Published in: French, J. A., Kamil, A. C., & Leger, D. W. (Eds.). (2001). *Evolutionary psychology and motivation* (Nebraska Symposium on Motivation,Vol. 48, pp. 113–143). Lincoln: University of Nebraska Press.

The Adaptive Toolbox: Toward a Darwinian Rationality

Gerd Gigerenzer

A cartoon shows an early *Homo sapiens* standing in front of a cave. He is calculating the trajectory of a lion's jump and the magnitude of the impulse the lion will have in order to decide what to do. The last picture shows a sated, happy lion. The cartoon makes us smile because its message conflicts with our ideal of rational decision making, which demands that we go through all the available information, deduce all the possible consequences, and compute the optimal decision. Good decision making, from this point of view, is based on the ideals of omniscience and optimization. An organism aiming for these heavenly ideals, however, might not survive on earth. Nevertheless, the majority of models of rational decision making in the social, behavioral, and cognitive sciences, as well as in economics, rely on some version of this doctrine. Even when empirical studies show that actual human beings cannot live up to it, the doctrine is not abandoned as other models would be—it is retained and declared a norm, that is, how we *should* reason.

In this chapter, I introduce an alternative to this doctrine of rational choice. In my view, intelligent behavior is the product of the "adaptive toolbox" of a species, which hosts a collection of heuristics—rather than one general intelligence or an optimizing calculus. Applied in the right situation, these heuristics can be fast and effective. As we will see, the rationality of the adaptive toolbox is not logical but ecological.

I begin with two examples of heuristics. They illustrate that, in the real world, lack of omniscience need not be a bad thing. Some heuristics can accomplish a lot with little knowledge and time.

Fast and Frugal Decision Making

A man is rushed to a hospital in the throes of a heart attack. The doctors need to make a decision, and they need to make it quickly: Should the victim be treated as a low-risk or a high-risk patient? How does one make such a decision? Theories of rational choice as well as common sense dictate that the doctors determine all the known relevant predictors—and there are at least 20 of them—and then combine these measures into a final conclusion, preferably with the aid of a fancy statistical software package. Now consider the simple decision tree in Figure 1, designed by Leo Breiman and his colleagues. It asks, at most, only three questions. If a patient has a systolic blood pressure of 91 or less, he is immediately classified as high risk—no other variables are ascertained. If systolic blood pressure is higher than 91, a second variable is considered—age. If the patient is 62.5 years old or younger, he is immediately classified as low risk. No further information is sought. If he is older, a third variable is measured that will finally classify him as high or low risk.

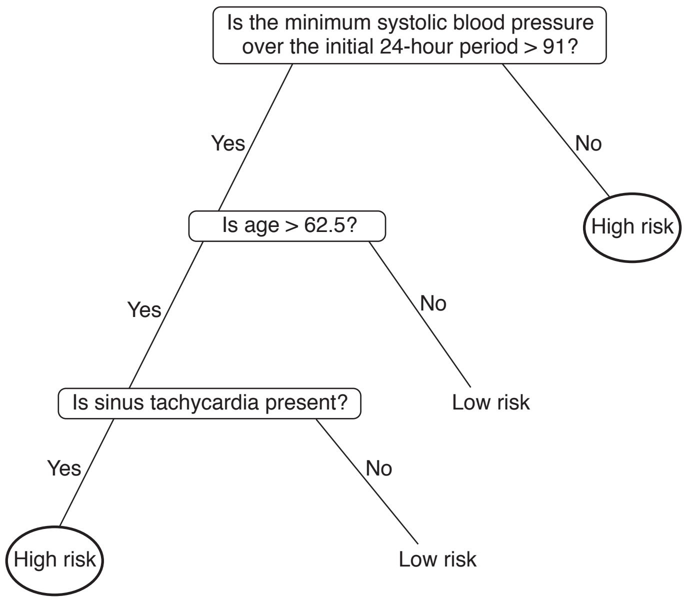

Figure 1. A simple decision tree for classifying incoming heart attack victims as high-risk or low-risk patients (adapted from Breiman et al., 1993). By ignoring a great part of the information, this tree can make more accurate classifications than standard statistical models that use all the information available. This tree is a version of one-reason decision making—the decision is based on only one variable, although up to three variables may be looked up.

This decision tree is simple in three respects. First, it uses, at most, three predictors and ignores the rest. Second, it dichotomizes each predictor, that is, it dispenses with quantitative information, such as whether a patient is 60 or 40 years old. Third, the three predictors are not combined; for instance, lower blood pressure cannot be compensated for by younger age. Only one predictor determines each decision. I call this *one-reason decision making.*

But how accurate is one-reason decision making? Would you want to be classified by three yes/no questions in a situation with such high stakes? The counterintuitive result is that this simple tree turns out to be more accurate in classifying actual heart attack patients than traditional statistical methods that use all the available predictors (Breiman, Friedman, Olshen, & Stone, 1993). Simplicity can pay off.[1]

Let us look at a second, quite different situation—sport. Imagine you want to build a robot that can catch balls—a robot that can play baseball, cricket, or soccer, depending on the nationality of your robot. It is a thought experiment—no such robots yet exist. If you follow a classical AI approach, you aim to give your robot a complete representation of its environment and the most sophisticated computational machinery. First, you might feed your robot the family of parabolas (because thrown balls have parabolic trajectories). In order to choose the right parabola,

[1] Decision trees such as the one in this example are easy to use but their construction is based on quite extensive computations. In this chapter, and in Gigerenzer, Todd, and the ABC Research Group (1999), we will see how fast and frugal heuristics can get around this costly construction phase.

