Ulrich Hoffrage, Torsten Reimer[*]

# Models of Bounded Rationality:
# The Approach of Fast and Frugal Heuristics[**]

In a complex and uncertain world, humans draw inferences and make decisions under the constraints of limited knowledge, resources, and time. Herbert Simon, with his call for models of bounded rationality, can be seen as one of the fathers of the recently initiated research program on "simple heuristics that make us smart" (Gigerenzer/Todd/the ABC Research Group, 1999). These heuristics perform well because they are ecologically rational: they explore the structure of environmental information and are adapted to this structure. The present review paper introduces the key concepts of this research tradition, and provides two examples: (1) The recognition heuristic, which exploits a partial lack of knowledge, and (2) Take The Best, a simple lexicographic strategy that deliberately ignores information although it is available. The paper explains their ecological rationality, provides empirical evidence of their use, and illustrates some of their applications in consumer behaviour and group decision making. Finally, this research program is related to various notions of rationality.

Key words:  **Bounded Rationality, Ecological Rationality, Simple Heuristics, Recognition Heuristic, Take The Best**

[*]     Ulrich Hoffrage, Max Planck Institute for Human Development, Berlin, Germany.

        Torsten Reimer, University of Basel, Basel, Switzerland.

        Address correspondence to:
        Ulrich Hoffrage, Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Lentzeallee 94, D – 14195 Berlin, Germany, Phone: ++49 30 82406 273, Fax: ++49 30 82406 394. E-Mail: Hoffrage@mpib-berlin.mpg.de.
        After December 2004 to: Ulrich Hoffrage, Faculty of Economics and Business Administration (Ecoles des Hautes Etudes Commerciales – HEC), University of Lausanne, BFSH1, CH-1015 Lausanne, Switzerland. E-mail: Ulrich.Hoffrage@hec.unil.ch.

In a course on judgment and decision making, one of the present authors confronted the students with the following fictitious scenario: After Thomas P. retired, he took all his savings, went to the roulette table at the nearest casino and placed all his money on the "12." Was this a good decision? The students did not hesitate for a moment to say no. After this they were told the outcome (the "12" won), and they were faced with the question again: Was it a good decision? This time it took them a bit longer to answer, but they did not change their minds. Although the outcome was favourable to Thomas P., his decision to take this risk was not considered to be rational. The lesson was that the rationality of a decision should *not* be evaluated by taking information into account that was *not* available to the decision maker when he made the decision.

The present review paper is centred on a question that appears to be quite similar, but is much more provocative: Can it be rational *not* to use information even when it *is* available? Common intuition says no. Francis Bacon's thesis that *knowledge is power* probably reflects what most of us think about this issue, and it is not accidental that these words are written above the entrance to the *Deutsches Museum* in Munich, a museum that recounts the history of the rise and the success of science and technology after the Renaissance. Having knowledge about the laws of physics or chemistry, for instance, gives one an advantage over those who lack this knowledge. Similarly, it is usually advantageous to have information about one's environment, including information about other people's states of knowledge and their intentions. This is why governments, companies, and individuals spend billions of dollars every year to acquire information. From the assumption that having more knowledge and information is better than having less, it is only a tiny step to the conclusion that using this knowledge and information is better than not using it.

We do not want to question this conclusion in general. However, we want to draw attention to situations in which it is beneficial not to have information in the first place or, if it is available, to deliberately not use it when making decisions. Moreover, we argue that such situations are not as rare as one may think. Note that in what we describe below, not using information is not a goal per se, but the result of using so-called fast and frugal heuristics. The research program on fast and frugal heuristics, often called simple heuristics (Gigerenzer/Todd/the ABC Research Group 1999), has attracted a considerable amount of attention and discussion over the past years (e.g., see the commentaries and the reply following Todd/Gigerenzer 2000). In the present article, we want to give a brief introduction to the core ideas of this program and an overview of some related research. It is structured as follows: In the first part we will explain some concepts fundamental to the approach of fast and frugal heuristics. Specifically, we will start with a short historical remark on the notion of heuristics, present the gaze heuristic as one example of a fast and frugal heuristic, and introduce the notions of bounded rationality and ecological rationality. In the second and third part, we will focus on the recognition heuristic and on Take The Best, respectively. The first exploits a lack of knowledge; the second deliberately ignores information although it is available. For each of these heuristics we specify their ecological rationality, provide empirical evidence for their use, and summarize some of their applications, including an application in a group setting. We conclude by relating the present research program to various notions of rationality.

## The Approach of Fast and Frugal Heuristics

*What is a heuristic?*

*Heuristic* comes from the Greek *heuriskein,* meaning "to find," hence *eureka,* meaning "I found it (out)." Since its introduction to English in the early 1800s, the term has acquired a range of meanings. For instance, in his Nobel prize-winning paper, "On a heuristic point of view concerning the generation and transformation of light," Albert Einstein (1905) used the term to indicate that his view served to find out or to discover something. Such a heuristic view may yield an incomplete and unconfirmed, eventually even false, but nonetheless useful picture. For the Gestalt psychologists who conceptualized thinking as an interaction between external problem structure and inner processes, heuristics (e.g., inspecting the problem and analyzing the conflict, the situation, the materials, and the goal) served the purpose of guiding the search for information in the environment and of restructuring the problem by internal processes (Duncker 1935). In the 1950s and 60s, Herbert Simon and Allen Newell, two pioneers of artificial intelligence and cognitive psychology, used the term to refer to methods for finding solutions to problems. In formalized computer programs (e.g., the General Problem Solver), they implemented heuristics such as the means-end analysis, which tried to set subgoals and find operations that would finally reduce the distance between the current state and the desired goal state (Newell/Simon 1972). With the advent of information theory in cognitive psychology, the term heuristic finally came to mean a useful shortcut, an approximation, or a rule of thumb for searching through a space of possible solutions. In the next section, we provide an example of a heuristic that can be used to solve a quite practical problem: how to catch a ball.

*The gaze heuristic*

Catching balls – balls that come in high, as in baseball, cricket, or soccer – seems to be quite easy, but if you imagine that your task is not to catch them yourself, but rather to build a robot that is able to achieve this goal, you will soon realize how hard it is to design and implement the processes involved. For the sake of simplicity, we consider situations where a ball is already high up in the air and will land in front of or behind the robot. How would you build such a robot? One vision is *omniscience*: Give your robot the most sophisticated computational machinery and a complete representation of its environment. First, you might feed your robot a parabolic equation because, in theory, balls have parabolic trajectories. You might also provide the joint probability distribution of the two parameters that define any parabola, and of course you would select the distribution for this particular game. In order to estimate the right parameter for the present parabola, the robot needs to be equipped with instruments that can measure the ball's initial distance and velocity, and its projection angle. In the real world, however, balls do not fly in parabolas, due to air resistance, wind, and spin. Thus, the robot would need further instruments that can measure the speed and direction of the wind at each point of the ball's flight in order to compute the resulting path and the point where the ball will land, and to then run (or roll) there. All this would have to be completed within a few seconds – the time a ball is in the air.

An alternative vision exists, which does not aim at complete representation and which can succeed without complicated calculation. This vision begins with the question: Is there a smart heuristic that can solve the problem? McLeod and Dienes (1996), who studied experienced players, observed that these players do not stand for some seconds and watch the ball, "compute" where it will come down, and then suddenly start to run. Instead, they start running immediately, while fixating their vision on the ball. The heuristic they use is to adjust the running speed so that the angle of gaze (i.e., between the eye and the ball, relative to the ground) remains constant – or within a certain range (Figure 1). In our thought experiment, a robot that uses this heuristic does not need to measure wind, air resistance, spin, or the other causal variables. It can safely ignore every piece of causal information, because all the relevant information is contained in one variable: the angle of gaze. Note that a robot using the gaze heuristic will not compute the point at which the ball will land. But it will be there.

