and effortless decisions? In contrast to these two theoretical questions, the third one focuses on practical applications: How can the study of people's repertoire of heuristics and their fit to environmental structure aid decision making in the real world?

Ecologically rational heuristics are studied in diverse domains. These include applied ones such as medicine (Wegewarth, Gaissmaier, & Gigerenzer, 2009), where heuristics can inform diagnosis, treatment prescription, and risk communication (e.g., Fischer et al., 2002; Marewski, Galesic, & Gigerenzer, 2009; Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz & Woloshin, 2007; Green & Mehr, 1997), and the library sciences, where heuristics can aid literature search (Cokely, Schooler, & Gigerenzer, in press; Lee, Loughlin, & Lundberg, 2002; Van Maanen & Marewski, 2009). In basic research, the fast and frugal heuristics program has proposed a range of heuristics for different tasks, such as mate search (Todd & Miller, 1999), parental investment (Hertwig, Davis, & Sulloway, 2002), and moral judgment (Coenen & Marewski, 2009; Marewski & Kroll, in press). Moreover, as we will explain in detail below, it has produced a large amount of research investigating whether people rely on given heuristics (e.g., Bröder & Schiffer, 2003; Göckner & Betsch, 2008), the environmental structures the heuristics perform well under (e.g., Hogarth & Karelaia, 2007; Katsikopoulos & Martignon, 2006), and how accurate they are for predicting quantities and events in the real world, such as the value of stocks (DeMiguel, Garlappi, & Uppal, 2009; Ortmann, Gigerenzer, Borges, & Goldstein, 2008), or which political parties voters will favor in political elections (Marewski, Gaissmaier, Schooler, Goldstein, & Gigerenzer, 2009). In what follows, we will discuss how the two theoretical questions—when is each heuristic used, and to which environmental structure is each adapted—have been explored for three heuristics. These are the *recognition heuristic*, the *fluency heuristic*, and the *take-the-best heuristic*.

Recognition Heuristic

Which car brand is of better quality, a BMW or a Fiat? Suppose you have heard of the German car manufacturer BMW before reading this article, but you have never heard of Fiat, an Italian car company. In this case, you could use the

recognition heuristic (Goldstein & Gigerenzer, 2002) to respond: You would choose the BMW, that is, the alternative you recognize. In its simplest form, this heuristic is designed for inferring which of two alternatives— one recognized and the other not—has a larger value on a quantitative criterion. When recognition correlates strongly with the criteria on which alternatives are evaluated, the heuristic is defined as follows.

If one alternative is recognized but not the other, infer that the recognized alternative has a larger value on the criterion.

When is it ecologically rational to rely on the recognition heuristic? The recognition heuristic is a specialized tool: Using it will only 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 & Gigerenzer, 2002), and in many competitive domains such as predicting which university is better (Hertwig & Todd, 2003). One reason why alternatives with larger criterion values are more often recognized is that they are more frequently mentioned in the environment: The BBC, CNN, The New York Times, and other *environmental mediators* make it probable that we will encounter and recognize alternatives with large criterion values. Figure 1 illustrates the ecological rationality of the recognition heuristic in terms of three correlations.

As shown, 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 the ecological or surrogate correlation or both are zero,

using the recognition heuristic is not ecologically rational.

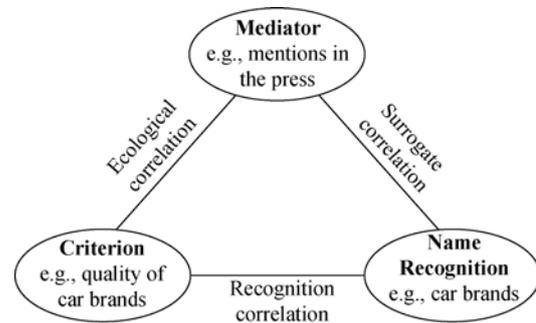


Figure 1. The ecological rationality of the recognition heuristic.

When do people rely on the recognition heuristic? When the correlation between their recognition of alternatives and the criterion is substantial, people tend to make inferences in accordance with the recognition heuristic (e.g., Hertwig, Herzog, Schooler, & Reimer, 2008). In contrast, when the correlation is 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 appeared 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., Hilbig, 2008; Hilbig, Pohl, & Bröder, in press; Marewski et al., 2009, in press; Newell & Fernandez, 2006; Oppenheimer, 2003; Pachur & Hertwig, 2006; Volz et al., 2006).

Fluency Heuristic

The recognition heuristic operates on a binary representation of recognition: An alternative is simply either recognized or it is unrecognized. But this heuristic essentially discards information that could be useful when two alternatives are

recognized but one is recognized more quickly than the other. A strategy that exploits such differences in retrieval speed is the fluency heuristic. This simple rule of thumb has been defined in various ways (e.g., by Jacoby & Brooks, 1984; Whittlesea, 1993). Here we use the term to refer to Schooler and Hertwig's (2005) model of it, which builds on these earlier definitions, a long research tradition on fluency (e.g., Jacoby & Dallas, 1981), and related notions such as *accessibility* (e.g., Bruner, 1957), *availability* (Tversky & Kahneman, 1973), and *familiarity* (e.g., Hintzman, 1988). The fluency heuristic reads as follows:

If one of two recognized alternatives is retrieved more quickly, then infer that this alternative has the higher value with respect to the criterion.

When is it ecologically rational to rely on the fluency heuristic? Retrieval time of an alternative largely depends on a person's history of past encounters with it. Roughly speaking, the more often and the more recently an alternative — say, the name of a company — has been encountered, the more quickly it will be retrieved. Using the fluency heuristic is only ecologically rational when the pattern of encounters with alternatives, and consequently, their retrieval time, correlates with the alternatives' values on a given criterion. Just like in the case of the recognition heuristic, environmental mediators can create such correlations by making it more likely we will encounter alternatives that have larger values on the criterion. And 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 musicians (Hertwig et al., 2008), countries' gross domestic product and companies' market capitalization (Marewski & Schooler, 2009), and the size of cities (Schooler & Hertwig, 2005).

When do people rely on the fluency heuristic?