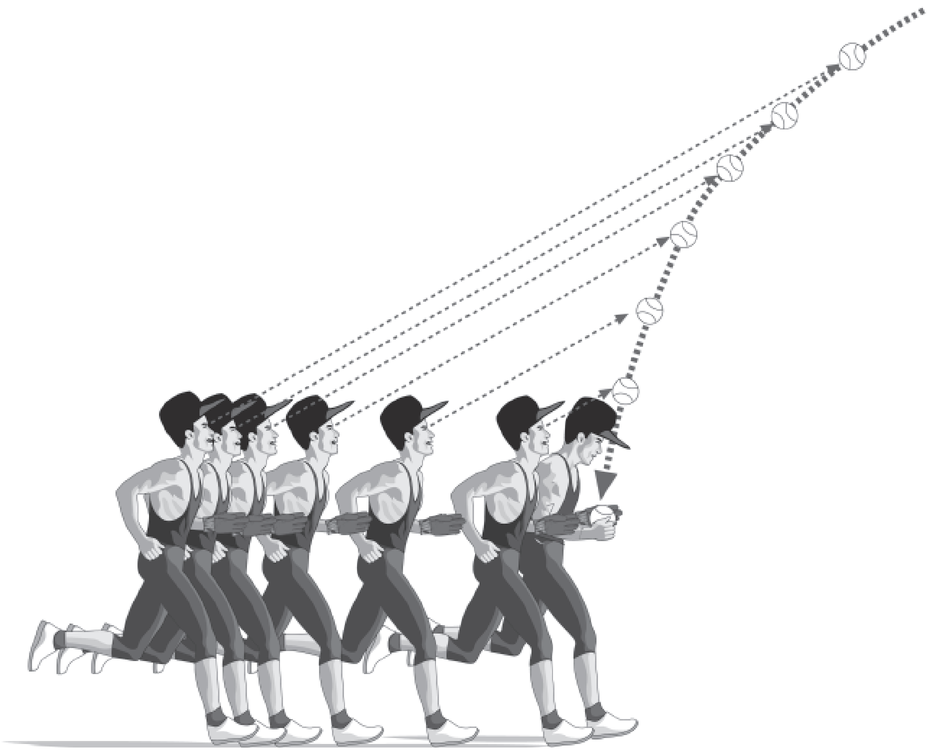

Figure 2. The gaze heuristic: A frugal strategy for the interception of moving objects, such as catching a ball while running. When the ball is descending, as shown here, the player only needs to fixate on the ball and adjust his running speed so that the angle of gaze remains constant. When the ball is ascending (not shown here), the player needs to adjust his running speed so that the angle of gaze remains within 0 and 90 degrees (McLeod & Dienes, 1996). In each case, only one variable needs to be attended to—another form of one-reason decision making in which causal variables can be ignored.

the robot needs instruments that can measure the ball's initial distance, its initial velocity, and its projection angle. But in the real world, balls do not fly in parabolas because of air resistance and wind. Thus, the robot would need additional instruments to measure the speed and direction of the wind at each point of the ball's flight and compute the resulting path. A true challenge. And there is more: Spin and myriad further factors would have to be measured and incorporated into a complete representation for the robot's use.

As in the heart attack situation, there is an alternative strategy that does not aim at complete information and representation, but rather at smart heuristics. One way to discover such heuristics is to study actual players. (On the other hand, if one assumes all the complex measurements and computations, one must further assume that these are unconscious and unobservable. This would obviate studying actual players. Have you ever interviewed a soccer player?) McLeod and Dienes (1996) discovered that experienced players use a simple heuristic. When a ball comes in high, the player fixates on the ball and starts running. The simple heuristic is to adjust the running speed so that the angle of gaze remains constant (or within a certain range)—that is, the angle between the eye and the ball (Figure 2). 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 get away with ignoring this information. All the relevant information is contained in one variable—the angle of gaze. Note that this robot, unlike its hypothetical, omniscient, and old-fashioned AI competitor, is not able to compute the point at which the ball will land. But the simple-minded robot will be there where the ball lands (and catch it or at least be hit by it).[2]

Visions of Rationality

These two examples illustrate two different visions of rationality. In Figure 3, I have labeled them demons and bounded rationality. Demons are popular in the social, cognitive, and behavioral sciences. There are two species of demons: those that exhibit unbounded rationality and those that optimize under constraints.

Unbounded Rationality

Unbounded rationality is about decision strategies that ignore the fact that humans (and other animals) have limited time, knowledge, and computational capacities. In this framework, the question is: If individuals were omniscient and had all eternity at their disposal, how would they behave? Maximizing expected utility, Bayesian models, and *Homo economicus* are examples of unbounded rationality frameworks. *Homo economicus,* for instance, chooses an action from a set of alternatives by first determining all possible consequences of each action, then computing the probabilities and utilities of these consequences, then calculating the expected utilities of each action, and finally choosing the action that maximizes the expected utility. Psychological theories have incorporated the same ideal. For instance, expectation-value theories of motivation assume that, of the many courses of action, the one chosen has the highest subjective expected value (see Heckhausen, 1991). Theories of causal attribution assume that a cause is attributed to an event in the same way that a statistician of the Fisherian school (Kelley, 1973) or a Bayesian statistician (e.g., Ajzen, 1977) would test a causal hypothesis. In general, unbounded rationality assumes some form of omniscience and optimization. Omniscience is epitomized in the assumption that, in order to make appropriate decisions, an individual must have a complete representation of its environment (as in good old-fashioned AI and in optimal foraging theories). Optimization means that, using this information, the maximum or minimum of a function (such as expected utility) is calculated. Thus, optimization is a process, not an outcome.

Unbounded rationality recreates humans in the image of God, or in a secularized version thereof—Laplace's superintelligence. The weakness of unbounded rationality is that it does not

2 Alan Kamil suggested that the gaze heuristic cannot be the whole story because human players give up on chasing balls that are out-of-bounds, which seems to imply that they compute the point where the ball will land. I do not think so. The gaze heuristic can also provide the information for when to stop trying. For instance, when the player realizes that he cannot run fast enough to keep the angle of gaze constant (or within a certain range), then he knows he will not catch the ball and stops running—without computing the point where the ball actually will land.

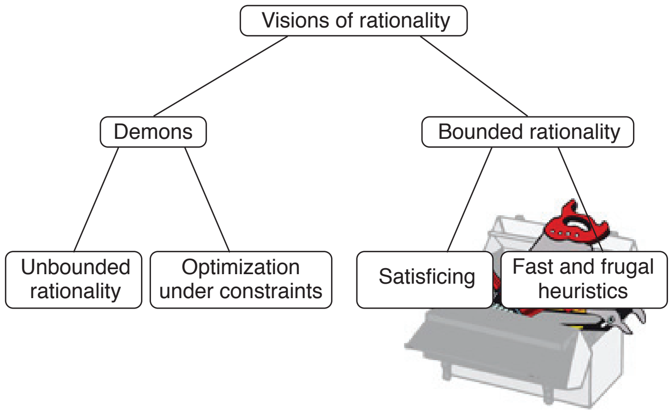

Figure 3. Visions of rationality. The label "demons" stands for models of human, animal, and artificial intelligence that assume that the agent has complete knowledge (or a complete "mental representation") of its environment, and uses optimization calculations (i.e., to compute a maximum or minimum of a function) to make decisions based on this knowledge. Omniscience and optimization are the key ideas of unbounded rationality, whereas models of optimization under constraints relax some of these strong assumptions by building in constraints such as limited time and information costs. However, the more constraints are built in, the more complex the optimization calculations tend to become, which can prevent both psychological plausibility and mathematical tractability. Models of bounded rationality, in contrast, dispense with optimization as the process of decision making—although, in the right environment, they can lead to optimal or good enough outcomes. Note that optimization does not guarantee optimal outcomes; for instance, some of the simplifying assumptions, on which optimization in the messy real world needs to be built, may be false.

describe the way real people think—not even philosophers, as the following anecdote illustrates. A philosopher from Columbia was struggling to decide whether to accept an offer from a rival university or to stay where he was. His colleague took him aside and said: "Just maximize your expected utility—you always write about doing this." Exasperated, the philosopher responded: "Come on, this is serious."

Optimization Under Constraints

In 1961, the economist George Stigler made the image of *Homo economicus* more realistic. He emphasized the fact that humans are not omniscient and therefore need to *search* for information—which costs time and money. However, Stigler chose to retain the ideal of optimization and assumed that search is stopped when the costs of further search exceed its benefits; in other words, an optimal stopping point is calculated. This vision of rationality is known as *optimization under constraints* (such as time). Few psychological theories have included search (a noteworthy exception is Anderson, 1990). Similarly, few experiments allow participants to search for information. Most of them lay all the relevant information out in front of the participant and thereby exclude search, either in memory or in the outside world. For instance, experiments on classification (see Berretty et al., 1999), reasoning, and judgment and decision making (see Gigerenzer, 1996a, 1996b) typically use artificial or hypothetical content, which makes search for information irrelevant. Note that elimination of search in experiments and the postulate of optimization go hand in hand. If search for information, in memory or in the outside world, were allowed, this would increase the two to four dimensions on which artificial stimuli are typically allowed to vary to a much larger and potentially infinite number, which can quickly make optimization computationally intractable.