**Figure 1:** **The gaze heuristic, a fast and frugal heuristic, is used by experienced players to catch a ball: "fixate the ball, start running, and adjust your running speed so that the angle of gaze remains constant"** (McLeod/Dienes 1996; figure adapted from Gigerenzer 2001).
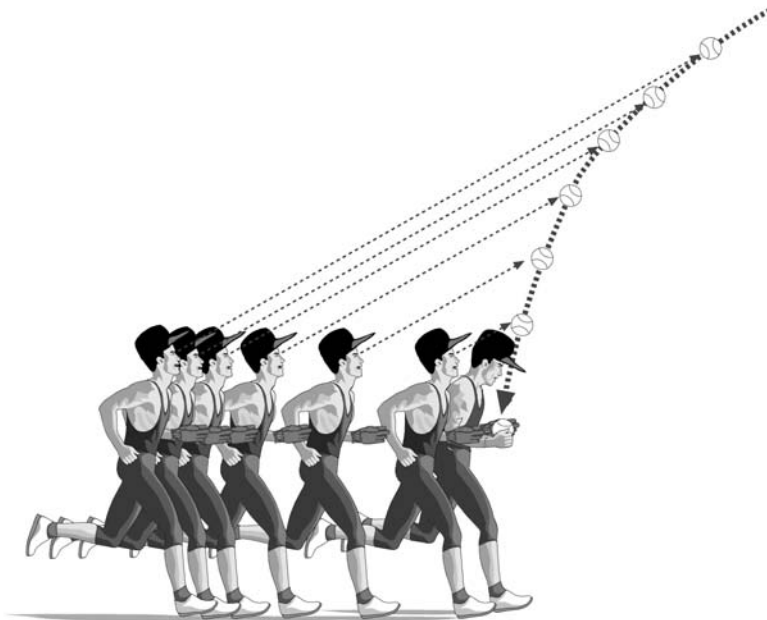


## Bounded Rationality

The gaze heuristic demonstrates some important aspects of many heuristics. First, it does not use all information that is potentially available. For instance, the robot does not have to be equipped with a device to measure wind speed. Second, the heuristic does not integrate variables in a complex way – in the present case there is no integra-

tion at all, and only one variable, the angle, is considered. Third, the success of a strategy depends on where and how it is implemented. To see this, imagine a situation in which catching the ball would only be possible if the robot immediately started running at its full (and thus limited) speed. One can think of situations in which a robot that uses the gaze heuristic would not be able to catch the ball, simply because being constrained to keep the angle constant would prevent him from running at full speed the entire time while the ball is in the air. In contrast, an "omniscient" robot could accomplish this task. This, however, would require the capacity for computing where the ball will land and for realizing that running at full speed is the only way to catch it. An equally important requirement is that the time to perform this calculation should not be too long, while still allowing the robot to arrive at the spot punctually.

If we consider humans, such computational constraints are often important to take into account when modelling behaviours, judgments or decisions. This led Herbert Simon (1947, 1982) to introduce the notion of *bounded rationality*. In contrast to models that aim at finding the optimal solution to a problem at hand, models of bounded rationality recognize that humans often have limited information, time, and computational capacities when making judgments or decisions. Given these constraints, the optimal solution is often unattainable. Moreover, many problems are too complex to solve within a reasonable amount of time, even if all the relevant information is available and the most powerful computers are used. Models of bounded rationality specify the (cognitive) processes that lead to a *satisficing* solution to a given problem, that is, to a solution that is both *satis*fying and suf*ficing*.

The gaze heuristic can be considered a model of bounded rationality. It can be – and, as experimental studies with experienced players have shown, is – used by humans who do not have the capacity to compute precisely where the ball will land after having seen it in flight for a few seconds (McLeod/Dienes 1996). A variant of this heuristic is used by pilots: When another plane is approaching, and a collision seems possible, they can look at a scratch in the windshield and observe whether the other plane moves away from that scratch. If it does, there is no need for concern. However, if the other plane does not move relative to the scratch, it means that it is heading straight for the windshield. For the outfielder, the goal is to produce a collision, whereas for the pilot, the goal is to avoid a collision. The nature of the heuristic is the same – keep the angle constant (outfielder) or avoid constant angles (pilot).

## Fast and Frugal Heuristics

A research program in the spirit of Simon's bounded rationality is the program of *simple heuristics* (Gigerenzer et al. 1999), often also referred to as *fast and frugal heuristics*. These heuristics are task-specific, that is, they are designed to solve a particular task (e.g., choice, numerical estimation, and categorization). They cannot, however, solve tasks that they are not designed for – just as a hammer is ideal for hammering in nails but is useless for sawing a board. In fact, this task-specificity is fundamental to the notion of the *adaptive toolbox* (Gigerenzer/Selten 2001), the collection of heuristics that has evolved and can be used by the human mind.

Although fast and frugal heuristics differ with respect to the problems they have been designed to solve, they share the same guiding construction principles. In par-

ticular, they are composed of *building blocks*, which specify how information, be it stored in memory or externally presented, is searched for (*search rule);* when information search is stopped (*stopping rule);* and how a decision is made based on the information acquired (*decision rule).* Thus, unlike models that assume all information is already known to the decision maker and that are merely used to predict the outcome of the decision making process, fast and frugal heuristics specify the cognitive processes, including those involved in information acquisition (for related programs that explicitly include information search, see Busemeyer/Townsend 1993, and Payne/Bettman/ Johnson 1993).

These heuristics are fast for two reasons. First, they do not integrate the acquired information in a complex and time-consuming way. In this respect, many heuristics of the adaptive toolbox are as simple as possible because they do not combine pieces of information at all; instead, the decision is based on just one single reason (*one-reason decision making*). Second, they are fast as a consequence of being frugal, that is, they stop searching for further information early in the process of information acquisition. The gaze heuristic, for instance, is a fast and frugal heuristic. It is fast because it can solve the problem within a few seconds, and it is frugal because it demands only one piece of information, the angle of gaze.

Studies on fast and frugal heuristics include (a) computer simulations to explore the performance of the heuristics in a given environment, in particular in real-world environments (e.g., Czerlinski/Gigerenzer/Goldstein 1999); (b) the use of mathematical or analytical methods to explore when and why they perform as well as they do (eventually supported by simulations, in particular in artificially created environments in which information structures are systematically varied, e.g., Martignon/Hoffrage 2002); and (c) experimental and observational studies to explore whether and when people actually use these heuristics (e.g., Rieskamp/Hoffrage 1999).

*Ecological Rationality*

Models of ecological rationality describe the structure and the representation of information in actual environments and their match with mental strategies, such as boundedly rational heuristics. In the present paper, we will only deal with the first aspect, namely the fit of heuristics to informational structures of the environment (for the effect of external representation of information on cognitive processes, see Gigerenzer/Hoffrage 1995; Hoffrage et al. 2000). To the degree that such a match between heuristics and informational structures exists, heuristics need not trade accuracy for speed and frugality. The importance of considering the environment when studying the human mind is best illustrated in Simon's analogy of a pair of scissors, with the mind and environment as the two blades: "Human rational behavior is shaped by a scissors whose blades are the structure of task environments and the computational capabilities of the actor" (Simon 1990: 7). If looking only at one blade, one cannot fully understand how the human mind works, just as one cannot understand how scissors with one single blade could function. The fit of a heuristic to the environment in which it is evaluated is an important aspect of fast and frugal heuristics, and has given rise to a series of studies and important insights (Todd/Gigerenzer/the ABC Research Group 2004).

