Models of Ecological Rationality: The Recognition Heuristic

Daniel G. Goldstein and Gerd Gigerenzer
Max Planck Institute for Human Development

One view of heuristics is that they are imperfect versions of optimal statistical procedures considered too complicated for ordinary minds to carry out. In contrast, the authors consider heuristics to be adaptive strategies that evolved in tandem with fundamental psychological mechanisms. The recognition heuristic, arguably the most frugal of all heuristics, makes inferences from patterns of missing knowledge. This heuristic exploits a fundamental adaptation of many organisms: the vast, sensitive, and reliable capacity for recognition. The authors specify the conditions under which the recognition heuristic is successful and when it leads to the counterintuitive less-is-more effect in which less knowledge is better than more for making accurate inferences.

What are heuristics? The Gestalt psychologists Karl Duncker and Wolfgang Koehler preserved the original Greek definition of "serving to find out or discover" when they used the term to describe strategies such as "looking around" and "inspecting the problem" (e.g., Duncker, 1935/1945). For Duncker, Koehler, and a handful of later thinkers, including Herbert Simon (e.g., 1955), heuristics are strategies that guide information search and modify problem representations to facilitate solutions. From its introduction into English in the early 1800s up until about 1970, the term *heuristics* has been used to refer to useful and indispensable cognitive processes for solving problems that cannot be handled by logic and probability theory (e.g., Polya, 1954; Groner, Groner, & Bischof, 1983).

In the past 30 years, however, the definition of heuristics has changed almost to the point of inversion. In research on reasoning, judgment, and decision making, heuristics have come to denote strategies that prevent one from finding out or discovering correct answers to problems that are assumed to be in the domain of probability theory. In this view, heuristics are poor substitutes for computations that are too demanding for ordinary minds to carry out. Heuristics have even become associated with inevitable cognitive illusions and irrationality (e.g., Piattelli-Palmerini, 1994).

The new meaning of heuristics—poor surrogates for optimal procedures rather than indispensable psychological tools—emerged in the 1960s when statistical procedures such as analysis of variance (ANOVA) and Bayesian methods became entrenched as the psychologist's tools. These and other statistical tools were transformed into models of cognition, and soon thereafter cognitive processes became viewed as mere approximations of statistical procedures (Gigerenzer, 1991, 2000). For instance, when Ward Edwards (1968) and his colleagues concluded that human reasoning did not accord with Bayes's rule (a normative standard for making probability judgments), they tentatively proposed that actual reasoning is like a defective Bayesian computer with wrongly

combined values (misaggregation hypothesis) or misperceived probabilities (misperception hypothesis). The view of cognitive processes as defective versions of standard statistical tools was not limited to Edward's otherwise excellent research program. In the 1970s, the decade of the ANOVA model of causal attribution, Harold Kelley and his colleagues suggested that the mind attributes a cause to an effect in the same way that experimenters draw causal inferences, namely, by computing an ANOVA:

> The assumption is that the man in the street, the naive psychologist, uses a naive version of the method used in science. Undoubtedly, his naive version is a poor replica of the scientific one—incomplete, subject to bias, ready to proceed on incomplete evidence, and so on. (Kelley, 1973, p. 109)

The view that mental processes are "poor replicas" of scientific tools became widespread. ANOVA, multiple regression, first-order logic, and Bayes's rule, among others, have been proposed as optimal or rational strategies (see Birnbaum, 1983; Hammond, 1996; Mellers, Schwartz, & Cooke, 1998), and the term *heuristics* was adopted to account for discrepancies between these rational strategies and actual human thought processes. For instance, the representativeness heuristic (Kahneman & Tversky, 1996) was proposed to explain why human inference is like Bayes's rule with the base rates left out (see Gigerenzer & Murray, 1987). The common procedure underlying these attempts to model cognitive processes is to start with a method that is considered optimal, eliminate some aspects, steps, or calculations, and propose that the mind carries out this naive version.

We propose a different program of cognitive heuristics. Rather than starting with a normative process model, we start with fundamental psychological mechanisms. The program is to design and test computational models of heuristics that are (a) ecologically rational (i.e., they exploit structures of information in the environment), (b) founded in evolved psychological capacities such as memory and the perceptual system, (c) fast, frugal, and simple enough to operate effectively when time, knowledge, and computational might be limited, (d) precise enough to be modeled computationally, and (e) powerful enough to model both good and poor reasoning. We introduce this program of fast and frugal heuristics here with perhaps the simplest of all heuristics: the recognition heuristic.