Even devoted proponents of optimization under constraints have pointed out that the resulting models generally become more demanding than models of unbounded rationality, both mathematically and psychologically. In optimization under constraints, humans are recreated in the image of econometricians, one step above the gods.

In contrast, Herbert Simon (1956, 1992), the father of bounded rationality, argued that a theory of rationality has to be faithful to the actual cognitive capacities of human beings—to their limitations in knowledge, attention, memory, and so on. To Simon's dismay, his term *limitations* has often been interpreted as being synonymous with *constraints for optimization,* and the term *bounded rationality* confused with optimization. In a personal conversation, he once remarked with a mixture of humor and anger that he had considered suing authors who misused his concept of bounded rationality to construct even more complicated and unrealistic models of the human mind.

Bounded Rationality: The Adaptive Toolbox

The metaphor of the *adaptive toolbox* can help to avoid the misapprehension that making rationality more realistic means making optimization more complex. The adaptive toolbox of a species contains a number of heuristics, not one general optimization calculus. Some are inherited, others learned or designed. The gaze heuristic and the medical decision tree are tools in the box. Like hammers and wrenches, they are designed for specific classes of problems; there is no general-purpose tool. The gaze heuristic, for instance, only works for a limited class of problems that involve the interception of moving objects, such as when an animal pursues potential prey. The heuristic also works for avoiding collisions. For instance, if you learn to fly an airplane, you will be taught a version of this heuristic: When another plane is approaching, look at a scratch in your windshield and see whether the other plane moves relative to that scratch. If it does not, dive away quickly.

There are various kinds of tools in the adaptive toolbox. One kind, Simon's "satisficing," involves search and an aspiration level that stops search. For instance, when searching for a house, satisficers search until they find the first house that meets their aspiration level, then stop searching, and go for it. I will talk today about a second kind: fast and frugal heuristics (Gigerenzer, Todd, & the ABC Research Group, 1999). The difference is this: Satisficing involves search across alternatives, such as houses and potential spouses, assuming that the criteria are given (the aspiration level). Fast and frugal heuristics, in contrast, search for criteria or cues, assuming that the alternatives are given. For instance, classifying heart attack patients into high- and low-risk categories is such a situation. The alternatives are given (high or low risk), and one has to search for cues that indicate to which of the alternative categories a patient belongs. Asking at most three yes/no questions is a fast and frugal heuristic: fast, because it does not involve much computation, and frugal, because it only searches for some of the information.

The adaptive toolbox is, in two respects, a Darwinian metaphor for decision making. First, evolution does not follow a grand plan, but results in a patchwork of solutions for specific problems. The same goes for the toolbox: Its heuristics are domain specific, not general. Second, the heuristics in the adaptive toolbox are not good or bad, rational or irrational, per se, only relative to an environment, just as adaptations are context-bound. In these two restrictions lie their potential: Heuristics can perform astonishingly well when used in a suitable environment. The rationality of the adaptive toolbox is not logical, but rather ecological.

How can one identify and experimentally study fast and frugal heuristics? I will first use the most frugal heuristic my research group at the Max Planck Institute has studied for illustration—the recognition heuristic, which is an instance of a class of heuristics I call ignorance-based decision making. It can only be applied if you are sufficiently ignorant—for example, if you are unable even to recognize relevant names.

Ignorance-Based Decision Making

The Recognition Heuristic

Which city—San Diego or San Antonio—has more inhabitants? Daniel G. Goldstein and I posed this question to undergraduates at the University of Chicago. Sixty-two percent of them got the answer right (San Diego). Then we asked German students. They not only knew very little about San Diego, many of them had not even heard of San Antonio. What percentage of the Germans got the answer right?—100%. How can this be? The answer is that the German students used the recognition heuristic: If one city is recognized and the other is not, then infer that the recognized city has the higher value. Note that the American students could not use the recognition heuristic because they had heard of both cities (Goldstein & Gigerenzer, 1999).

Now consider sports. Ayton and Önkal (1997) asked British and Turkish students to predict the results of all 32 English F. A. Cup third-round soccer matches. The Turkish students knew very little about English soccer and had not heard of many of the teams. In 95% of the cases where one team was recognized (familiar to some degree) but the other was not, the Turkish students bet that the team whose name they had heard of would win. Their predictions were almost as good as those of the experienced British students. As before, the recognition heuristic turned partial ignorance into reasonable inference.

When the task is to predict which of two objects has a higher value on some criterion (e.g., which team will win), the recognition heuristic can be simply stated: If one of two objects is recognized and the other is not, then infer that the recognized object has the higher value.

Note that the recognition heuristic can *only* be applied when one of the two objects is not recognized, that is, under partial ignorance. In a domain where recognition correlates negatively with the criterion, "higher" needs to be replaced with "lower" in the definition.

Ecological Rationality

Like all heuristics in the adaptive toolbox, the recognition heuristic is not foolproof. It works in certain situations, but would be useless in others. Its rationality depends on the environment, a term I use as shorthand for the structure of the environment as it is known to an agent. This notion of ecological rationality differs from the notion of rationality as internal coherence, in which rationality is defined by internal laws of judgment (such as transitivity) that do not relate to specific structures of environments. *The recognition heuristic is ecologically rational when ignorance is systematic rather than random, that is, when lack of recognition is correlated with the criterion.* This correlation, the recognition validity α, can be determined empirically.

How accurate is the recognition heuristic? Equation 1 specifies the proportion of correct predictions c that the recognition heuristic will make, such as in predicting the outcomes of a series of sports games or multiple choice tests.

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

All four variables, α, β, N and n, are empirically measurable; no parameter fitting is involved. A person's recognition validity α and her knowledge validity β are easily measured: The recognition validity is the proportion of correct choices among all pairs in which one alternative is recognized and the other is not; the knowledge validity is the same proportion when both alternatives are recognized. The right side of the equation breaks into three parts: The leftmost term equals the proportion of correct inferences made by the recognition heuristic; the middle term equals the proportion of correct inferences resulting from guessing; the rightmost term equals the proportion of correct inferences made when knowledge beyond mere recognition can be used. Thus, the three terms cover the three possible states: one, none, or both objects are recognized. Inspecting this equation, we see that if the number of objects recognized, n, is zero, then all questions will lead to guesses and the proportion correct will be .5. The total number of objects is N. If $n = N$, then the two leftmost terms become zero and the proportion correct will be β. We can also see that the recognition heuristic will come into play most when the participant is operating under "half ignorance," that is, when half of the objects are recognized ($n = N - n$), because this condition maximizes the number of pairs $n\,(N - n)$ in which one object is recognized and the other is not.

The Less-Is-More Effect

A little mathematics reveals that the recognition heuristic can lead to a counterintuitive phenomenon: the *less-is-more effect.* The less-is-more effect occurs when less knowledge leads to more accurate predictions. This happens when a person's recognition validity α is larger than her knowledge validity β: A less-is-more effect occurs when $\alpha > \beta$.