The simultaneous focus on the mind and its environment, past and present, puts research on decision making under uncertainty into an evolutionary and ecological framework, a framework that is missing in most theories of reasoning, both descriptive and normative. From such a perspective it is straightforward to study the adaptation of mental and social strategies to real-world environments rather than compare strategies to the norms of probability theory (e.g., *Bayes's rule*, which can be used to update prior beliefs in the light of new data) and logic (e.g., the *conjunction rule*, according to which the probability that an object belongs both to the classes A and B cannot exceed the probability that it belongs to class A). Rather, the performance of a heuristic is evaluated against a criterion that exists in the environment (see Hammond 1996, for the distinction between internal consistency versus external correspondence). For instance, the QuickEst heuristic (Hertwig/Hoffrage/Martignon 1999) that makes fast and frugal inferences about the numerical values of objects (e.g., number of employees of firms) by sequentially checking cues until a cue with a positive value is found and by estimating the criterion based on what that cue was, is evaluated by comparing estimated and true values.

After this general introduction, we now turn to two heuristics in more detail: the recognition heuristic and Take The Best. We will elucidate their ecological rationality, demonstrate some applications, and report on the empirical evidence that people use these heuristics.

## The Recognition Heuristic

Most people would agree that it is usually better to have more information than to have less. There are, however, situations in which partial ignorance is informative, which the recognition heuristic exploits. Consider the following question: Which city has more inhabitants, San Antonio or San Diego? If you grew up in the U.S., you probably have a considerable amount of knowledge about both cities, and should do far better than chance when comparing the cities with respect to their populations. Indeed, about two-thirds of University of Chicago undergraduates got this question right (Goldstein/Gigerenzer 2002). In contrast, German citizens' knowledge of the two cities is negligible. How much worse will they perform? The amazing answer is that within a German sample of participants, 100% answered the question correctly (Goldstein/Gigerenzer 2002). How could this be? Most Germans might have heard of San Diego, but do not have any specific knowledge about it. Even worse, most have never even heard of San Antonio. However, this difference with respect to name recognition was already sufficient to make an inference, namely that San Diego has more inhabitants. Their lack of knowledge allowed them to use the recognition heuristic, which, in general, says: "If one of two objects is recognized and the other is not, then infer that the recognized object has the higher value with respect to the criterion" (Goldstein/Gigerenzer 2002, p. 76). The Chicago undergraduates could not use this heuristic, because they had heard of both cities – they knew too much.

*Ecological rationality of the recognition heuristic.*

The ecological rationality of the recognition heuristic lies in the positive correlation between criterion and recognition values of cities (if such a correlation were negative,

the inference would necessarily go in the opposite direction). In the present case, the correlation is positive, because larger cities (as compared to smaller cities) are more likely to be mentioned in mediators such as newspapers, which, in turn, increases the likelihood that their names are recognized by a particular person. It should thus be clear that the recognition heuristic only works if recognition is correlated with the criterion. Examples include size of cities, length of rivers, or productivity of authors. In contrast, the heuristic will probably not work when, for instance, cities have to be compared with respect to their mayor's age or their altitude above sea level.

### The dominance effect

Like humans, animals have a capacity for recognition memory, and they tend to rely on the recognition heuristic in situations where recognition is ecologically valid. For instance, when wild Norway rats choose between two diets, one that they recognize from having previously smelt it on a fellow rat's breath and one that is novel, they tend to choose the first one. The evolutionary rationale seems to be that the heuristic helps to avoid food poisoning (Galef 1987). But what if one provides the rat with information that conflicts with what the recognition heuristic suggests? Galef, McQuoid, and Whiskin (1990) again gave "observer" rats a choice between the two diets; this time, however, the fellow rat was ill. The fellow rat was injected with a nauseant, but as far as the observer rat knew, it could have been food poisoning. Which diet would the observer rat choose? The surprising finding was that the observer rats still chose the diet they recognized from their fellow rats' breath, and avoided the diet that they did not recognize. Recognition information dominated illness information.

A similar dominance effect could be demonstrated for humans. In a series of experiments in which participants had to compare cities with respect to their number of inhabitants, it turned out that they chose the object predicted by the recognition heuristic in more than 90% of the cases – the amazing finding was that in another study in which participants were taught knowledge about the cities that contradicted the choices implied by the recognition heuristics, participants' choices followed the recognition heuristic in practically the same percentage of cases (Goldstein/Gigerenzer 2002).

### The less-is-more effect

As the example with the German and American students demonstrates, recognizing less objects can sometimes enhance performance. By means of mathematical analysis, it is possible to specify the conditions in which such a *less-is-more effect* can be obtained. In order to better understand this condition we first have to consider the percentage of correct inferences ($c$) in a complete paired comparison task among the $N$ objects of a given reference class, from which $n$ objects are recognized. This percentage can be calculated as follows (Goldstein/Gigerenzer 2002):

$$c = 2 \left( \frac{n}{N} \right) \left( \frac{N-n}{N-1} \right) \alpha + \left( \frac{n}{N} \right) \left( \frac{n-1}{N-1} \right) \beta + \left( \frac{N-n}{N} \right) \left( \frac{N-n-1}{N-1} \right) \frac{1}{2}$$

The first term on the right side of the equation accounts for the correct inferences made by the recognition heuristic (in front of the *a* is the percentage of pairs in which only one object is recognized), the second term equals the proportion of correct inferences made when knowledge beyond recognition is used (in front of the *β* is the percentage of pairs in which both objects are recognized), and the third accounts for guessing (in front of the ½ is the percentage of pairs in which neither of the two objects is recognized). The *recognition validity a* is the relative frequency of pairs in which the recognized object is the correct answer (out of all pairs in which only one object is recognized), and the *knowledge validity β* is the relative frequency of getting a correct answer when both objects are recognized. Note that all parameters in this equation can be determined without knowing *c*.

A less-is-more effect occurs if *a > β*. If this is the case, then recognizing all objects will not yield the maximum performance, and, as a consequence, one may be better off when recognizing less then all objects. This is simply because the proportion of pairs by which *a* is multiplied is largest if half of the objects are recognized, while recognizing less than all objects (as compared to recognizing all objects) reduces the term by which *β* is multiplied. A less-is-more effect can emerge in at least three different situations: First, between two groups of people, when a more knowledgeable group makes worse inferences than a less knowledgeable group in a given domain; second, between domains, that is, when the same group of people achieve higher accuracy in a domain in which they know little than in a domain in which they know a lot; and third, during knowledge acquisition, that is, when an individual begins to make more erroneous inferences as a result of learning. Goldstein and Gigerenzer (2002) could empirically demonstrate two less-is-more effects, namely of the second and third type just mentioned. Specifically, in one study they found that participants performed better in a domain in which they recognized a lower percentage of objects. For instance, American participants performed better when comparing German cities with respect to their population size than when comparing American cities; for German participants it was the other way round. In another study, they found that performance decreased through successively working on the same questions. With this procedure, recognition of objects increased during the course of the experiment. As a consequence, the proportion of cases that is multiplied by *a* decreased at the expense of cases that are either multiplied by *β* or ½.

## The recognition heuristic in group decision making

Does the dominance effect – prioritizing recognition information over cue-based information – also hold in a group context? What happens if an individual who can use the recognition heuristic and other individuals who recognize both cities have to reach a joint group decision? Reimer and Katsikopoulos (in press) developed and tested various models that can be used by groups who are faced with dichotomous choice tasks. These models were tested in a study in which three-member groups performed the city population task. On the basis of a pilot study, individuals could be categorized as persons who recognized both, one, or neither of the two cities of each inference task. Because it is difficult to "prove" which of two cities is more populated in a group discussion, one reasonable assumption is that a group uses a simple majority rule (see

Gigone/Hastie 1997). However, there are also good reasons to assume that the group decision will be dominated by the "experts" who recognize both cities and have obviously more knowledge on the task than those "ignorant members" who do not even recognize one of the cities; the knowledge-based majority model assumes that members who can use their knowledge will dominate the group decision process. But if recognition plays a special role in the process of making inferences, it may be that groups exploit the lack of knowledge of their semi-ignorant members who recognize only one of the two cities; according to the recognition-based majority model, those who can apply the recognition heuristic will dominate the group decision process.