A number of mechanisms might guide people's use of the fluency heuristic (see Hertwig et al., 2008). Most recently, in a series of experimental and

computer simulation studies, Marewski and Schooler (2009) showed that this heuristic is most likely relied on (i) when using it is likely to help a person make accurate inferences, and (ii) when a person has little or no knowledge about the alternatives in question.

Betting on one Good Cue: The Take-the-best Heuristic

Whereas the fluency and the recognition heuristic rely on retrieval fluency and recognition, other strategies use knowledge about alternatives' attributes as *cues* to make 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 potential attribute to consider is 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 model, which belongs to the family of lexicographic strategies. 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. The heuristic can be described in terms of three rules: one for searching

information, one for stopping this search, and one for making a decision:

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: Choose the alternative that this cue favors.

When is it ecologically rational to rely on take-the-best? When evaluating the performance of parameterized models such as take-the-best, one has to distinguish between two situations. Here, we use the term *predicting new data* (or prediction) to refer to situations in which a model's free parameters are fixed so that they cannot adjust to the data on which the model is tested. In contrast, we use the term *fitting existing data* to refer to situations in which a model's parameters are allowed to adapt to the test data. For each of these two situations, a few general principles have been identified that allow to understand when relying on take-the-best will be successful and when it will fail (see Gigerenzer & Brighton, 2009, for a recent overview).

In fitting existing data, take-the-best will be as successful 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 subsequent cues (Martignon & Hoffrage, 1999). If this is the case, one calls the cue weights *noncompensatory*. For example, a cue weight of 1 would not be compensated if the weight of each subsequent cue were always half the weight of each previous one, which would be the case with weights 1/2, 1/4, 1/8, and so forth.

In predicting new data, take-the-best is often as good or better than models that take into account more information if the cues are highly intercorrelated with each other (Dieckmann & Rieskamp, 2007), that is, if there is much redundancy in the environment. To illustrate, imagine the extreme case of cue intercorrelations of 1. All cues then carry identical information, and

looking at only one cue or looking at all of them makes no difference. In environments with high redundancy and cue-criterion correlations, the biased way take-the-best orders cues — by simple validity rather than by *conditional validity*, that is, by ignoring rather than taking account of the dependencies between cues — has been shown to lead to more accurate predictions. When redundancy and cue-criterion correlations, however, are low, the opposite result is obtained and, effort aside, the complex estimation of dependencies would pay (Brighton, 2009; Gigerenzer & Brighton, 2009).

When do people use take-the-best? Numerous experiments have been conducted that investigate people's reliance on this simple decision strategy (e.g., Bergert & Nosofsky, 2007; Bröder & Schiffer, 2003, 2006; Glöckner & Betsch, 2008; Mata, Schooler, & Rieskamp, 2007; Newell & Shanks, 2003; Rieskamp & Otto, 2006; see Bröder, in press, for an overview). In general, people tend to make inferences consistent with take-the-best when applying it is ecologically rational, for instance, when this heuristic is easier to execute than other strategies. Several studies have suggested that retrieving cue information from memory (as opposed to reading it on a computer screen) fosters people's reliance on take-the-best, especially when cues are represented verbally and when working memory load is high (see Bröder & Gaissmaier, 2007).

In short, in order to learn when people use a given heuristic, one needs to find out when applying this heuristic is ecologically rational. In what follows, we will dig deeper into the methodology of studying people's reliance on heuristics. We will consider five closely interrelated methodological points. These *principles*, as we will call them, are not meant to represent an exhaustive checklist for conducting sound research on heuristics. Corresponding empirical work may also make use of other approaches than the ones described here. The principles we present next highlight that models of heuristics should be (i) precisely defined; (ii) tested comparatively; (iii)

studied in line with theories of strategy selection; (iv) evaluated by how well they predict new data; and (vi) tested in the real world or informed by models of the world.

First Principle: Build Precise Models of Heuristics

The fast and frugal heuristics research program emphasizes specifying precise formal models of heuristics that can be submitted to vigorous testing. For instance, in a two-alternative choice situation, say, whether to read this paper or another one, the computer code or mathematical equations formally specifying a model of a heuristic decision strategy should predict both which alternative will be chosen and how different reasons to choose one alternative over the other will be processed in order to derive a decision. A number of research programs in judgment and decision making and other disciplines take a similar approach, precisely specifying formal models of behavior (e.g., Anderson & Lebiere, 1998; Busemeyer & Myung, 1992; Dougherty, Gettys, & Ogden, 1999; Juslin & Persson, 2002; Payne, Bettman, & Johnson, 1993; Raaijmakers & Shiffrin, 1981; Ratcliff, Van Zandt, & McKoon, 1999; Rumelhart, McClelland, & the PDP Research Group, 1986).

In the fast and frugal heuristics framework, a model of a heuristic specifies (i) process rules such as search, stopping, and decision rules; (ii) the kinds of problems the heuristic can solve, that is, the structures of environments in which it is successful; and (iii) the capacities that the heuristic exploits. The latter two are Simon's above-mentioned two blades.

Corresponding formal models of heuristics need to be distinguished from broad notions (see Gigerenzer, 1996, 2000; Gigerenzer & Murray, 1987; Gigerenzer & Regier, 1996; Marewski, Gaissmaier, & Gigerenzer, in press b). Consider the *representativeness heuristic* (Kahneman & Tversky, 1972), which has been proposed in the aforementioned heuristics-and-biases program. A probability assessed by this decision strategy, such

as whether a newly encountered animal is a dog, is derived from how representative this animal is of the target category—in this case, dogs. However, exactly *how* the category is represented or *how* representativeness is derived was not precisely defined when the heuristic was proposed. This made it possible to successfully apply the broad notion of representativeness to a wide range of phenomena, such as misperception of regression, the conjunction fallacy, and base-rate neglect. At the same time, the lack of specification made representativeness difficult — if not impossible — to test. For instance, as pointed out by Ayton and Fischer (2004), the heuristic has even been invoked to explain opposing events (e.g., A and $\neg A$). After Kahneman and Frederick (2002) made the definition of the heuristic more precise, a number of studies found that models assuming different psychological processes outperform it in predicting people's behavior (e.g., Nilsson, Olsson, & Juslin, 2005).