Figure 4 shows an example of a less-is-more effect: With increasing knowledge, performance increases up to a certain point and then drops, as the recognition heuristic can be used less and less often. That's mathematics, you may say, but can the effect be observed in the real world? Can it be that there are situations in which more knowledge can hurt? If you know significantly *more* about one domain than another, can it be that you will systematically perform *worse?* Equation1 specifies the conditions under which one can produce a less-is-more effect experimentally. For instance, Daniel Goldstein and I gave University of Chicago students the names of the 22 largest American cities and asked them, for each of the resulting 231 pairs of cities, which one has the larger population. Then the American students were asked to do the same with the largest German cities, about which they knew very little. To their own surprise, more answers were accurate for German cities than for American cities—less is more (Goldstein & Gigerenzer, 1999).

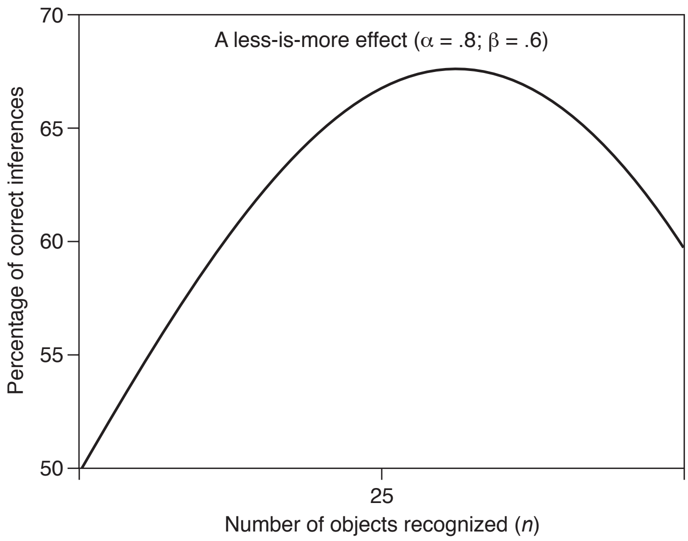

Figure 4. Illustration of a less-is-more effect. The recognition validity is .8, that is, among all pairs of objects where one is recognized by a person and the other is not, the recognized object scores higher on a criterion in 80% of the cases—for example, wins the game. The knowledge validity is .6, that is, among all pairs of objects where both are recognized by a person, the person makes 60% correct predictions. When a person has not heard of any of the objects ($n = 0$), performance is at chance level; when the number of objects known increases, performance increases. But from some point that can be computed by Equation 1, performance counterintuitively decreases with increasing knowledge.

Recognition Dominates Contradicting Information

The use of this simple heuristic can lead to other surprising behavioral results. For instance, the recognition heuristic is a strategy that several species employ for food choice. Wild Norway rats rely on recognition when choosing between two foods: They prefer the one they recognize from having tasted it, or from having smelled it on the breath of a fellow rat. This heuristic is followed even if the fellow rat is sick at the time (Galef, McQuoid, & Whiskin, 1990). That is, recognition dominates illness information. In technical terms, the recognition heuristic is noncompensatory. What is the empirical evidence for the heuristic in humans? In various experiments, typically some 90% of the participants rely on this heuristic in appropriate situations, that is, where recognition is correlated with the criterion (Goldstein & Gigerenzer, 1999). The noncompensation phenomenon that has been reported for rats—they choose the recognized object (e.g., the food smelled on the breath of a fellow rat) despite negative information (the fellow rat is sick)—has also been observed in experiments with humans. The proportion of people who followed the recognition heuristic remained unchanged when they received information that indicated that the recognized city would not be the larger—for instance, that it has no soccer team in the major league (Goldstein & Gigerenzer, 1999). Recognition dominated contradictory information.

Brand Name Recognition

Naturally, if organisms and institutions rely on recognition, from animal foraging and kin recognition to the hiring of star professors, there are also others who exploit this heuristic. Advertising is a case in point. Firms such as Benetton do not waste time describing their product; they just try to increase brand name recognition. Oliviero Toscani, the designer of the Benetton ads, pointed out that already in 1994 the ads had pushed Benetton beyond Chanel into the top five, best-known brand names worldwide (Toscani, 1997), and Benetton's sales increased by a factor of 10. For instance, the Benetton advertising campaign featuring pictures of prison inmates sentenced to death would otherwise make little sense. The recognition heuristic offers a rationale for the Benetton strategy. Consumer behavior relies on name recognition, and this fact can be exploited by firms who increase their name recognition rather than the quality of their products.

Brand name recognition is also relevant to investing in the stock market. If you read the *Wall Street Journal,* you know that experts are often outperformed by randomly selected stocks. Can the recognition heuristic do better than both? To answer this question, one needs sufficiently ignorant people. In a large study, we interviewed several hundred pedestrians in downtown Chicago and downtown Munich and created portfolios from the stocks that 90% of them recognized. In the period investigated, the eight portfolios of U.S. and German stocks chosen by the recognition heuristic outperformed the randomly picked stocks and less recognized stocks, and, in six out of eight cases, also outperformed major mutual funds and the market as a whole (Borges, Goldstein, Ortmann, & Gigerenzer, 1999).

In conclusion, the recognition heuristic is one of the fast and frugal heuristics in the adaptive toolbox. It feeds on an adaptation, the capacity to recognize—face, smell, and name recognition. Face recognition, for instance, is so complex that there is, as yet, no artificial system that can perform as well as a three-year-old child. The recognition heuristic itself, however, is very simple; it can be written in one line of a computer program. The heuristic can exploit ignorance, that is, lack of recognition, and is ecologically rational when recognition is correlated with what needs to be predicted. Ecological rationality defines the domains in which the heuristic works, and those in which it does not.

One-Reason Decision Making

Take The Best

A second heuristic I discuss is Take The Best (Gigerenzer & Goldstein, 1996, 1999). It belongs to the class of *one-reason decision making,* and has the same sequential structure as the heart-attack decision tree. However, the way in which the order of cues is generated is much simpler. The task of Take The Best is to infer, or to predict, which of two objects or alternatives scores higher on a criterion. The recognition heuristic can be the initial step of Take The Best, which illustrates the *nesting* of heuristics in the adaptive toolbox:

Step 0: If applicable, use the recognition heuristic; that is, if only one object is recognized, predict that it has the higher value on the criterion. If both are recognized go on to Step 1.

Step 1: Ordered search: Choose the cue with the highest validity that has not yet been tried for this choice task. Look up the cue values of the two objects.