Overall, the simple majority rule described the data well (Reimer/Katsikopoulos in press). This was particularly the case when knowledge and recognition converged – for example, when members who could use the recognition heuristic agreed with members who recognized both cities. However, when the models made contrasting predictions – that is, when recognition-based criteria conflicted with knowledge-based criteria, as in the situation described earlier – the groups typically adhered to the recognition heuristic. In 65% of the cases where the recognition-based and the knowledge-based models disagreed, the group inference matched the predictions of the recognition-based majority model, which explained 90% of the group inferences when it could be applied.

That group members let their knowledge be dominated by the semi-ignorant members who recognized only one of the two cities may seem odd. But in fact this apparently irrational decision increased the overall accuracy of the group, due to the ecological rationality of the recognition-based majority model. More precisely, this model is functional because the less-is-more effect also holds in groups that exploit the recognition information. The less-is-more effect was also empirically observed in the group study – when two groups had almost identical average $a$ and $\beta$, the group who recognized *fewer* cities (smaller $n$) typically had *more* correct answers. For instance, the members of one group recognized on average only 60% of the cities, and those in a second group 80%; but the first group got 83% answers correct in a series of over 100 questions, whereas the second only 75%. In sum, the groups seemed to apply the recognition heuristic. Exploiting the lack of knowledge of some group members was adaptive in the sense that it improved the group's accuracy. This was possible because the recognition validity was higher than the knowledge validity, which can also account for the counterintuitive less-is-more effect that was observed between groups.

### The recognition heuristic in advertisement

In a world full of unboundedly rational consumers, advertisement would not make a difference – they already have full information and will integrate it such that they maximize their own expected utility. In the real world, however, consumers have imperfect information, including misinformation (Stigler/Becker 1977). Here, advertisement can provide (additional) information and firms are well advised to select this information such that their products sell better. But if providing information is the foremost – if not the only – goal of advertisement, why then would firms pay for advertisement that does not provide any information about their product at all? The Benetton campaign, for instance, presented shocking photos alongside the name of

their company, such as the face of a man who has been "sentenced to death." From seeing those ads one could not even infer what products Benetton sells. However, the ads ensured that the name of Benetton will stick in people's memory. Such noninformative advertisement tries to exploit consumer's reliance on the recognition heuristic when choosing what to buy. Oliviero Toscani (1997), the designer of many Benetton ads, pointed out that the ads had pushed Benetton beyond Chanel into the top-five best-known brand names across the world, and Benetton's sales increased by a factor of 10. Less-known politicians, universities, cities, sport teams, and even small nations go on crusades for a place in the recognition memory of the general public. For still another successful application of the recognition heuristic, namely for picking stocks for a portfolio, see Ortmann et al. (in press).

## Take The Best

If both objects in a pair-comparison task are recognized, the recognition heuristic does not discriminate between them. When comparing Hanover and Heidelberg with respect to their number of inhabitants, Americans who have heard of Heidelberg but not of Hanover can use it, but Germans who recognize both cities cannot. A fast and frugal heuristic that could be used in such a case is Take The Best. For simplicity, it is assumed that all cues (i.e., predictors) are binary (positive or negative), with positive cue values indicating higher criterion values. In the present example, information that could serve as a cue included whether a city has an airport or an opera house, as well as whether the city is a state capital or the site of a major exposition. Take The Best is a simple, lexicographic strategy that consists of the following building blocks:

(0) Recognition heuristic: see above.

(1) Search rule: If both objects are recognized, choose the cue with the highest validity (where validity is defined as the percentage of correct inferences among those pairs of objects in which the cue discriminates) among those that have not yet been considered. Look up the cue values of the two objects.

(2) Stopping rule: If one object has a positive value and the other does not (i.e., has either a negative or unknown value) then stop search and proceed to Step 3. Otherwise return to Step 1 and search for another cue. If no further cue is found, then guess.

(3) Decision rule: Infer that the object with the positive cue value has the higher value on the criterion.

In the Hanover/Heidelberg comparison, Take The Best would start with the most valid cue – provided that both cities are recognized. Assume this is the capital cue (since Berlin is the largest German city, this cue would only make correct inferences and thus have a validity of 1). However, because neither Hanover nor Heidelberg is the capital of Germany, Take The Best will proceed with the next cue. Let us assume this is the exposition site cue (in fact, the cue hierarchy depends on several factors, for instance, on the set of cues under consideration, the set of pairs among which the validity is computed, and the accuracy of knowledge about the cue values). The exposition site cue discriminates: Hanover is an exposition site; Heidelberg is not. At this point, Take The Best stops search and decides in favour of Hanover – no further information is considered.

Note that Take The Best's search rule ignores cue dependencies and therefore will most likely not establish the optimal ordering.[1] Further note that the stopping rule does not attempt to compute an optimal stopping point after which the costs of further search exceed its benefits. Rather, the motto of the heuristic is "Take the best, ignore the rest." Finally note that Take The Best uses "one-reason decision making" because its decision rule does not weight and integrate information, but relies on one cue only.

By varying the building blocks, one can generate "siblings" of Take The Best. For instance, Minimalist is another fast and frugal heuristic that also employs one-reason decision making. It is even simpler than Take The Best because it does not try to order cues by validity, but chooses them in random order. Its stopping rule and its decision rule are the same as those of Take The Best.

### The performance of fast and frugal heuristics.

What price does one-reason decision making have to pay for being fast and frugal? How much more accurate are strategies that use all cues and combine them? To answer these questions, Czerlinski et al. (1999) evaluated the performance of various strategies in 20 data sets containing real-world structures rather than convenient multivariate normal structures; they ranged from having 11 to 395 objects, and from 3 to 19 cues. The predicted criteria included demographic variables, such as mortality rates in U.S. cities and population sizes of German cities; sociological variables, such as drop-out rates in Chicago public high schools; health variables, such as obesity at age 18; economic variables, such as professors' salaries and selling prices of houses; and environmental variables, such as the amount of rainfall, ozone, and oxidants. In the tests, half of the objects from each environment were randomly drawn. From all possible pairs within this training set, the order of cues according to their validities was determined (Minimalist used the training set only to determine whether a positive cue indicates a higher or lower criterion). Thereafter, performance was tested both on the training set (fitting) and on the other half of the objects (generalization). Two linear models were introduced as competitors: multiple regression and a simple unit-weight linear model (Dawes 1979). To determine which of two objects has the higher criterion value, multiple regression estimated the criterion of each object, and the unit-weight model simply added up the number of positive cue values.

---

[1]     Indeed, when Martignon and Hoffrage (2002) used the set of all German cities with more than 100,000 inhabitants (N=83; this was the set that has been used by Gigerenzer and Goldstein, 1996), and computed the performance for each of the 362,880 orderings that could be constructed from nine cues, the ordering based on the validities of those cues was not the best. A lexicographic algorithm that used this ordering (i.e., Take The Best) achieved 74.2% correct inferences, whereas the maximum performance of 75.8% was achieved by another ordering. However, this difference was only 1.6 percentage points, and only 1.8% of the 362,879 remaining possible orderings performed better than the one established by validity. Moreover, the problem of finding this optimal ordering is NP-hard, which implies that there is no simpler way of finding this optimal ordering than by computing and comparing the performances of all possible orderings.

**Table 1:** **Performance of two fast and frugal heuristics (Take The Best and Minimalist) and two linear models (multiple regression and a unit-weight linear model) across 20 data sets.** *Frugality* **denotes the average number of cue values looked up;** *fitting* **and** *generalization* **refer to the performance in the training set and the test set, respectively (see text)** (table adapted from Czerlinski et al. 1999).

| Strategy | Frugality | Accuracy (% Correct) | |
|---|---|---|---|
| | | Fitting | Generalization |
| Take The Best | 2.4 | 75 | 71 |
| Minimalist | 2.2 | 69 | 65 |
| Multiple Regression | 7.7 | 77 | 68 |
| Unit-weight linear model | 7.7 | 73 | 69 |

Table 1 shows the counterintuitive results obtained by averaging across frugality and percentages of correct choices in each of the 20 different prediction problems. The two simple heuristics were most frugal: they looked up less than a third of the cues (on average, 2.2 and 2.4 as compared to 7.7). What about their accuracy? Multiple regression was the winner when the strategies were tested on the training set, that is, on the set in which their parameters were fitted. However, when it came to predictive accuracy, that is, to accuracy in the hold-out sample, the picture changed. Here, Take The Best was not only more frugal, but also more accurate than the two linear strategies (and even Minimalist, which looked up the least cues, performed closely behind the two linear strategies).