In short, models of heuristics should enable making precise predictions. They can then provide strong bridges between theories and empirical evidence. We believe that this principle of developing precise models should not only guide research on people's adaptive toolbox of heuristics, but also psychological theorizing and testing in general.

Second Principle: Test Heuristics Comparatively

Once heuristics have been developed into precise models, they should be tested comparatively. In what follows, we will explain why. We will argue that comparative model tests (i) lead to the identification of better models of behavior, and (ii) provide a yardstick for model evaluation.

First, research on heuristics should not be about testing just one model in isolation, proclaiming whether it fits the data or not, as has been done with the recognition heuristic on numerous occasions (e.g., Bröder & Eichler, 2006; Goldstein & Gigerenzer, 2002; Newell &

Fernandez, 2006; Pachur et al., 2008; Pohl, 2006; Richter & Späth, 2006). Rather, research should be about identifying better models of behavior than those that already exist, aiding scientific progress in developing the theory of the adaptive toolbox. For instance, comparative model tests may show that a newly proposed heuristic is a better model than another one, which in turn might add a new model to the adaptive toolbox.

Second, formal model comparisons establish yardsticks for evaluating the descriptive adequacy of competing models, with the models being each other's benchmarks in model evaluation. When just one model is tested, a seemingly large discrepancy between the model's predictions and the observed data might lead a researcher to reject that model. In contrast, with a comparison, the researcher may find that all models suffer, enabling her to find out which model suffers least. At times, sources of variation which affect all models, such as memory variables, may actually be of theoretical interest. This point is illustrated by a set of model comparisons reported by Marewski et al. (in press), which show what dramatically different conclusions one can make from experimental results, depending on whether alternative models are formally specified and tested or merely described without proper comparative testing. Previous findings that people do not always make decisions consistent with the recognition heuristic not only raised doubts about the adequacy of this heuristic as a model of behavior, but were also used to propose that people rely on compensatory strategies instead (e.g., Pohl, 2006; Richter & Späth, 2006). Yet, no study—including any we coauthored (e.g., Goldstein & Gigerenzer, 2002; Pachur et al., 2008; Volz et al., 2006)—tested a corresponding alternative model against the heuristic (but see Pachur & Biele, 2007). Instead, the authors of previous work only provided verbally formulated alternative hypotheses of how people might make their decisions if they did not use the recognition heuristic. While Marewski et al. (in press) were able to replicate several of the previous findings, namely, that the heuristic does

not always predict people's decisions, they also showed that for most people, it predicted behavior better than did each of six alternative models that implemented some of the verbal alternative hypotheses. In doing so, Marewski et al. provided evidence that memory variables such as the strength of the recognition signal are responsible for systematic variations in the frequency of inferences consistent with the recognition heuristic, pointing to mechanisms of heuristic choice rather than to shortcomings in the descriptive adequacy of the heuristic.

In short, the comparative study of heuristics can aid in identifying better models of behavior than those that already exist and establish criteria for evaluating the descriptive adequacy of competing models. In our view, the principle of testing psychological models comparatively represents a good research strategy in general.

Third Principle: Conduct Comparative Model Tests Guided by Theories of Strategy Selection

When discussing above why heuristics should be tested comparatively, we mentioned that such tests led researchers to identify mechanisms that may guide people's use of the recognition heuristic. As we will point out next, the identification of mechanisms of heuristic choice should not only be seen as a by-product of comparative model testing. Rather, comparative tests should actually be informed by theories of heuristic use — or, to use a more general term — by theories of strategy selection (e.g., Cooper, 2000; Feeney, 2000; Gigerenzer, Hoffrage, & Goldstein, 2008; Glöckner, Betsch, & Schindler, in press; Luce, 2000; Marewski, in press). Without such a theory, rejecting a model of a heuristic simply because it does not predict behavior in a certain situation is problematic. There are at least two potential reasons why a decision strategy does not predict behavior. One is that the strategy is not used because people (or the corresponding selection

mechanisms) *choose* not to use it in a particular situation; an alternative reason is that the decision strategy is *generally* not a good model of behavior.

This point is illustrated by the results of a recent study series on the choice between the fluency heuristic and knowledge-based strategies, such as take-the-best. Using the ACT-R theory of cognition (*adaptive control of thought-rational*; Anderson et al., 2004), Marewski and Schooler (2009) developed a theory of strategy selection, called the *cognitive niche framework*. According to it, the ways in which memory represents the structure of information in the environment constrain the set of strategies a decision maker can choose from, defining for each strategy a “cognitive niche”, that is, a range of situations in which the strategy can be applied. Before Marewski and Schooler build a formal model of the cognitive niche of the fluency heuristic, it was reasonable to assume that this heuristic is equally applicable in all situations, that is, both when no knowledge about alternatives is available and when knowledge can be retrieved (see Hertwig et al., 2008, who did not distinguish between these situations). Comparative model tests in which these situations are not examined separately would have shown that knowledge-based strategies predict people’s decisions systematically better than the fluency heuristic does. Yet, it would have been mistaken to then conclude that the fluency heuristic is not a good model of behavior: As Marewski and Schooler showed in experiments and computer simulations with the ACT-R memory model, the interplay between memory and the environment constrains the choice between the this heuristic and knowledge-based strategies such that the fluency heuristic can most likely be relied upon when knowledge is sparse or unavailable, representing an instance of strategy selection.

In fact, assessments of people’s reliance on different heuristics have progressed as research has shifted from asking questions such as whether people use *one* heuristic in *all* situations to testing heuristics comparatively, examining *when* a given

heuristic might be applied (see Bröder, in press). To illustrate, Bröder (2000) began by asking whether all people use take-the-best in probabilistic inferences. But, as he and others later pointed out, “hypothesis rejections at the group level may throw out the baby with the bath water if individual strategy differences are not taken into account” (Pachur, Bröder, Marewski, 2008, p. 204). Indeed, almost all studies on take-the-best have suggested that the proportion of participants who rely on it depends on the characteristics of the decision task. In the light of such findings, the focus of research on take-the-best has shifted toward explorations of variables that might guide strategy use (e.g., Bröder & Schiffer, 2003, 2006; Rieskamp & Otto, 2006).