Correspondence concerning this article should be addressed to Daniel G. Goldstein and Gerd Gigerenzer, Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. E-mail: goldstein@mpib-berlin.mpg.de and gigerenzer@mpib-berlin.mpg.de

In this article, we define the recognition heuristic and study its behavior by means of mathematical analysis, computer simulation, and experiment. We specify the conditions under which the recognition heuristic leads to less-is-more effects: situations in which less knowledge is better than more knowledge for making accurate inferences. To begin, we present two curious findings illustrate the counterintuitive consequences of the recognition heuristic.

Can a Lack of Recognition Be Informative?

In the statistical analysis of experimental data, missing data are an annoyance. However, outside of experimental designs—when data are obtained by natural sampling rather than systematic sampling (Gigerenzer & Hoffrage, 1995)—missing knowledge can be used to make intelligent inferences. We asked about a dozen Americans and Germans, "Which city has a larger population: San Diego or San Antonio?" Approximately two thirds of the Americans correctly responded that San Diego is larger. How many correct inferences did the more ignorant German group achieve? Despite a considerable lack of knowledge, 100% of the Germans answered the question correctly. A similar surprising outcome was obtained when 50 Turkish students and 54 British students made forecasts for all 32 English F. A. Cup third round soccer matches (Ayton & Önkal, 1997). The Turkish participants had very little knowledge about (or interest in) English soccer teams, whereas the British participants knew quite a bit. Nevertheless, the Turkish forecasters were nearly as accurate as the English ones (63% vs. 66% correct).

At first blush, these results seem to be in error. How could more knowledge be no better—or worse—than significantly less knowledge? A look at what the less knowledgeable groups knew may hold the answer. All of the Germans tested had heard of San Diego; however, about half of them did not recognize San Antonio. All made the inference that San Diego is larger. Similarly, the Turkish students recognized some of the English soccer teams (or the cities that often make up part of English soccer team names) but not others. Among the pairs of soccer teams in which they rated one team as completely unfamiliar and the other as familiar to some degree, they chose the more familiar team in 627 of 662 cases (95%). In both these demonstrations, people used the fact that they did not recognize something as the basis for their predictions, and it turned out to serve them well.

The strategies of the German and Turkish participants can be modeled by what we call the *recognition heuristic.* The task to which the heuristic is suited is selecting a subset of objects that is valued highest on some criterion. An example would be to use corporate name recognition for selecting a subset of stocks from Standard and Poor's 500, with profit as the criterion (Borges, Goldstein, Ortmann, & Gigerenzer, 1999). In this article, as in the two preceding laboratory experiments, we focus on the case of selecting one object from two. This task is known as paired comparison or two-alternative forced choice; it is a stock-in-trade of experimental psychology and an elementary case to which many other tasks (such as multiple choice) are reducible.

The recognition heuristic is useful when there is a strong correlation—in either direction—between recognition and criterion. For simplicity, we assume that the correlation is positive. For two-alternative choice tasks, the heuristic can be stated as follows:

Recognition heuristic: If one of two objects is recognized and the other is not, then infer that the recognized object has the higher value with respect to the criterion.

The recognition heuristic will not always apply, nor will it always make correct inferences. Note that the Americans and English in the experiments reported could not apply the recognition heuristic—they know too much. It is also easy to think of instances in which an object may be recognized for having a small criterion value. Yet even in such cases the recognition heuristic still predicts that a recognized object will be chosen over an unrecognized object. The recognition heuristic works exclusively in cases of limited knowledge, that is, when only some objects—not all—are recognized.

The effectiveness of the apparently simplistic recognition heuristic depends on its *ecological rationality:* its ability to exploit the structure of the information in natural environments. The heuristic is successful when ignorance, specifically a lack of recognition, is systematically rather than randomly distributed, that is, when it is strongly correlated with the criterion. (When this correlation is negative, the heuristic leads to the inference that the unrecognized object has the higher criterion value.)

The direction of the correlation between recognition and the criterion can be learned from experience, or it can be genetically coded. The latter seems to be the case with wild Norway rats. Galef (1987) exposed "observer" rats to neighbors that had recently eaten a novel diet. The observers learned to recognize the diet by smelling it on their neighbors' breath. A week later, the observers were fed this diet and another novel diet for the first time and became ill. (They were injected with a nauseant, but as far as they knew it could have been food poisoning.) When next presented with the two diets, the observer rats avoided the diet that they did not recognize from their neighbors' breath. Rats operate on the principle that other rats know what they are eating, and this helps them avoid poisons. According to another experiment (Galef, McQuoid, & Whiskin, 1990), this recognition mechanism works regardless of whether the neighbor rat is healthy or not when its breath is smelled. It may seem unusual that an animal would eat the food its sick neighbor had eaten; however, rats seem to follow recognition without taking further information (such as the health of the neighbor) into account. In this article, we investigate whether people follow the recognition heuristic in a similar noncompensatory way. If reasoning by recognition is a strategy common to humans, rats, and other organisms, there should be accommodations for it in the structure of memory. We explore this question next.