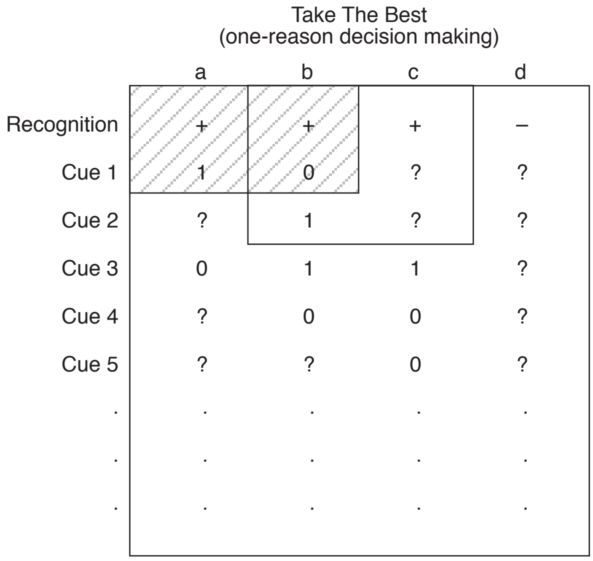

Figure 5. Illustration of Take The Best. Objects *a, b,* and *c* are recognized (+), *d* is not (–). Cue values are binary (0 or 1); missing knowledge is shown by a question mark. For instance, to compare *a* to *b,* Take The Best looks up the values in the lined space and concludes *a* > *b.* To compare *b* to *c,* search is limited to the dotted space and the conclusion is *b* > *c.* The other cue values are not looked up and so are shown within the diagram as shrouded in the fog of memory.

Step 2: Stopping rule: If one object has a positive cue value ("1") and the other does not (that is, either "0" or unknown value), then stop search and go on to Step 3. Otherwise go back to Step 1 and search for another cue. If no further cue is found, then guess.

Step 3: Decision rule: Predict that the object with the positive cue value has the higher value on the criterion.

Figure 5 illustrates how these rules for search, stopping, and decision work. The cue values of the four objects *a, b, c,* and *d* represent the knowledge that an individual can retrieve from searching long-term memory, or, alternatively, from searching in the environment. For simplicity, we treat the cue values as binary, with "1" indicating a higher value on the criterion, "0" indicating a lower value, and "?" representing lack of knowledge about cue values. Not all objects are recognized (*d* is not), thus the recognition heuristic can come into play. Consider first how Take The Best infers which of *a* or *b* scores higher on a criterion. Both objects are recognized, thus the recognition heuristic cannot be used. Take The Best searches in memory (or alternatively, in its external environment) for the value of Cue 1 (Step 1). One is positive, the other is not, thus search is stopped (Step 2) and the inference is made that *a* has the higher criterion value (Step 3). All other cue values of the two objects are ignored, or more precisely, not even searched for. Consider now objects *b* and *c.* Both are recognized, thus the recognition heuristic is again of no use. None of the objects has a positive value on Cue 1, and therefore search continues. On Cue 2, *b* has a positive value ("1") and *c*'s value is unknown, thus search stops and the inference is made that *b* has a higher value than *c.* The information below the dotted area in Figure 5 is ignored, that is, not looked up. Finally, consider *c* and *d*. If there is a positive correlation between recognition and the criterion in this domain, the recognition heuristic applies and the inference is made that *c* has a higher value on the criterion than *d*.

Take The Best is fast (it does not involve much calculation) and frugal (it searches for only part of the information, that is, cues). The ordering of the cues can be learned by a simple but robust criterion that ignores dependencies between cues (Gigerenzer & Goldstein, 1999), or it may be genetically coded, as in mate choice in various animal species (e.g., Dugatkin, 1996).

There is evidence that the Take The Best heuristic is in the toolbox of several species. Female guppies, for instance, choose males on the basis of both physical and social cues, such as bright orange color, large body size, and whether they have observed the male in question mating with another female (Dugatkin, 1996). These cues seem to be organized in a dominance order, as in Figure 5, with the orange-color cue dominating the social cue. If a female has a choice between two males, one of them much more orange than the other, she will choose the more orange one. If the males, however, are close in "orangeness," she prefers the one she has seen mating with another female. Mate choice in guppies illustrates limited search, simple stopping rules, and one-reason decision making. Humans also tend to use this heuristic. Bröder (2000) reported that when the search for information is costly, about 65% of the participants' choices were consistent with Take The Best, compared to fewer than 10% with a linear rule (for similar results, see Rieskamp & Hoffrage, 1999).

Accuracy and Frugality

But how accurate is this heuristic? After all, it does not follow the prescriptions of rational choice theory: It does not look up most of the information, does not calculate an optimal order of cues, does not calculate an optimal stopping point, and relies on one-reason decision making. To answer this question, Czerlinski, Gigerenzer, and Goldstein (1999) tested its predictive accuracy in 20 different situations with varying numbers of cues and varying difficulties of the problem. These situations included: predicting homelessness rates in American cities based on six cues, including rent control and temperature; predicting dropout rates in Chicago public high schools based on 18 cues, such as average SAT Scores and the percentage of low-income students; predicting the mortality rates in U.S. cities based on 15 cues, including pollution levels and the percentage of nonwhites; predicting professors' salaries based on five cues, such as gender and rank; predicting the number of eggs of female Arctic char based on three cues, including each fish's weight and age; and predicting obesity at age 18 from 10 cues measured from age two and older, such as leg circumference and strength. The task for Take The Best was always to predict which of two objects had the higher value on the criterion.

As with the heart-attack decision tree described earlier, the cues were treated as yes/no alternatives, and all cue values and objects were known (i.e., with no "?" values), which excludes the recognition heuristic. Figure 6 illustrates one of these 20 tests predicting which of two American cities had a higher homelessness rate based on six powerful cues. For instance, the best predictor for homelessness was rent control—if there is rent control, homelessness rates tend to be high. In the case of Los Angeles and Chicago, Take The Best stopped search after the first cue, because Los Angeles has rent control and Chicago does not. Take The Best inferred that Los Angeles had the higher homelessness rate, which happens to be correct. When comparing Los Angeles and New York, search is extended until the last cue, and the inference is made that Los Angeles has

	Los Angeles	Chicago	New York	New Orleans
Rent control	1	0	1	0
Vacancy rate	1	0	1	0
Temperature	1	0	1	1
Unemployment	1	1	1	1
Poverty	1	1	1	1
Public housing	1	1	0	0

Figure 6. Predicting which of two American cities has a higher homelessness rate with Take The Best (without recognition and missing data). All cues and 4 out of 50 cities are pictured.

a higher rate, which again is correct. When comparing Chicago and New York, however, Take The Best made an error.

Take The Best is certainly fast and frugal, but is it any good? How close does its predictive accuracy come to that of multiple regression, a linear strategy that uses all predictors, weights them, and combines them? How close does it come to a simpler linear strategy, which also uses all predictors but uses unit weights, that is, +1 or –1, instead of computing the optimal regression weights? We tested the performance of these strategies on 50 American cities and the six predictors shown in Figure 6, using cross-validation, that is, the strategies learned their parameters on half of the data (learning sample), and were tested on the other half (test sample). The surprising result was that Take The Best was more accurate in predicting homelessness than multiple regression and the unit-weight strategy.

Figure 7 shows that this result holds across all 20 problems. The horizontal (*x*) axis shows the frugality of each strategy, that is, the number of cues looked up, and the vertical (*y*) axis shows its predictive accuracy. Take The Best was more frugal than the linear strategies: It searched through only 2.4 cues on average, whereas the linear strategies used all cues, which numbered 7.7 on average. Figure 7 also shows a trade-off region, spanned by the performance of multiple regression. The idea of the trade-off is that if a strategy is more frugal than regression, it has to pay some price in accuracy. Therefore, a more frugal strategy should lie within that region, as indeed one other heuristic, the Minimalist, does. The Minimalist differs from Take The Best only in Step 1. It searches for cues randomly rather than according to an order estimated from the learning set. Take The Best, by contrast, performed outside of the trade-off region. Compared to the two linear strategies, it was both more frugal and more accurate.