On a related note, if one assumes that people select from a repertoire of strategies, it is useful to examine the descriptive adequacy of a given heuristic in comparison to others, guided by theories of strategy selection, because it is possible that several strategies might be *equally* able to help a person to behave adaptively. For example, in situations of flat maxima, take-the-best and other strategies could result in equally accurate, effortless, and fast decisions so that it matters little which heuristic a decision maker chooses to employ — applying any of them would yield ecologically rational decisions. People’s use of different decision strategies possibly shows the greatest variability in such situations, because the mechanisms that can otherwise systematically influence strategy selection in the majority of individuals might not be at work. It might be worth speculating whether in such situations social practices, individual preferences, habits, or even personality dispositions channel people’s choice of different decision strategies, giving rise to large individual differences in strategy use.

In short, the comparative testing of heuristics should be informed by theories of strategy selection. If they come accompanied by such theories, comparative tests of heuristics can not only help evaluating the descriptive adequacy of models of heuristics but they can also enhance our

understanding of the ecological rationality of a decision maker's strategy choices. We believe that the principle of testing models informed by theories of strategy selection should not only guide research on heuristics, but it may also be helpful when working from related theoretical perspectives that assume people to come equipped with a repertoire of psychological mechanisms.

Fourth Principle: Examine how Well Models of Heuristics Predict New Data

We have pointed out that the heuristics in the adaptive toolbox should be tested comparatively. In what follows, we will explain *how*. Consider two models of heuristics that compete as explanations for a behavior in a task. How can one decide which model provides a better explanation for the data? This comparison of alternative models is called *model selection*. Model selection can have various technical meanings in different fields, but for our purposes it suffices to say that it is the task of choosing a model from a set of potential models, given available data.

A number of model selection criteria are available (see, e.g., Jacobs and Grainger, 1994, for a detailed overview). These include falsifiability, that is, whether the models can be proven wrong, and the number of assumptions the models make. For instance, one could ask which of many competing models accounts for the data in the simplest way. Other criteria address standards for psychological plausibility, such as whether the computations postulated by a model are tractable in the world beyond the laboratory (e.g., Gigerenzer et al., 2008). Moreover, one could also ask whether a model is consistent with overarching theories of cognition (see, e.g., Dougherty, Franco-Watkins, & Thomas, 2008). Integrative architectures such as ACT-R can impose precise theoretical constraints on which models represent acceptable developments of a theory. Possibly the most widely used model selection criterion is *descriptive adequacy* — which is the yardstick for model selection we will focus on in the remainder of this section. Descriptive

adequacy is often evaluated in terms of *goodness of fit*, that is, when two or more models are compared, the model that provides the smallest deviation from existing data — measured, for example, in terms of R^2 — is favored over a model that results in a larger deviation from that data. Yet, there is a limitation of model selection procedures that are based exclusively on such measures of fit.

The Problem of Overfitting

To conclude that one model provides a better account of data than another based on R^2 or other standard goodness of fit indices might be reasonable if psychological measurements were free of noise. However, noise-free data are practically impossible to obtain. Hence, researchers are confronted with the problem of disentangling the variation in data caused by noise and the variation caused by psychological mechanisms. Goodness of fit measures alone cannot make this distinction. As a result, a model can end up *overfitting* the data; that is, it can capture not only the variance caused by the cognitive process under investigation but also that caused by random error. Figure 2 illustrates a situation in which one model, here Model A, overfits existing data by chasing after idiosyncrasies in that data. This model fits the existing data perfectly but does a poor job of predicting new data. Model B, albeit not as good at fitting the existing data, 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.

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 cognitive process, rather than fitting 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*, that is, a model's inherent

flexibility that enables it to fit diverse patterns of data (see Pitt et al., 2002). Among the factors that contribute to a model's complexity are (i) the number of free parameters it has, (ii) how the parameters are combined in it, (iii) and the extension of the allowable parameter space.

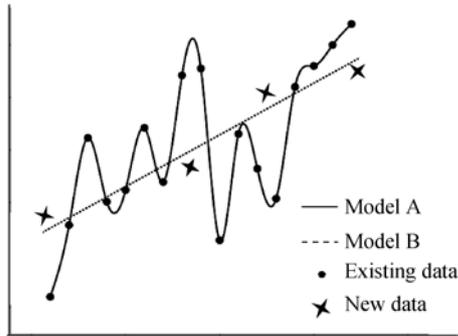


Figure 2. Schematic illustration of how two models fit existing data and how they predict new data. Model A overfits the existing data and is not as accurate as Model B in predicting new data (see Pitt, Myung, & Zhang, 2002).

The impact of the number of free parameters can be illustrated by comparing two regression models. A first model with the cues (i.e., the predictor variables), x_i , might read: $y = w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + w_5x_5 + z$, where the weights, w_i , as well as the constant, z , are free parameters. A second model might look like this: $y = w_1x_1 + w_2x_2 + w_3x_3 + z$. This second model represents a special case of the first model (i.e., with $w_4 = 0$ and $w_5 = 0$) but it is less flexible in fitting existing data than the first model. The impact of the number of free parameters is also shown in Figure 2, where the model that overfits the existing data (Model A) has more free parameters than the model that captures the main tendencies in the new data (Model B).

The impact of how a model's parameters are combined can be explained by comparing Stevens' (1957) and Fechner's (1860/1966) famous models of the relationship between physical dimensions (e.g., the intensity of light, called x here) and their psychological counterparts (e.g., brightness, called y here). In both models, there are two free parameters, a and b , but they have different places in the models' equations (Stevens' model: $y = ax^b$;

Fechner's model: $y = a \ln[x+b]$). Townsend (1975) noted that Stevens' model is more complex than Fechner's model. Since it assumes that a power function relates the psychological and physical dimensions, Stevens' 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 impact of the extension of the parameter space can be illustrated, once more, by comparing two regression models. Both are weighted additive models, in which the cues (i.e., the predictor variables), x_i , are combined linearly, and multiplied by their validities, v_i , and a set of weights, w_i (i.e., $c_1v_1w_1 + c_2v_2w_2 + c_iv_iw_i$, assuming unequal validities). In the first model, let us call it Model C, the weights are chosen such that the prediction of the most valid discriminating cue cannot be overruled by the rest of the cues in the model. For instance, the weights could take the values 1/2, 1/4, 1/8, and so on. Model C is a noncompensatory model. The second model — let us refer to it as Model D — has the same equation and the same number of free parameters as Model C. However, its weights are not constrained; rather they are allowed to take any value. As a result, Model D not only includes the noncompensatory Model C as special case, but it can also fit a larger range of data. Model D is thus more complex than Model C.