## The robustness of Take The Best

The result presented in Table 1 may sound paradoxical because multiple regression processed all the information that Take The Best did and more. However, by being sensitive to too many features of the data – for instance, by taking correlations between cues into account – multiple regression suffers from overfitting, especially with small data sets. Take The Best, on the other hand, uses few cues. The first cues tend to be highly valid and, in general, they will remain so across different subsets of the same class of objects. The stability of highly valid cues is a main factor for the robustness of Take The Best, that is, its low danger of overfitting in cross-validation as well as in other forms of incremental learning. Thus, Take The Best can have an advantage against more savvy strategies that capture more aspects of the data, when the task requires making out-of-sample predictions.

A similar finding was obtained when Take The Best's search rule was compared to other search rules. Martignon and Hoffrage (2002) split one of the 20 environments mentioned above (namely the set of German cities, which was also used by Gigerenzer/Goldstein 1996) into two halves: in the training set, various cue hierarchies were established, and in the remaining half, these orderings were tested. In the training set (averaged across 100 random splits of the data set), Take The Best's ordering performed 1.2 percentage points worse than an ordering based on conditional validities.

Unlike cue validities that are used to establish Take The Best's ordering and that are simply computed from the set of all pairs, conditional validities are computed from those pairs that are left undiscriminated by cues that have already been used. This ordering, based on conditional cue validities, was still outperformed by the optimal ordering by 0.9 percentage points. In the test set, however, Take The Best's ordering outperformed the other two. Thus, the simplest ordering, which ignored dependencies between cues, turned out to be the best one when the task was to generalize to new objects. It strikes a balance between the dangers of overfitting (i.e., extracting too much information from the training set, as optimal ordering and conditional validity do) and underfitting (extracting too little information).

*The ecological rationality of Take The Best*

The robustness of simple heuristics can be viewed as one aspect of their ecological rationality: Seen through the eyes of a model, information structures in the ecology contain noise that does not transfer to new situations or objects. By being simple, fast and frugal heuristics have – loosely speaking – a lower noise-to-information ratio than complex strategies. While this seems to be true independent of the particular environment, we can also ask whether there are specific characteristics of environments that give Take The Best an advantage (or disadvantage) when compared to other strategies. Here we discuss five properties.

First, Take The Best is equivalent in performance to a linear model whose weights form a noncompensatory set (and decay in the same order as that of Take The Best). A set of cues is noncompensatory if each weight is larger than the sum of all other weights to come (such as ½, ¼, ⅛, …). For such sets of cues, Take The Best is in fact a shortcut to a linear strategy: It uses less information, but will always make the same inference. Therefore, if an environment consists of cues that are noncompensatory, then no linear model can have higher predictive accuracy than Take The Best (Martignon/Hoffrage 2002).

Second, Take The Best has an advantage over a unit-weight linear model when information in an environment is scarce. To illustrate the concept of scarce information, let us recall a fact from information theory: A class of N objects contains $\log_2 N$ bits of information. This means that if we were to encode each object in the class by means of binary cue profiles of the same length, this length should be at least $\log_2 N$ if each object is to have a unique profile. For instance, in order to encode eight objects, it is sufficient to use three ($\log_2 8 = 3$) binary variables. This leads to the following definition: A set of $M$ cues provides scarce information for a reference class of $N$ objects if $M \leq \log_2 N$. By means of exhaustive counting, we found that in the majority of small environments (i.e., environments with fewer than 1,000 objects) with scarce information, Take The Best is more accurate than a unit-weight linear model. Thus, in such scarce environments, a unit-weight linear model can take little advantage of its strongest property, namely compensation.

Third, Take The Best has a disadvantage over linear models in environments with *abundant* information. Adding cues to a scarce environment will do little for Take The Best if the best cues in the original environment already have high validity. For a unit-weight linear model, however, adding cues may help because they can compensate for

various mistakes this rule would have made if restricted to using only the first cues. That was the intuition; here is the corresponding theorem: If an environment consists of all possible cues, a unit-weight linear model will discriminate among all objects and make only correct inferences (for the proof, see Martignon/Hoffrage 2002). Note that in the present context, the term "cue" denotes a binary-valued function in the reference class. Therefore, the number of different cues in a finite reference class is also finite. The theorem can be generalized from the simple linear model with unit weights to linear models that use cue validities as weights. Moreover, it is even true if all certain cues, that is, cues with a validity of 1, are excluded.

Fourth, the differences in performance between Take The Best, a unit-weight linear model, and a linear model that uses cue validities as weights does not seem to depend on the average validity of the cues. We did not find any substantial effects of these environmental properties across 10,000 environments consisting of 16 objects and 4 cues, for which the cue values have been randomly generated (Martignon/ Hoffrage 2002).

Fifth, the number of positive cue intercorrelations had an effect, albeit only a small one. Specifically, in those 10,000 environments with the randomly generated cue values, Take The Best outperformed the linear model with unit weights and the weighted linear model across all environments, in which at least five of the six intercorrelations between the four cue were positive. For each class of environments in which the number of positive cue intercorrelations was 4, 3, 2, 1, and 0, the weighted linear model outperformed Take The Best, and – by a far greater margin – the unit-weight model as well.

*Empirical Evidence*

How do people know when to apply which heuristic? Can mere feedback select heuristics? In an experiment by Rieskamp and Otto (2004), participants took the role of bank consultants with the task of evaluating which of two companies applying for a loan was more creditworthy. Six cues, such as profitability and financial flexibility, were provided for each company. For the first 24 pairs of companies, no feedback was provided as to the correctness of the participants' inferences. Thereafter feedback was given. For one group of participants, the correct answer was determined in about 90% of the cases by Take The Best, that is, feedback was obtained from the cues in a noncompensatory way. For the second group, the more creditworthy company was determined in about 90% of the cases by a weighted additive rule, that is, the feedback was generated in a compensatory way. Did people intuitively adapt their heuristics to the feedback structure of the environments? As can be seen from Figure 2, this was the case: Feedback changed the frequency of responses consistent with Take The Best. Note that in this experiment, participants could acquire information without paying for it. This fosters compensatory strategies, as can be seen from the low initial frequency of around 30% for Take The Best. People learned without instruction that different heuristics are successful in different environments.