Note also that the simple linear strategy (which uses unit weights rather than regression weights) also did slightly better than multiple regression, showing the robustness reported earlier by Dawes and Corrigan (1974). This confirms the counterintuitive finding that the choice of weights, except for their signs, does not matter much. The demonstration that Take The Best outperformed both of these linear strategies is new. This result is stable across various changes in the way the strategies are tested (Czerlinski et al., 1999; Gigerenzer & Goldstein, 1996).

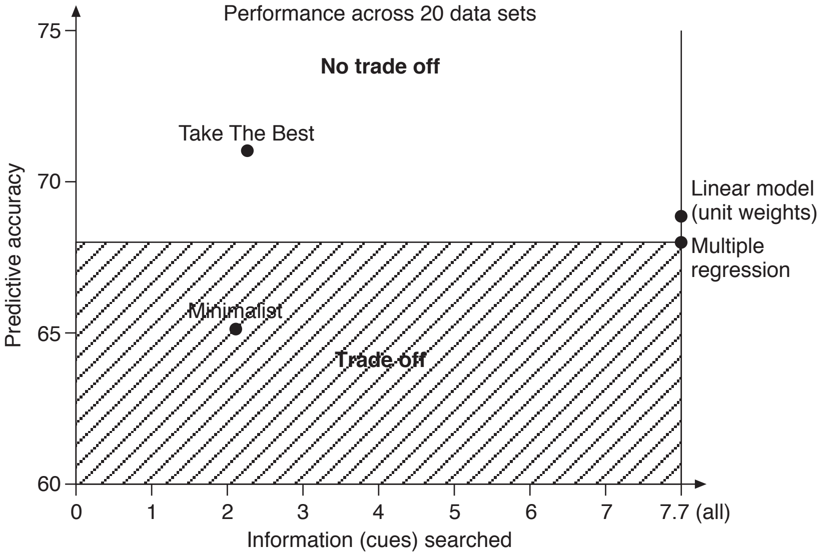

Figure 7. Average accuracy and frugality of Take The Best in predicting a total of 20 criteria, including homelessness, compared to two linear strategies, multiple regression and a simple linear model with unit weights (Czerlinski et al., 1999).

Ecological Rationality

How can one reason be better than many? There are two answers. One is the concept of ecological rationality—that is, the match of a heuristic with the structure of an environment. Figure 8 (left graph) shows one structure that Take The Best can exploit (there are others, see Martignon & Hoffrage, 1999); Figure 8 (right graph) shows one that it cannot. Recall that Take The Best is a noncompensatory strategy. It relies on one cue, and even if all others point in the opposite direction, they cannot compensate. Figure 8 shows examples for noncompensatory and compensatory structures. For instance, binary cues with weights that decrease exponentially, such as 1/2, 1/4, 1/8, and so on, are noncompensatory—the sum of all cue weights to the right of a cue can never be larger than its own weight. When the environment has the same noncompensatory structure as Take The Best, one can prove mathematically that no linear model, including multiple regression, can outperform the faster and more frugal Take The Best (Martignon & Hoffrage, 1999).

The research program on ecological rationality is in the spirit of the earlier ecological programs of Egon Brunswik and J. J. Gibson. Both were studying the structure of environments, although with different tools. Brunswik was looking for the correlational texture of environments and Gibson for invariants in the ambient light. Both were behaviorists. They hesitated to model mental strategies; that is, they did not want to open the "black box." Here the program of ecological validity differs: It studies not just environmental structure, but the degree of match between the heuristics in the adaptive toolbox and the structure of environments (Gigerenzer & Todd, 1999). The black box is a toolbox.

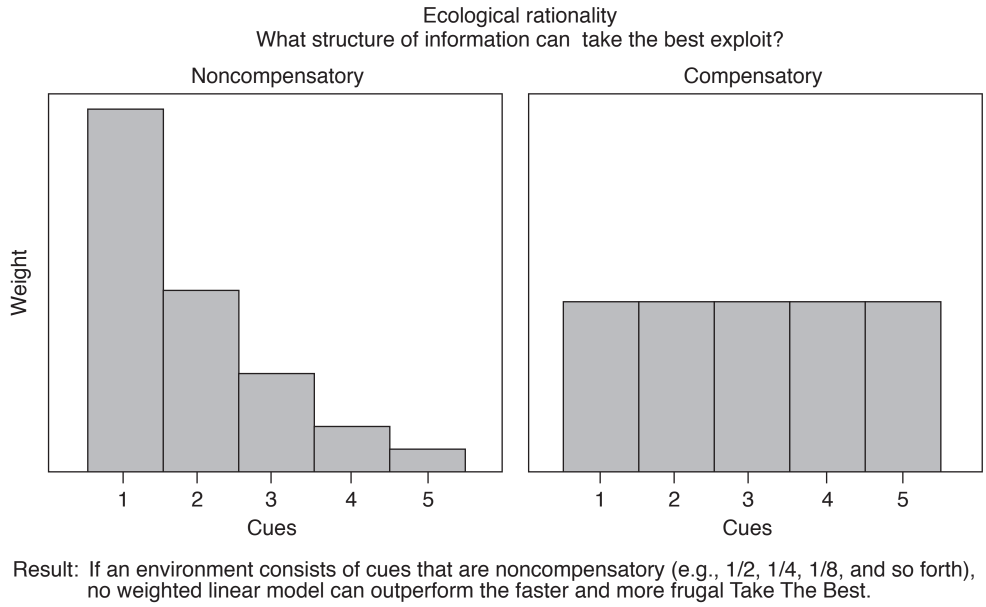

Figure 8. Ecological rationality of Take The Best. One of the structures of environments that Take The Best can exploit is cues with noncompensatory weights, as shown on the left side (Martignon & Hoffrage, 1999).

Robustness

The second answer to the question, "How can one reason be better than many?" is robustness. A strategy is robust to the degree that it can be used in new situations. In a situation where there is uncertainty—and there is, for instance, a lot of uncertainty in predicting homelessness—only part of the information available today will be of predictive value for the future. For instance, if one records the temperature of each day this year in Chicago, one can find a mathematical equation with sufficiently complex exponential terms that represents the jagged temperature curve almost perfectly. However, this equation may not be the best predictor of next year's temperature; a simpler curve that ignores much of this year's measurements may do better. In other words, only part of the information available in one situation generalizes to another. To make good decisions or predictions under uncertainty, one *has* to ignore much of the information available, and the art is to find that part that generalizes. Since Take The Best relies only on the best cue on which the two objects differ, its chances of ignoring less robust information are good.