The relation between model complexity and generalizability can be summarized in the following way. Increased complexity makes a model more likely to end up overfitting the data while its generalizability to new data decreases. At the same time, a model's generalizability can also increase with the model's complexity—but only to the point at which the model is complex enough to capture systematic variations in the data. Beyond that point, additional complexity can result in decreased generalizability, because the model may then also start to absorb random variations in the data (Pitt et al., 2002).

In short, a good fit to existing data does not

necessarily imply good generalizability to new data, making it important to consider more than standard goodness of fit indices when comparing different heuristics as models of behavior. If descriptive adequacy is relied on as a model selection criterion, heuristics should also be tested using *cross-validation* (Browne, 2000; Stone, 1974, 1977), *Bayesian model selection* (e.g., Myung & Pitt, 1997), *minimum description length (MDL)* (Pitt et al., 2002; see Grünwald, 2007, for a comprehensive treatment of MDL), or other tools to assess the models' generalizability (for overviews of different approaches, see, e.g., Forster, 2000; Marewski & Olsson, 2009; Pitt, Kim, Navarro, & Myung, 2006; Shiffrin, Lee, Kim, & Wagenmakers, 2008). We believe that this principle of examining the generalizability of models should not only guide research on heuristics, but also psychological testing in general.

Fifths Principle: Test Heuristics in the Real World or Guided by Models of the World

The next issue we address also focuses on models' generalizability. This time, however, rather than examining the degree to which a model is capable of predicting all samples generated by the same cognitive process — which is how we had defined the term above — we are concerned with the generalizability of results across different experiments. To ensure generalizability, many psychologists sample representatively from populations of potential study participants. But do they sample stimuli (e.g., car brand names) as carefully as they sample participants? The answer is “not always”. Although there are notable exceptions, for instance in corpus-based memory research (e.g., Anderson & Schooler, 1991; Burgess & Lund, 1997; Griffiths, Steyvers, & Tenenbaum, 2007; Landauer & Dumais, 1997), the sampling of stimuli rarely receives the attention it should (Gigerenzer, 2006; see also, e.g., Clark, 1973; Maher, 1978; Wells & Windschitl, 1999). But for many of the claims about cognitive errors and illusions that have been made in the heuristics-and-

biases tradition, the sampling of stimuli does matter. Research on what has been called the *overconfidence bias* illustrates the point. In experiments, participants are typically given a sample of general knowledge questions, such as “Which city has more inhabitants, Hyderabad or Islamabad?” Participants choose one alternative, such as “Islamabad,” and then give a confidence judgment, such as “70%,” that their answer is correct. Average confidence is usually higher than the proportion correct, which is termed “overconfidence bias” and typically attributed to a cognitive or motivational flaw. How and from what population the knowledge questions are sampled is often not specified in these studies. Yet one can always demonstrate good or bad performance, depending on the items selected: The first researchers who conducted these studies went through almanacs and chose the questions with answers that surprised them (Gigerenzer, 2006). When Gigerenzer, Hoffrage, and Kleinböling (1991) instead introduced random sampling from a defined population (cities in Germany), the “overconfidence bias” largely disappeared in their experiments.

Why is comparatively little attention paid to the sampling of stimuli in certain fields of Psychology? One reason may be that since the publication of Woodworth's (1938) classic textbook *Experimental Psychology*, the methodological dictate in psychology — *systematic design* — has prescribed the isolation and manipulation of a few independent variables whereas all others are kept constant or varied randomly (Dhmi, Hertwig, & Hoffrage, 2004). This has led to the wide acceptance of highly controlled experimental tasks, often entailing only a few impoverished, artificial stimuli that yield a maximum of control, for example, of participant's pre-experimental exposure to the stimuli. Ecological theorizing, however, has motivated strong criticism of this methodology. Brunswik (1955) suggested that it destroys the natural covariation of variables in the organism's habitat,

making it hard to generalize from such experiments to the conditions under which the organism actually performs in its environment. In the real world, people hardly ever interact with only a few impoverished stimuli. People, newspaper ads, or features of cars rarely come in isolated packages; rather, they are often accompanied by a wealth of other information, such as the contexts in which we encounter other humans, read ads, or look at cars. Long before the above mentioned research on overconfidence was conducted (i.e., Gigerenzer et al., 1991), Brunswik lamented that his colleagues followed a double standard by being concerned with sampling participants but not stimuli. His alternative to systematic design — *representative design* — seeks to sample stimuli while preserving their natural covariation and other environmental properties (for a review, see Dhimi et al., 2004). For Brunswik, such representative sampling of stimuli meant random sampling from a defined population. To illustrate, in a classic experiment on size constancy, he walked with the participant through her natural environment and asked her at random intervals to estimate the size of objects she was looking at. As neo-Brunswikian research shows, observations made in the wild can also be used to build formal models that can be tested by predicting new observations. In one such study, Dhimi (2003) observed judges in London courts to examine how punitive decisions are made. Based on her findings she constructed different models of these juridical judgments, which she then validated by predicting new observations and by testing the models against each other. She found that a simple heuristic provided the best formal explanation for the judges' behavior.

Apart from going out into the natural world, one can also try to bring the world into the lab. For instance, by ensuring that a laboratory task reflects the statistical structure of information inherent in the natural environments, one could try turning an experimental task itself into a model of the world. Given that there may often be different ways in which the world can be represented in the lab, this

approach can require deciding between competing models of the world, just as one has to decide between alternative models of the mind.

This is not to say that systematic and more representative designs are mutually exclusive routes to sound research. We believe, in fact, that the two are *complementary* and their use should be tailored to the research question being asked. For example, while priming experiments can disrupt natural correlations (e.g., between different words' retrieval times and the environmental frequency of occurrence of the words in the world), they have proven to be extremely helpful in explaining phenomena such as people's reliance on the fluency heuristic (Hertwig et al., 2008), or the role of ease of processing in affect (e.g., Winkielman, Knutson, Paulus, & Trujillo, 2007).