Is Take The Best also spontaneously used without previous learning trials? Two conditions that should foster the use of fast and frugal heuristics include (a) situations when decisions have to be made under time pressure (i.e., when it is important to be

**Figure 2:**      **Do people intuitively learn when to use Take The Best? Participants had to make choices based on cue values, followed by feedback that was generated from those cue values. When choices had been made in an environment in which the cues formed a noncompensatory set, the frequency of choices consistent with Take The Best increased over time; when the set of cues was compensatory, this frequency decreased** (Rieskamp/Otto 2004; figure taken from Gigerenzer at al., in press).
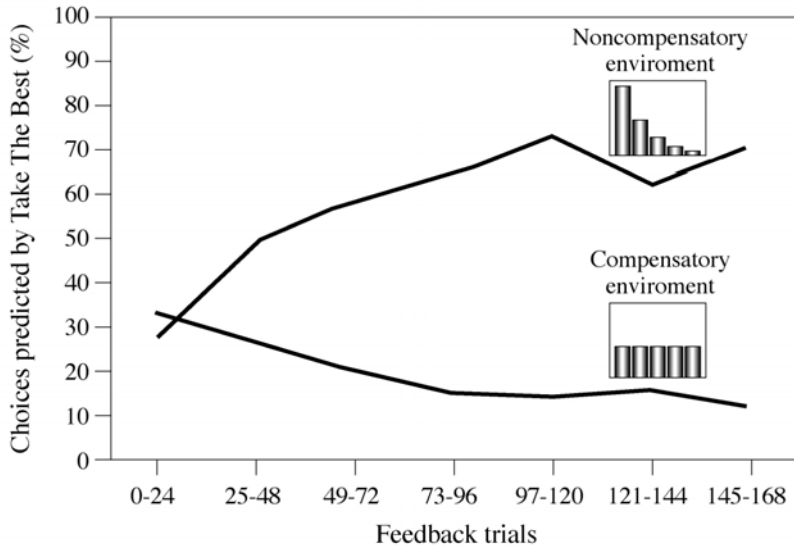


fast), and (b) situations in which information is costly (i.e., when it is important to be frugal). Rieskamp and Hoffrage (1999) studied decision making under the first condition. Specifically, they asked how well eight strategies proposed in the literature predicted people's decisions under low and high time pressure. The participants' task was to predict which of four companies had the highest yearly profit. They could sequentially look up the information from six cues (e.g., amount of investments, the number of employees, etc.). Two strategies modelled participants' choices best: a generalization of Take The Best from binary choices to choices among several alternatives (lexicographic heuristic or LEX) and Weighted Pros (Huber 1979). Weighted Pros considers only the highest value on each cue (i.e., ignores all other values). Under time pressure, participants' choices conformed better to LEX, which is also the computationally less expensive strategy. Only very few participants could be best described by any of the other six models (for other studies on time pressure that provide empirical test for the use of simple heuristics, see Payne/Bettman/Johnson 1988, 1993; Edland 1994).

Bröder (2000, Experiments 3 and 4) studied decision making in a situation in which acquisition of information was costly. Under this condition, more than 60% of the participants were classified as using Take The Best, whereas none was classified as using a compensatory linear model with unit weights (for more evidence on the effect of costs, see Newell/Shanks 2003). For a more elaborate discussion of how heuristics

are acquired either through the course of evolution, through social learning, or through individual learning and how they are selected from the adaptive toolbox when facing a particular task at hand, see Gigerenzer (2003).

**Figure 3:** **A test of how well eight strategies predict people's choices under low (50-second) and high (20-second) time pressure. The four strategies to the left are compensatory, the three to the right are noncompensatory, and LEX-ADD consists of a first, noncompensatory phase and a second, compensatory phase. LEX is a variant of Take The Best, generalized to a choice situation with more than two alternatives** (figure adapted from Rieskamp/Hoffrage 1999).
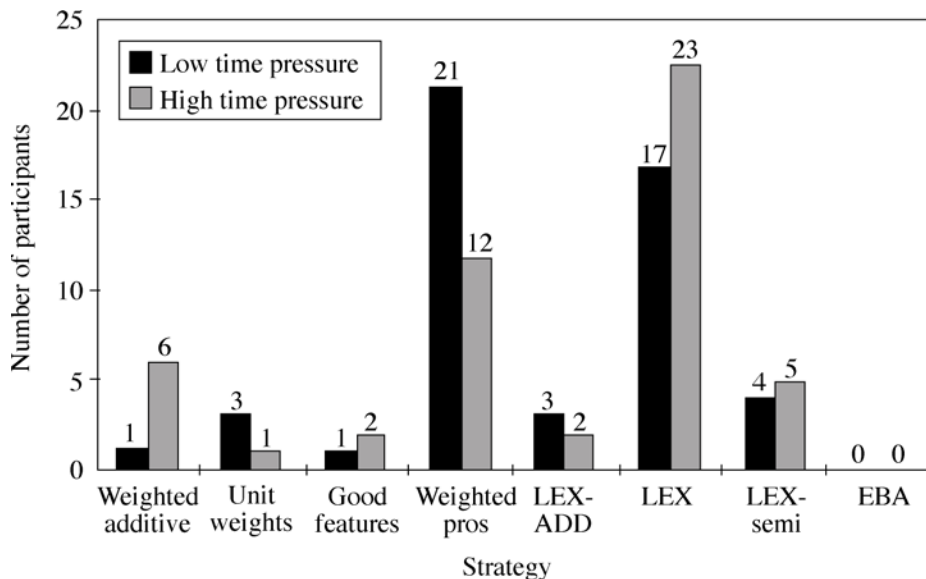


*Simple group heuristics*

The comparison of fast and frugal heuristics and complex strategies with respect to their performance supports the claim that the quality of a decision may not primarily depend on the amount of information that is integrated into it – as long as the heuristics that use only few cues pick the good ones. How does this claim relate to group decision making as compared to individual decision making? Typically, groups have access to much more information than the individual group members do. More information can provide an advantage for group decision-making, such as in cases when the knowledge of the individual group members is biased. Consider the following situation: Two candidates, Candidate *A* and Candidate *B*, apply for a position and a four-member committee has to select one of them. Overall, most arguments are in favour of Candidate *A*. However, no single group member is aware of this because no member knows all arguments. Instead, information is distributed among the four members in a biased way such that each group member has more arguments in favour of Candidate *B* (see Stasser 1992; Reimer/Hoffrage 2003). Such a situation can arise when all group members know the arguments in favour of Candidate *B* (shared in-

formation), whereas each argument in favour of Candidate *A* is known by only one group member (unshared information). Are groups able to detect the "hidden profile," that is, are they able to detect that there are more arguments in favour of Candidate *A* overall (Stasser 1992)? Many experimental studies, which tend to use significantly more pieces of information and more choice alternatives, have revealed that groups usually fail to detect such a hidden profile – in our example, most groups would decide for Candidate *B*. The hidden-profile effect is a robust empirical phenomenon that has been observed across various samples of tasks and groups (for an overview, see Wittenbaum/Stasser 1996).

According to the most prominent explanation of this effect, groups fail to detect hidden profiles because they fail to pool and integrate all available pieces of information (Stasser/Titus 1985). However, in this research tradition, the question of how the information should be processed by the group has been largely ignored. Despite the assumption that the detection of a hidden profile requires exhaustive information processing, several sets of simulation studies revealed that a group version of Take The Best very effectively identifies concealed alternatives in the hidden-profile task (Reimer/Hoffrage 2003). This communication-based Take The Best heuristic requires that group members pool and exchange information on the most valid cues. Thus, a lack of individual knowledge does not necessarily prevent a group from detecting hidden profiles. Rather, performance depends on whether group members share information on valid or on invalid cue dimensions and on how the information is processed and integrated into a group decision. Such a heuristic is particularly effective if cue validities follow a J-shaped distribution in which some cues have a high validity and most cues have low validities (Reimer/Hoffrage 2004).

*Fast and frugal memory inferences*

The fast and frugal decision making approach was recently used to model a well-known phenomenon in memory research, namely the hindsight bias. Once people know the outcome of an event, they tend to overestimate what could have been anticipated in foresight. Rather than thinking of this so-called hindsight bias as a flaw of human cognition, as previous research has done, we argue that it is a by-product of an adaptive mechanism, namely updating of knowledge. According to the RAFT model (which stands for *Reconstruction After Feedback with Take The Best*; Hoffrage/Hertwig/Gigerenzer 2000), hindsight bias occurs when people attempt to reconstruct their predictions of an event's outcome. Given that we do not have unlimited memory capacities, recollection can be replaced by reconstruction in many cases. Imagine you were once asked about the result of a simple arithmetic calculation, such as 26*48. Further imagine that at some later point in time, your task is to remember what you had responded. If you are not able to retrieve it from memory, you still have a good chance of correctly "recalling" your initial answer by simply redoing the calculation.