Note that ecological rationality and robustness are two independent concepts that both explain when simple heuristics work and when they do not. Figure 8 (left) illustrates an environmental structure that Take The Best can exploit, that is, where its use is ecologically rational. This structure makes Take The Best *as* accurate as any linear strategy of whatever complexity. But this noncompensatory structure does not yet explain how Take The Best can actually be *more* accurate than multiple linear regression (for structures that lead to this result, see Martignon &

Hoffrage, 1999). This result follows when we consider that the results in Figure 7 are about how well the strategies predict new data, rather than fit old data. When making predictions about noisy environments, simpler strategies (e.g., with fewer free parameters) tend to be more robust than more complex ones. The details of this relationship are given in Geman et al. (1992) and Forster and Sober (1994). Thus, the example in Figure 8 illustrates when a simple heuristic can be as accurate as any linear strategy (ecological rationality), and the concept of robustness that enters when predictions need to be made in noisy environments explains the additional edge that the heuristic has over strategies that use more knowledge and computational power.

Why Models With The Best Fit Are Not Necessarily The Best

Assume you have several competing models and you want to determine the one that most likely describes the "true" strategy an individual uses. You have a body of data and find that one model fits the data significantly better than the others. You conclude that the empirical evidence supports this model and propose it as the likely actual decision strategy. Isn't this how science works? Not exactly. There are two ways to select a model: to choose the model with the best fit, or the most robust one. Data fitting tells us how well a model can fit *given* data; the generalizability to *new* data is not evaluated. Robustness refers to a situation wherein a model estimates its parameters from a learning sample, but is tested on a new sample, such as in Figure 7. Surprisingly, most research programs in the behavioral and social sciences never proceed beyond data fitting and take a good fit as evidence of the validity of the model tested (Roberts & Pashler, 2000). The same strategy can be observed in animal research, such as when the data on avoidance learning in goldfish is explained by a theory with three equations and six adjustable parameters (Zhuikov, Couvillon, & Bitterman, 1994). However, a good fit by itself is not a good reason to choose between competing models. Why is this?

First, mathematical models with a sufficiently large number of adjustable parameters always lead to an excellent fit—*here, a good fit is a mathematical truism, not an empirical validation of a model.* If a model is too powerful (such as a neural network model with numerous hidden units and adjustable parameters), it can fit any data, even those generated by contradictory underlying processes. These models are largely immune to being falsified. Success in fitting comes at the price of *overfitting,* that is, fitting noise and idiosyncratic parts of the data that do not generalize to new situations. In contrast, fast and frugal heuristics such as the recognition heuristic, Take The Best, and the Minimalist have no adjustable parameters; all concepts such as the recognition validity and the cue validities are empirically measurable. As a consequence, predictions such as in Equation 1 can be proven wrong. In statistical terminology, models of heuristics show "bias" whereas models with numerous adjustable parameters show "variance" (Geman et al., 1992).

Second, from a Darwinian point of view, the program of identifying behavioral strategies by means of data fitting neglects the function of strategies. For an organism, the best strategy (e.g., in foraging or mate search) is not the one that best fits past data in the individual's history or in the evolutionary history of a species. A better strategy is one that predicts future data. In an uncertain world, these two strategies are *not* the same. (In a certain world they would be the same.) To be useful for new situations, a strategy needs to be robust, that is, not to overfit—but the strategy with the best fit is often the one that overfits most. For instance, in data fitting, multiple regression had the best fit across the 20 problems mentioned before, but in predictive accuracy it took the highest loss and was outperformed by simpler and more robust strategies (see Figure 7). Multiple regression overfitted the data.

Can Cognitive Limitations Be Adaptive?

Thus, from an evolutionary point of view, heuristics need to generalize to new situations, not to fit memories of past experiences. This argument leads to an—admittedly speculative—answer to the question "Why did humans and other animals not evolve 'perfect' cognitive functions, such as perfect memory, attention span, and computational skills?" In principle, these abilities might have evolved, as the occasional person with an astonishing memory or computational powers indicates. The answer is that in uncertain environments, precise monitoring and recording of past data is neither necessary nor desirable, because perfect data fitting can be counterproductive. A robust strategy must *ignore* part of the available information. This can be achieved by limited information search, forgetting, or other tools that prevent omniscience. The more uncertain an environment is, the more information that needs to be ignored. The art is to ignore the right information, that is, to pay attention to the proper, powerful cues and forget the rest. Thus, so-called limited information processing capacities can actually be adaptive, not merely a sign of shoddy mental software.

Ecological rationality and robustness are key research tools of a Darwinian approach to decision making. Ecological rationality differs from logical rationality. It defines the reasonableness of a heuristic by its fit to an environmental structure, not by its fit to laws of logic and internal coherence, such as transitivity and additivity of probabilities. However, a glance through today's journals and textbooks on thinking, intelligence, judgment, and decision making reveals that the structure of environments is not part of the investigation (for an exception, see Anderson, 1990). For instance, if an individual ignores relevant cues, ignores the dependencies between cues, and does not even integrate the few cues he or she knows, it is treated in this literature as an illustration of human irrationality. These fallacies are usually attributed to "limited information processing capacities," "confirmation biases," and other shoddy mental software (see the extensive literature on so-called cognitive illusions, e.g., Piattelli-Palmerini, 1994). Individuals who use Take The Best commit all these three "sins." However, as Figure 7 shows, their decisions can actually be more frugal and more accurate than strategies that look rational by traditional standards. A rethinking of rationality is needed—the ecological way.

The Building Blocks of The Adaptive Toolbox

Recombining Building Blocks

The building blocks of the heuristics in the adaptive toolbox include rules for search, stopping search, and decision making. By recombining different building blocks, the adaptive toolbox can create new heuristics. For instance, in a situation in which Take The Best cannot be used because an individual does not have the knowledge to order the cues according to their validity, a less demanding search rule can be used instead that searches for cues in random order or simply tries the cue first that stopped search the last time. This simplification of the search rule results in the Minimalist (see Figure 7) and Take the Last heuristic, respectively (Gigerenzer & Goldstein, 1999). The adaptive toolbox, therefore, has a large number of heuristics at its disposal built from a smaller number of building blocks.

In this chapter, I have described only a few of the heuristics in the adaptive toolbox, and I have focused on heuristics for choice, such as Take The Best. Similar building blocks underlie heuristics for categorization, such as Categorization by Elimination (Berretty et al., 1999) and estimation, such as QuickEst (Hertwig et al., 1999). Simple heuristics for various important adaptive problems have been identified recently, such as how humans infer intentions from movements (Blythe et al., 1999), how honey bees choose a location for a new hive (Seeley, 2001), and how to find a mate without optimization (Miller & Todd, 1999). For an overview of what we know about the adaptive toolbox, see Gigerenzer, Todd, and the ABC Research Group (1999) and Gigerenzer and Selten (2001).

Nesting of Heuristcs

New heuristics can be created not only by recombining building blocks, but also by nesting heuristics. For instance, the recognition heuristic can function as the initial step for Take The Best (Figure 5). The recognition heuristic draws on recognition memory, whereas Take The Best uses recall memory. Recognition memory seems to develop earlier than recall memory both ontogenetically and evolutionarily, and the nesting of heuristics can be seen as analogous to the addition of a new adaptation on top of an existing one. In other words, a heuristic can become a building block of another heuristic.