In short, research on heuristics should not only be concerned with the generalizability of results across participant populations, but also with the generalizability of results across stimuli. In our view, this principle applies beyond research on heuristics — models of human cognition should generally be evaluated in the real world or guided by models of the world.

Summary and Conclusion

We began with an overview of the fast and frugal heuristics framework, an approach to decision making that assumes the mind comes equipped with a repertoire of simple, fast, and frugal decision strategies. These heuristics can lead to good decisions because they can exploit the structure of information in the environment as well as the ways basic cognitive capacities such as memory work. We formulated five principles that may help guide the study of these heuristics. We emphasized that models of heuristics should be (i) precisely defined; (ii) tested comparatively; (iii) studied in line with theories of strategy selection; (iv) evaluated by how well they predict new data; and (v) tested in the real world or informed by models of the world.

In concluding, we would like to highlight that

often a universe of different models of decision strategies exist, all of which are equally capable of reproducing and explaining behavior — a dilemma that is also known as the *identification problem* (see Anderson, 1976). Consequently, it appears unreasonable to ask which of many models of heuristics is more “truthful”; rather, one needs to ask which model is better than another given a set of criteria, for instance, the models’ practical relevance, simplicity, or usability. As Box (1979) puts it — and we agree — “All models are false, but some are useful” (p. 202).

References

- Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*, 1036–1060.
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.
- Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, *2*, 396–408.
- Ayton, P., & Fischer, I. (2004). The hot hand fallacy and the gambler’s fallacy: Two faces of subjective randomness? *Memory & Cognition*, *32*, 1369–1378.
- Bergert, F. B., & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and take-the-best models of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(1), 107–129.
- Box, G. E. P. (1979). Robustness in the strategy of scientific model-building. In R. L. Launer & G. N. Wilkinson (Eds.), *Robustness in statistics* (pp. 201–236). New York: Academic Press.
- Brighton, H. (2009). *Robust cognition and the structure of the environment*. Unpublished manuscript.
- Bröder, A. (2000). Assessing the empirical validity of the “take-the-best” heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1332–1346.
- Bröder, A. (in press). The quest for take the best: Insights and outlooks from experimental research. In P. M. Todd, G. Gigerenzer, & the ABC Research Group, *Ecological rationality: Intelligence in the world*. New York: Oxford University Press.
- Bröder, A., & Eichler, A. (2006). The use of recognition information and additional cues in inferences from memory. *Acta Psychologica*, *121*, 275–284.
- Bröder, A., & Gaissmaier, W. (2007). Sequential processing of cues in memory-based multi-attribute decisions. *Psychonomic Bulletin and Review*, *14*, 895–900.
- Bröder, A., & Schiffer, S. (2003). Take the best versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General*, *132*, 227–293.
- Bröder, A., & Schiffer, S. (2006). Stimulus format and working memory in fast and frugal strategy selection. *Journal of Behavioral Decision Making*, *19*, 361–380.
- Browne, M. W. (2000). Cross-validation methods. *Journal of Mathematical Psychology*, *44*, 108–132.
- Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, *64*, 123–152.
- Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, *62*, 193–217.
- Burgess, C., & Lund, K. (1997). Modelling parsing constraints with high-dimensional context space. *Language and Cognitive Processes*, *12*, 177–210.
- Busemeyer, J. R., & Myung, I. J. (1992). An adaptive approach to human decision making: Learning theory, decision theory, and human performance. *Journal of Experimental Psychology: General*, *121*, 177–184.
- Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335–359.
- Coenen, A., & Marewski, J. N. (2009). Predicting moral judgments of corporate responsibility with formal decision heuristics. In N.A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1524–1528). Austin, TX: Cognitive Science Society.
- Cokely, E. T., Schooler, L. J., & Gigerenzer, G. (in press). Information use for decision making. In M.N. Maack & M.J. Bates (Eds.), *Encyclopedia of Library and Information Sciences*.
- Cooper, R. (2000). Simple heuristics could make us smart: But which heuristics do we apply when? (Open peer commentary). *Behavioral and Brain Sciences*, *23*, 746.
- Dhami, M. K. (2003). Psychological models of professional decision making. *Psychological Science*, *14*, 175–180.
- Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin*, *130*, 959–988.
- Dieckmann, A. & Rieskamp, J. (2007). The influence of information redundancy on probabilistic inferences. *Memory & Cognition*, *35*, 1801–1813.
- Dougherty, M. R. P., Franco-Watkins, A. N., & Thomas, R. (2008). Psychological plausibility of the theory of probabilistic mental models and the fast and frugal heuristics. *Psychological Review*, *115*, 199–213.
- Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). Minerva-DM: A memory processes model for judgments of likelihood. *Psychological Review*, *106*, 180–209.
- DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, *22*, 1915–1953.
- Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, *51*, 380–417.
- Fechner, G. T. (1966). *Elements of psychophysics*. (H. E. Adler, Trans.). New York: Holt, Rinehart and Winston. (Original work published in 1860)
- Feeney, A. (2000). Simple heuristics: From one infinite regress to another? (Open peer commentary). *Behavioral and Brain Sciences*, *23*, 749–750.
- Fischer, J. E., Steiner, F., Zucol, F., Berger, C., Martignon, L., Bossart, W., Altwegg, M., & Nadal, D. (2002). Use of simple heuristics to target macrolide prescription in children with community-acquired pneumonia. *Archives of Pediatrics and Adolescent Medicine*, *156*, 1005–1008.
- Forster, M. R. (2000). Key concepts in model selection: Performance and generalizability. *Journal of Mathematical Psychology*, *44*, 205–231.
- Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review*, *98*, 254–267.
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996). *Psychological Review*, *103*, 592–596.
- Gigerenzer, G. (2006). What’s in a sample? A manual for building cognitive theories. In K. Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition* (pp. 239–260). New York: Cambridge University Press.
- Gigerenzer, G. & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, *1*, 107–143.
- Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest*, *8*, 53–96.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *104*, 650–669.