The RAFT model applies the logic that recollection can be replaced by reconstruction to knowledge-based probabilistic inferences. To the extent that the mechanisms involved in such inferences are reliable, the same outcome should be obtained. However, the model also assumes that we constantly update our knowledge, thereby overwriting previous (and possibly outdated) knowledge states. In particular, the

model assumes that feedback or correct information about things we have previously inferred changes the knowledge base underlying our original judgment and causes a bias toward the new information. In two studies, we could show that feedback about a criterion variable led to systematic changes in memory of the cue values that were previously used to infer the criterion variable (Hoffrage/Hertwig/Gigerenzer 2000). We further showed that reconstructions based on this updated knowledge base could account for the hindsight bias that we obtained for recollections of the criterion variable. Note that these tests were performed for each participant's individual items, thereby reaching a level of analysis that is rarely achieved in cognitive psychology.

Another advantage of precisely formulated process models such as RAFT is that they allow for computer simulations. In a series of such simulations, we demonstrated that knowledge updating is adaptive, as it increases the accuracy of inferences (Hertwig/Fanselow/Hoffrage 2003). These simulations also showed that the RAFT model is able to account for well-known effects (such as the expertise effect, which says that the more comprehensive people's knowledge is in foresight, the smaller their hindsight bias is), and to make new and counterintuitive predictions (e.g., even if foresight knowledge is false, it can reduce hindsight bias). In sum, the RAFT model conceives the hindsight bias as a small price to pay for a much larger gain: a well-functioning memory that is able to forget what we do not need – such as outdated knowledge – and that constantly updates our knowledge.

## Notions of Rationality

In the early 19th century, the astronomer and philosopher Pierre Simon Laplace imagined an omniscient superintelligence that "could comprehend all the forces of which nature is animated and the respective situation of the beings who compose it – an intelligence sufficiently vast to submit these data to analysis ... nothing would be uncertain and the future, the past, would be present to its eyes" (Laplace 1814, *Essai Philosophique:* 1325). More than 100 years before Laplace, the English philosopher John Locke (1690) had contrasted this secularized version of God with us humble humans living in the "twilight of probability." Even if the world were deterministic and such a superintelligence as described by Laplace existed, it would not make any difference for us mortals to whom the world appears to be uncertain and who have to make inferences with limited information, limited computational capacities, and often also with limited time. While Laplace's vision of a superintelligence has inspired models of unbounded rationality, Locke's view describes a situation that led Herbert Simon (1982) to call for models of bounded rationality.

The adaptive toolbox of us mortal human beings contains fast and frugal heuristics that can be considered such models of bounded rationality. These heuristics offer a theoretical alternative to three different notions of rationality. First, they should be distinguished from models of unbounded rationality, such as maximization of expected utility and Bayesian models. While proponents of unbounded rationality generally acknowledge that their models assume unrealistic mental abilities, they nevertheless defend them by arguing that humans act *as if* they were unboundedly rational. Within this interpretation, the laws of probability do not describe the process but merely the outcome of reasoning. Fast and frugal heuristics, in contrast, not only re-

quire less computational capabilities and less information, but they also specify the cognitive processes, that is, how information is searched for, when information search is stopped, and how a decision is made based on the acquired information.

Second, fast and frugal heuristics should be distinguished from models that optimize under constraints. Both assume that information is not just given, but needs to be searched for. Moreover, both specify how search should proceed. Unlike fast and frugal heuristics with their simple search rules and simple stopping rules, however, models that optimize under constraints assume that the stopping rule optimizes search with respect to the time, computation, money, and other resources being spent. More specifically, this vision of rationality holds that the mind should calculate the benefits and costs of searching for each further piece of information and stop search as soon as the costs outweigh the benefits (e.g., Sargent 1993). The rule "stop search when costs outweigh benefits" sounds plausible at first glance. But a closer look reveals that optimization under constraints invites unbounded rationality to sneak in through the back door in that computing the optimal stopping point can require even more knowledge and computation, and leads, at least in theory, to an infinite regress (Vriend 1996).

Third, fast and frugal heuristics should be distinguished from heuristics in the tradition of the heuristics-and-biases program (Kahneman/Slovic/Tversky 1982). Although both research programs embrace the idea of simple mechanisms, and both are concerned with identifying the situations in which these heuristics are used, they are different on several dimensions (also see the debate between Kahneman/Tversky 1996, and Gigerenzer 1996). Within the heuristics-and-biases program, heuristics were invoked as explanations for systematic errors found in human reasoning – mainly deviations from the laws of probability. Although Tversky and Kahneman repeatedly asserted that heuristics sometimes succeed and sometimes fail, they and many of their colleagues focused on the latter category and interpreted their experimental findings as indicating some kind of fallacy. These fallacies were attributed (usually post hoc) to one of three main heuristics: representativeness (judgments influenced by what is typical), availability (judgments based on what comes easily to mind), or anchoring and adjustment (judgments relying on what comes first). Fast and frugal heuristics, in contrast, are not associated with the value-laden term "bias." On the contrary, by taking advantage of the structure of information in the environment, these heuristics can lead to accurate and useful inferences; hence, they do not necessarily lead to biases but they can "make us smart" (Gigerenzer et al. 1999). Another difference is that in the heuristics-and-biases program, heuristics are indeterminate and imprecise. They are one-word labels that at once explain too little and too much: too little, because the underlying processes are left unspecified, and too much, because, with sufficient imagination, one of them can be fit post hoc to almost any empirical result. One and the same heuristic may even be used to explain opposite findings (Gigerenzer, 2004). For instance, the representativeness heuristic has been used to explain both the "gambler's fallacy" (the belief that a series of binary events is more likely to stop) and the "hot hand fallacy" (the belief that a series of binary events is more likely to continue). The simple heuristics program, in contrast, formulates precise process models with zero adjustable parameters that completely lack such flexibility. Moreover, as has already

been mentioned in the first part of the present paper, the simple-heuristics program dispenses with the focus on coherence criteria (e.g., some normative answer, which is derived by applying a law of probability in a content-blind way) as the yardsticks of rationality. Rather, fast and frugal heuristics are designed to make inferences about the real world, thus, the standard against which they are evaluated is their correspondence with external criteria.

Let us summarize how fast and frugal heuristics as models of bounded rationality relate to other notions of rationality: They are – consistent with traditional views about what constitutes a heuristic – useful tools for finding reasonable solutions to a given problem. In contrast to models of unbounded rationality, they are psychologically plausible process models that take the limits of human information processing into account. In contrast to models that optimize under constraints, they consist of simple building blocks, which dispense with unrealistic assumptions that inconspicuously introduce complex calculations when it comes to specifying stopping criteria for information search. In contrast to heuristics proposed in the heuristics-and-biases program, they are formalized enough to allow for computer simulations to determine their fit in a particular real-world environment. Given their ecological rationality, which exploits the structure of the environment in which they function, fast and frugal heuristics do not need to sacrifice accuracy for speed and frugality. Thus, it can be rational not to use information even when it is available: Simplicity does not imply irrationality.

## References

Bröder, A. (2000): Assessing the empirical validity of the "Take-the-best" heuristic as a model of human probabilistic inference. In: Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 1332–1346.

Busemeyer, J./Townsend, J. T. (1993): Decision Field Theory: A dynamic cognition approach to decision making. In: Psychological Review, 100, 432–459.

Czerlinski, J./Gigerenzer, G./Goldstein, D. G. (1999): How good are simple heuristics? In: G. Gigerenzer/P. M. Todd/the ABC Research Group, Simple heuristics that make us smart (pp. 97–118). New York: Oxford University Press.

Dawes, R. M. (1979): The robust beauty of improper linear models in decision making. In: American Psychologist, 34, 571–582.

Duncker, K. (1935): Zur Psychologie des produktiven Denkens [The psychology of productive thinking]. Berlin: Julius Springer.

Edland, A. (1994): Time pressure and the application of decision rules: Choices and judgments among multiattribute alternatives. In: Scandinavian Journal of Psychology, 35, 281–291.