Emotions and Social Norms

In the examples given in this chapter, the building blocks of heuristics were cognitive, such as recognition and ordered search. However, emotions can also function as building blocks for guiding and stopping search. For instance, falling in love can be a powerful stopping rule that ends search for a partner and strengthens commitment to the loved one. Similarly, feelings of parental love, triggered by one's infant's presence or smile, can be seen as commitment tools, which *prevent* cost-benefit computations with respect to proximal goals, so that the question of whether to endure all the sleepless nights and other challenges associated with baby care never arise. For important adaptive tasks, emotion can be more efficient than cognition (Gigerenzer & Todd, 1999; Tooby & Cosmides, 1990). For instance, the stopping rule in satisfying—stop search after the first person is found that meets or exceeds an aspiration level—does not generate the commitment to a partner that love can. When a new and slightly more attractive partner comes along, nothing prevents the satisficer from leaving her partner on the spot. Emotions, like motivations, are substantially domain-specific and are part of the heuristics in the adaptive toolbox. Social norms can also function as tools for bounded rationality, freeing individuals from making a large number of potential decisions (Cosmides & Tooby, 1992). Building blocks and heuristics can be learned socially through imitation, word of mouth, or cultural heritage—a topic dealt with in Gigerenzer and Selten (2001).

Beyond Demons

In this chapter, I introduced the main concepts for the study of the adaptive toolbox: ecological rationality, frugality, robustness, and the building blocks of heuristics—simple rules for search, stopping, and decision. The underlying vision of rationality is that of domain-specific heuristics that do not involve optimization and are ecologically rational when used in a proper environment.

The perspective of the adaptive toolbox conflicts with several attractive ideals. It conflicts with Laplace's superintelligence and Leibniz's dream of a universal calculus and its modern offspring. For instance, if you open a contemporary textbook on human reasoning and decision making, you will notice the predominance of mental logic, probability theory, and the maximization of expected utility—all attempts at attaining the dream of a universal calculus of reason. Heuristics play a little role, and if they do, it is mainly in the form of vague words that supposedly "explain" errors in logic and probability theory (see Gigerenzer, 1996b). The emphasis on simplicity and transparency conflicts with the preference of many cognitive scientists who are in love with complex mathematical models: The more mathematically sophisticated and nontransparent a model is, the better. For instance, what happens in a neural network is nontransparent, whereas simple heuristics are transparent (Regier, 1996). Finally, simplicity and robustness can conflict with legal values. A doctor who classifies heart attack patients without having measured all variables runs the risk of being sued. Legal systems, like bureaucracies, often run on the defensive vision that more is always better.

The surprising performance of the heuristics—such as the less-is-more effect and the absence of a trade-off between frugality and accuracy—may give us pause and cause us to rethink the notion of bounded rationality. For many, boundaries come from within the human mind—limited capacities for memory, attention, and other constraints within which evolution had to work. However, a Darwinian view would emphasize that the selective forces impinging on our cognitive evolution came largely from outside our minds, from interaction with our physical and social world (Todd, 2001). The notion of ecological rationality provides a framework for understanding the match between heuristics and environment. Simple heuristics are not the shoddy software of a limited mind. Rather, they enable adaptive behavior.

Rational choice theory—the idea that sound decisions are reached by optimization, with or without constraints—has been criticized as *descriptively* inadequate, but maintained as the only *normative* standard. The research program of studying the adaptive toolbox goes one step further. It analyzes how sound decisions can actually be made without omniscience, optimization, or a general logical calculus. Psychological theories need less Aristotle and more Darwin.

References

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on predictions. *Journal of Personality and Social Psychology, 35,* 303–314.

Anderson, J. R. (1990). *The adaptive character of thought.* Hillsdale, NJ: Erlbaum.

Ayton, P., & Önkal, D. (1997). *Forecasting football fixtures: Confidence and judged proportion correct.* Unpublished manuscript.

Berretty, P. M., Todd, P. M., & Martignon, L. (1999). Categorization by elimination: Using few cues to choose. In G. Gigerenzer, P. M. Todd, & the ABC Research Group, *Simple heuristics that make us smart.* New York: Oxford University Press.

Blythe, P. W., Todd, P. M., & Miller, G. E. (1999). How motion reveals intention: Categorizing social interactions. In G. Gigerenzer, P. M. Todd, & the ABC Research Group, *Simple heuristics that make us smart.* New York: Oxford University Press.

Borges, B., Goldstein, D. G., Ortmann, A., & Gigerenzer, G. (1999). Can ignorance beat the stock market? In G. Gigerenzer, P. M. Todd, & the ABC Research Group, *Simple heuristics that make us smart* (pp. 59–72). New York: Oxford University Press.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1993). *Classification and regression trees.* New York: Chapman & Hall.

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.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). New York: Oxford University Press.

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., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81,* 95–106.

Dugatkin, L. A. (1996). Interface between culturally based preferences and genetic preferences: Female mate choice in Poecilia reticulata. *Proceedings of the National Academy of Sciences, 93,* 2770–2773.

Forster, M., & Sober, E. (1994). How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions. *British Journal of Philosophical Science, 45,* 1–35.

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

Geman, S. E., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural Computation, 4,* 1–58.

Gigerenzer, G. (1996a). Rationality: Why social context matters. In P. B. Baltes & U. M. Staudinger (Eds.), *Interactive minds: Lifespan perspectives on the social foundation of cognition* (pp. 319–346). Cambridge, UK: Cambridge University Press.

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

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

Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The Take The Best heuristic. In G. Gigerenzer, P. M. Todd, & the ABC Research Group, *Simple heuristics that make us smart* (pp. 75–96). New York: Oxford University Press.

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

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

Goldstein, D. G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. In G. Gigerenzer, P. M. Todd, & the ABC Research Group, *Simple heuristics that make us smart* (pp. 37–58). New York: Oxford University Press.

Heckhausen, H. (1991). *Motivation and action.* (Peter K. Leppmann, Trans.). Berlin: Springer.

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.* New York: Oxford University Press.

Kelley, H. H. (1973). The process of causal attribution. *American Psychologist, 28,* 107–128.

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.

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

Piatelli-Palmerini, M. (1994). *Inevitable illusions: How mistakes of reason rule our minds.* New York: Wiley.

Regier, T. (1996). *The human semantic potential: Spatial language and constrained connectionism.* Cambridge, MA: MIT Press.

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. 141–168). New York: Oxford University Press.

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review, 107,* 358–367.

Seeley, T. D. (2001). Decision making in superorganisms: How collective wisdom arises from the poorly informed masses. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (pp. 249–261). Cambridge, MA: MIT Press.

Simon, H. A. (1956). Rational choice and the structure of environments. *Psychological Review, 63,* 129–138.

Simon, H. A. (1992). *Economics, bounded rationality, and the cognitive revolution.* Aldershot Hantsuk: Elgar.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy, 69,* 213–225.

Todd, P. (2001). Fast and frugal heuristics for environmentally bounded minds. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (pp. 51–70). Cambridge, MA: MIT Press.

Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology, 11,* 375–424.
Toscani, O. (1997). *Die Werbung ist ein lächelndes Aas.* Frankfurt a.M.: Fischer.
Zhuikov, A. Y., Couvillon, P. A., & Bitterman, M. E. (1994). Quantitative two-process analysis of avoidance conditioning in goldfish. *Journal of Experimental Psychology: Animal Behavior Processes, 20,* 32–43.