- Gigerenzer, G., Hoffrage, U., & Goldstein, D. G. (2008). Fast and frugal heuristics are plausible models of cognition: Reply to Dougherty, Franco-Watkins, & Thomas (2008). *Psychological Review*, *115*, 230–239.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*, 506–528.
- Gigerenzer, G., & Murray, D. J. (1987). *Cognition as intuitive statistics*. Hillsdale, NJ: Erlbaum.
- Gigerenzer, G., & Regier, T. (1996). How do we tell an association from a rule? Comment on Sloman (1996). *Psychological Bulletin*, *119*, 23–26.
- Glöckner, A., & Betsch, T. (2008). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1055–1075.
- Glöckner, A., Betsch, T., & Schindler, N. (in press). Coherence shifts in probabilistic inference tasks. *Journal of Behavioral Decision Making*.
- Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*, 75–90.
- Green, L., & Mehr, D. R. (1997). What alters physicians' decisions to admit to the coronary care unit? *The Journal of Family Practice*, *45*, 219–226.
- Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, *114*, 211–244.
- Grünwald, P. D. (2007). *The minimum description length principle*. Cambridge, MA: MIT Press.
- Hammond, K. R. (1996). *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*. Oxford: Oxford University Press.
- Hertwig, R., Davis, J. N., & Sulloway, F. (2002). Parental investment: How an equity motive can produce inequality. *Psychological Bulletin*, *128*, 728–745.
- Hertwig, R., Herzog, S. M., Schooler, L. J., & Reimer, T. (2008). Fluency heuristic: A model of how the mind exploits a by-product of information retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1191–1206.
- Hertwig, R., & Todd, P. M. (2003). More is not always better: The benefits of cognitive limits. In D. Hardman & L. Macchi (Eds.), *Thinking: Psychological perspectives on reasoning, judgment and decision making* (pp. 213–231). Chichester, UK: Wiley.
- Hilbig, B. E. (2008). Individual differences in fast-and-frugal decision making: neuroticism and the recognition heuristic. *Journal of Research in Personality*, *42*, 1641–1645.
- Hilbig, B. E., Pohl, R. F. & Bröder, A. (in press). Criterion knowledge: A moderator of using the recognition heuristic? *Journal of Behavioral Decision Making*.
- Hintzman, D. L. (1988). Judgment of frequency and recognition memory in a multiple trace memory model. *Psychological Review*, *95*, 528–551.
- Hogarth, R. M., & Karelaia, N. (2007). Heuristics and linear models of judgment: Matching rules and environments. *Psychological Review*, *114*, 733–758.
- Jacobs, A. M., & Grainger, J. (1994). Models of visual word recognition. Sampling the state of the art. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1311–1334.
- Jacoby, L. L., & Brooks, L. R. (1984). Nonanalytic cognition: Memory, perception and concept learning. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 18, pp. 1–47). New York: Academic Press.
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*, 306–340.
- Juslin, P., & Persson, M. (2002). PROBABILITIES from EXemplars (PROBEX): A "lazy" algorithm for probabilistic inference from generic knowledge. *Cognitive Science*, *26*, 563–607.
- Kahneman, D. (2003). Maps of bounded rationality. Psychology for behavioral economics. *American Economic Review*, *93*, 1449–1475.
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). New York: Cambridge University Press.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*, 430–454.
- Katsikopoulos, K. V., & Martignon, L. (2006). On the accuracy of lexicographic strategies for pair comparison. *Journal of Mathematical Psychology*, *50*, 116–122.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.
- Lee, M. D., Loughlin, N., & Lundberg, I. B. (2002). Applying one reason decision-making: The prioritisation of literature searches. *Australian Journal of Psychology*, *54*, 137–143.
- Lopes, L. L. (1991). The rhetoric of irrationality. *Theory & Psychology*, *1*, 65–82.
- Lopes, L. L. (1992). Three misleading assumptions in the customary rhetoric of the bias literature. *Theory & Psychology*, *2*, 231–236.
- Luce, D. R. (2000). Fast, frugal, and surprisingly accurate heuristics. (Open peer commentary). *Behavioral and Brain Sciences*, *23*, 757–758.
- Maher, B. A. (1978). Stimulus sampling in clinical research: Representative design reviewed. *Journal of Consulting and Clinical Psychology*, *46*, 643–647.
- Marewski, J. N. (in press). On the theoretical precision, and strategy selection problem of a single-strategy approach: A comment on Glöckner, Betsch, and Schindler. *Journal of Behavioral Decision Making*.
- Marewski, J. N., Gaissmaier, W., & Gigerenzer, G. (in press a). Good judgments do not require complex cognition. *Cognitive Processing*.
- Marewski, J. N., Gaissmaier, W., & Gigerenzer, G. (in press b). We Favor Formal Models of Heuristics Rather than Yin Yang Lists of Dichotomies: A Reply to Evans and Over. *Cognitive Processing*.
- Marewski, J. N., Gaissmaier, W., Schooler, L. J., Goldstein, D. G., & Gigerenzer, G. (2009). Do Voters Use Episodic Knowledge to Rely on Recognition? In N.A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2232–2237). Austin, TX: Cognitive Science Society.
- Marewski, J. N., Gaissmaier, W., Schooler, L. J., Goldstein, D. G., & Gigerenzer, G. (in press). From recognition to decisions: extending and testing recognition-based models for multi-alternative inference. *Psychonomic Bulletin and Review*.
- Marewski, J. N., Galesic, M., & Gigerenzer, G. (2009). Fast and frugal media choices. In T. Hartmann (Ed.), *Media choice: A theoretical and empirical overview* (pp. 107–128). New York & London: Routledge.
- Marewski, J. N. & Krol, K. (in press). Models of ecological rationality: Towards studying the heuristics of morality. [Modelle ökologischer Rationalität: Auf dem Weg zu einer Theorie der Moralheuristiken] In M. Iorio & R. Reisenzein (Eds.), *Regel, Norm, Gesetz. Eine interdisziplinäre Bestandsaufnahme [Rules, norms, and laws. An interdisciplinary review]*. Frankfurt/Main, Germany: Peter Lang Verlag.
- Marewski, J. N., & Olsson, H. (2009). Beyond the null ritual: Formal modeling of psychological processes. *Journal of Psychology*, *217*, 49–60.
- Marewski, J. N., & Schooler, L. J. (2009). *How memory aids strategy selection*. Manuscript submitted for publication.
- Marr, D. (1982). *Vision*. New York: Freeman.
- Martignon, L., & Hoffrage, U. (1999). Why does one-reason decision making work? A case study in ecological rationality. In G. Gigerenzer, P. M. Todd, & the ABC Research Group, *Simple heuristics that make us smart* (pp. 119–140). New York: Oxford University Press.
- Mata, R., Schooler, L. J., & Rieskamp, J. (2007). The aging decision