Einstein, A. (1905): Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. In: Annalen der Physik, 17, 132.

Galef, B. G., Jr. (1987): Social influences on the identification of toxic foods by Norway rats. In: Animal-Learning & Behavior, 15, 327–332.

Galef, B. G., Jr./McQuoid, L. M./Whiskin, E. E. (1990): Further evidence that Norway rats do not socially transmit learned aversions to toxic baits. In: Animal Learning & Behavior, 18, 199–205.

Gigerenzer, G./Hoffrage, U. (1995): How to improve Bayesian reasoning without instruction: Frequency formats. In: Psychological Review, 102, 684–704.

Gigerenzer, G. (1996): On narrow norms and vague heuristics: A reply to Kahneman and Tversky. In: Psychological Review, 103, 592–596.

Gigerenzer, G. (2001): The adaptive toolbox: Toward a Darwinian rationality. In: J. A. French/A. C. Kamil/D. W. Leger (Eds.): Evolutionary psychology and motivation (pp. 113–143). Nebraska Symposium on Motivation, Volume 48. Lincoln: University of Nebraska Press.

Gigerenzer, G. (2003): The adaptive toolbox and life span development: common questions? In: U. M. Staudinger/U. Lindenberger (Eds.): Understanding human development: Dialogues with lifespan psychology (pp. 423–435). Boston, MA: Kluwer.

Gigerenzer, G. (2004): Fast and frugal heuristics: The tools of bounded rationality. In: D. Koehler/N. Harvey (Eds.): Handbook of judgment and decision making (pp. 62–88). Oxford, UK: Blackwell.

Gigerenzer, G./Goldstein, D. G. (1996): Reasoning the fast and frugal way: Models of bounded rationality. In: Psychological Review, 103, 650–669.

Gigerenzer, G./Selten, R. (2001): Bounded rationality: The adaptive toolbox. Cambridge, MA: The MIT Press.

Gigerenzer, G./Martignon, L./Hoffrage, U./Rieskamp, J./Czerlinski, J./Goldstein, D. (in press): One-reason decision making. In: C. R. Plott/V. L. Smith (Eds.): Handbook of experimental economics results. North Holland/Elsevier Press.

Gigerenzer, G./Todd, P. M./the ABC Research Group (1999): Simple heuristics that make us smart. New York: Oxford University Press.

Gigone, D./Hastie, R. (1997): The impact of information on small group choice. In: Journal of Personality and Social Psychology, 72, 132–140.

Goldstein, D. G./Gigerenzer, G. (2002): Models of ecological rationality: The recognition heuristic. In: Psychological Review, 109, 75–90.

Hammond, K. R. (1996): Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice. New York: Oxford University Press.

Hertwig, R./Fanselow, C./Hoffrage, U. (2003): Hindsight Bias: How knowledge and heuristics affect our reconstruction of the past. In: Memory, 11, 357–377.

Hertwig, R./Hoffrage, U./Martignon, L. (1999): Quick estimation: Letting the environment do the work. In: G. Gigerenzer/P. M. Todd/the ABC Research Group. Simple heuristics that make us smart (pp. 209–234). New York: Oxford University Press.

Hoffrage, U./Hertwig, R./Gigerenzer, G. (2000): Hindsight bias: A by-product of knowledge updating? In: Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 566–581.

Hoffrage, U./Lindsey, S./Hertwig, R./Gigerenzer, G. (2000): Communicating statistical information. In: Science, 290, 2261–2262.

Huber, O. (1979): Nontransitive multidimensional preferences: Theoretical analysis of a model. In: Theory and Decision, 10, 147–165.

Kahneman, D./Slovic, P./Tversky, A. (Eds.) (1982): Judgment under uncertainty: Heuristics and biases. New York: Cambridge University Press.

Kahneman, D./Tversky, A. (1996): On the reality of cognitive illusions: A reply to Gigerenzer's critique. In: Psychological Review, 103, 582–591.

Laplace, P. S. (1814/1951): A philosophical essay on probabilities. (F. W. Truscott/F. L. Emory, Trans.). New York: Dover. (Original work published 1814).

Locke, J. (1690/1959): An essay concerning human understanding. (A. C. Fraser, Ed.) New York: Dover.

Martignon, L./Hoffrage, U. (2002): Fast, frugal and fit: Simple heuristics for paired comparison. In: Theory and Decision, 52, 29–71.

McLeod, P./Dienes, Z. (1996): Do fielders know where to go to catch the ball or only how to get there? In: Journal of Experimental Psychology: Human Perception and Performance, 22, 531–543.

Newell, A./Simon, H. (1972): Human problem solving. Englewood Cliffs, NJ: Prentice Hall.

Newell, B./Shanks, D. (2003): Take the best or look at the rest? Factors influencing 'one-reason' decision-making. In: Journal of Experimental Psychology: Learning, Memory, and Cognition, 29, 53–65.

Ortmann, A./Gigerenzer, G./Borges, B./Goldstein, D. G. (in press): Recognition heuristic: A fast and frugal way to investment choice? In: C. R. Plott/V. L. Smith (Eds.): Handbook of experimental economics results. North-Holland: Elsevier.

Payne, J. W./Bettman, J. R./Johnson, E. J. (1988): Adaptive strategy selection in decision making. In: Journal of Experimental Psychology: Learning, Memory, and Cognition, 14, 534–552.

Payne, J. W./Bettman, J. R./Johnson, E. J. (1993): The adaptive decision maker. New York: Cambridge University Press.

Reimer, T./Hoffrage, U. (2003): Information aggregation in groups: The approach of simple group heuristics (SIGH). In: R. Alterman/D. Kirsch (Eds.): Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society (pp. 982–987). Boston: Cognitive Science Society.

Reimer, T./Hoffrage, U. (2004): Can simple group heuristics detect hidden profiles in randomly generated environments? Manuscript submitted for publication.

Reimer, T./Katsikopoulos, K. (in press): The use of recognition in group decision-making. Cognitive Science.

Rieskamp, J./Hoffrage, U. (1999): When do people use simple heuristics, and how can we tell? In: G. Gigerenzer/P. M. Todd/the ABC Research Group. Simple heuristics that make us smart (pp. 141–167). New York: Oxford University Press.

Rieskamp, J./Otto, P. (2004): SSL: A theory of how people learn to select strategies. Manuscript submitted for publication.

Sargent, T. J. (1993): Bounded rationality in macroeconomics. Oxford: Oxford University Press.

Simon, H. A. (1947): Administrative behavior. New York: Macmillan.

Simon, H. A. (1982): Models of bounded rationality. Cambridge, MA: The MIT Press.

Simon, H. A. (1990): Invariants of human behavior. In: Annual Review of Psychology, 41, 1–19.

Stasser, G. (1992): Information salience and the discovery of hidden profiles by decision–making groups: A "thought experiment." In: Organizational Behavior and Human Decision Processes, 52, 156–181.

Stasser, G./Titus, W. (1985): Pooling of unshared information in group decision making: Biased information sampling during discussion. In: Journal of Personality and Social Psychology, 48, 1467–1478.

Stigler, G./Becker, G. (1977): De gustibus non est disputandum. In: American Economic Review, 67, 76–90.

Todd, P. M/Gigerenzer, G./the ABC Research Group (in preparation): Ecological rationality: From simple heuristics to environment structures.

Todd, P. M./Gigerenzer, G. (2000): Precis of Simple heuristics that make us smart. In: Behavioral and Brain Sciences, 23, 727–780.

Toscani, O. (1997): Die Werbung ist ein lächelndes Aas. Frankfurt a. M.: Fischer.

Vriend, N. J. (1996): Rational behavior and economic theory. In: Journal of Economic Behavior & Organization, 29, 263–285.

Wittenbaum, G. M./Stasser, G. (1996): Management of information in small groups. In: J. L. Nye/A. B. Brower (Eds.): What's social about social cognition? (pp. 3–28). London: Sage.