- maker: Cognitive aging and the adaptive selection of decision strategies. *Psychology and Aging*, 22, 796–810.
- Myung, I. J., & Pitt, M. A. (1997). Applying Occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin & Review*, 4, 79–95.
- Newell, B. R., & Fernandez, D. (2006). On the binary quality of recognition and the inconsequentiality of further knowledge: Two critical tests of the recognition heuristic. *Journal of Behavioral Decision Making*, 19, 333–346.
- Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing "one-reason" decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 53–65.
- Nilsson, H., Olsson, H., & Juslin, P. (2005). The cognitive substrate of subjective probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 600–620.
- Oppenheimer, D. M. (2003). Not so fast! (and not so frugal!): Rethinking the recognition heuristic. *Cognition*, 90, B1–B9.
- Ortmann, A., Gigerenzer, G., Borges, B., & Goldstein, D. G. (2008). The recognition heuristic: A fast and frugal way to investment choice? In C. R. Plott & V. L. Smith (Eds.), *Handbook of experimental economics results: Vol. 1* (Handbooks in Economics No. 28) (pp. 993–1003). Amsterdam: North-Holland.
- Pachur, T., & Biele, G. (2007). Forecasting from ignorance: The use and usefulness of recognition in lay predictions of sports events. *Acta Psychologica*, 125, 99–116.
- Pachur, T., Bröder, A., & Marewski, J. N. (2008). The recognition heuristic in memory-based inference: Is recognition a non-compensatory cue? *Journal of Behavioral Decision Making*, 21, 183–210.
- Pachur, T., & Hertwig, R. (2006). On the psychology of the recognition heuristic: Retrieval primacy as a key determinant of its use. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 983–1002.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York: Cambridge University Press.
- Pitt, M. A., Kim, W., Navarro, D. J., & Myung, J. I. (2006). Global model analysis by parameter space partitioning. *Psychological Review*, 113, 57–83.
- Pitt, M. A., Myung, I. J., & Zhang, S. (2002). Toward a method for selecting among computational models for cognition. *Psychological Review*, 109, 472–491.
- Pohl, R. (2006). Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making*, 19, 251–271.
- Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, 88, 93–134.
- Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review*, 106, 261–300.
- Richter, T., & Späth, P. (2006). Recognition is used as one cue among others in judgment and decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 150–162.
- Rieskamp, J., & Otto, P. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135, 207–236.
- Rumelhart, D. E., McClelland, J. L., & the PDP Research Group. (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1). Cambridge, MA: MIT Press.
- Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, 112, 610–628.
- Shiffrin, R. M., Lee, M. D., Kim, W., & Wagenmakers, E.-J. (2008). A survey of model evaluation approaches with a tutorial on hierarchical Bayesian methods. *Cognitive Science*, 32, 1248–1284.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63, 129–138.
- Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, 41, 1–19.
- Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, 64, 153–181.
- Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions (with discussion). *Journal of the Royal Statistical Society, Series B*, 36, 111–147.
- Stone, M. (1977). Asymptotics for and against cross-validation. *Biometrika*, 64, 29–35.
- Todd, P. M., & Miller, G. F. (1999). From pride and prejudice to persuasion: Realistic heuristics for mate search. In G. Gigerenzer, P. M. Todd, & the ABC Research Group, *Simple heuristics that make us smart* (pp. 287–308). New York: Oxford University Press.
- Townsend, J. T. (1975). The mind-body problem revisited. In C. Cheng (Ed.), *Philosophical aspects of the mind-body problem* (pp. 200–218). Honolulu, HI: Honolulu University Press.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207–232.
- Van Maanen, L., & Marewski, J. N. (2009). Recommender systems for literature selection: A competition of decision making and memory models. In N.A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2914–2919). Austin, TX: Cognitive Science Society.
- Volz, K. G., Schooler, L. J., Schubotz, R. I., Raab, M., Gigerenzer, G., & von Cramon, D. Y. (2006). Why you think Milan is larger than Modena: Neural correlates of the recognition heuristic. *Journal of Cognitive Neuroscience*, 18, 1924–1936.
- Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin*, 25, 1115–1125.
- Wegwarth, O., Gaissmaier, W., & Gigerenzer, G. (2009). Smart strategies for doctors and doctors-in-training: Heuristics in medicine. *Medical Education*, 43, 721–728.
- Winkielman, P., Knutson, B., Paulus, M. P., & Trujillo, J. T. (2007). Affective influence on decisions: Moving towards the core mechanisms. *Review of General Psychology*, 11, 179–192.
- Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1235–1253.
- Woodworth, R. (1938). *Experimental psychology*. New York: Holt.

研究人类使用启发式之五原则

Julian N. Marewski, Lael J. Schooler and Gerd Gigerenzer

(Max Planck Institute for Human Development, Berlin, Germany)

摘要 快速节俭启发式框架假定人们使用一套简捷的决策策略——启发式——进行推理、选择、评价及其他决策。这些启发式策略能顺应任务情境结构中的规律，利用人类的基本认知能力。正基于此，启发式成就了适应性行为。本文拟对启发式框架进行回顾，并简要陈述引导研究者研究人类适应性工具箱的五条原则。我们强调，启发式模型应(i)精确界定(ii)对照检验(iii)与策略选择理论相符(iv)能评估其对新资料的预测力(v)能既在实验室又在现实世界中得以检验。

关键词 简捷启发式；实验设计；模型检验

分类号 B842.5; B841

(中文摘要翻译: 汪祚军)