The Capacity for Recognition

As we wander through a stream of sights, sounds, tastes, odors, and tactile impressions, some novel and some previously experienced, we have little trouble telling the two apart. The mechanism for distinguishing between the novel and recognized seems to be specialized and robust. For instance, recognition memory often remains when other types of memory become impaired. Elderly people suffering memory loss (Craik & McDowd, 1987; Schonfield & Robertson, 1966) and patients suffering certain kinds of brain damage (Schacter & Tulving, 1994; Squire, Knowlton, & Musen, 1993) have problems saying what they know about an

object or even where they have encountered it. However, they often know (or can act in a way that proves) that they have encountered the object before. Such is the case with R. F. R., a 54-year-old policeman who developed such severe amnesia that he had great difficulty identifying people he knew, even his wife and mother. One might be tempted to say he had lost his capacity for recognition. Yet on a test in which he was shown pairs of photographs consisting of one famous and one nonfamous person, he could point to the famous persons as accurately as could a healthy control group (Warrington & McCarthy, 1988). His ability to distinguish between the unrecognized people (whom he had never seen before) and the recognized people (famous people he had seen in the media) remained intact. However, his ability to recall anything about the people he recognized was impaired. Laboratory research has demonstrated that memory for mere recognition encodes information even in divided-attention learning tasks that are too distracting to allow more substantial memories to be formed (Jacoby, Woloshyn, & Kelley, 1989). Because recognition continues to operate even under adverse circumstances, and it can be impaired independently from the rest of memory, we view it as a primordial psychological mechanism (for cases involving a selective loss of recognition, see Delbecq-Derousné, Beauvois, & Shallice, 1990).

Because the word *recognition* is used in different ways by different researchers, it needs to be precisely defined. Consider three levels of knowledge an organism may have. First, one may have no knowledge of an object or event because one has never heard, smelled, touched, tasted, or seen it before. Such objects we call "unrecognized." Second are objects one has experienced before but of which one has absolutely no further knowledge beyond this initial sense of recognition. These we will call "merely recognized." (The Scottish verb "to tartle" gives a name to this state of memory; people tartle when they recognize another's face but cannot remember anything else about him or her.) The third level of knowledge comprises mere recognition plus further knowledge; not only does one recognize the object, but one can provide all sorts of additional information about it, such as where one encountered it, what it is called, and so on. Thus, with the term *recognition*, we divide the world into the novel and the previously experienced.

This use of the term needs to be distinguished from another use, which might be characterized as "recognition of items familiar from a list" (e.g., Brown, Lewis, & Monk, 1977). Here the behavior of interest is a person's ability to verify whether a common thing (usually a word such as *house*) had been presented in a previous experimental session. Studies dealing with this meaning of recognition often fail to touch on the distinction between the novel and the previously experienced because the stimuli—mostly numbers or everyday words—are certainly not novel to the participant before the experiment. Experiments that use nonwords or never-seen-before photographs capture the distinction of interest here: that between the truly novel and the previously experienced.

Recognition also needs to be distinguished from notions such as availability (Tversky & Kahneman, 1974) and familiarity (Griggs & Cox, 1982). The availability heuristic is based on recall, not recognition. People recognize far more items than they can recall. Availability is a graded distinction among items in memory and is measured by the order or speed with which they come to mind or the number of instances of categories one can generate. Unlike availability, the recognition heuristic does not address comparisons between items in memory, but rather the difference between items in and out of memory (Goldstein, 1997). The term *familiarity* is typically used in the literature to denote the degree of knowledge (or amount of experience) a person has of a task or object. The recognition heuristic, in contrast, treats recognition as a binary, all-or-none distinction; further knowledge is irrelevant.

A number of studies demonstrate that recognition memory is vast, easily etched on, and remarkably retentive despite short presentation times. Shepard's (1967b) experiment with 612 pairs of novel photographs shown with unlimited presentation time resulted in participants correctly recognizing them with 99% accuracy. This impressive result was made to appear ordinary by Standing's experiments 6 years later. Standing (1973) increased the number of pictures (photographs and "striking" photographs selected for their vividness) to 1,000 and limited the time of presentation to 5 s. Two days later, he tested recognition memory with pairs of pictures in which one had been presented previously and one was novel. Participants were able to point to the previously presented picture 885 times of 1,000 with normal pictures and 940 times of 1,000 with "striking" pictures. Standing then outdid himself by a factor of 10. In perhaps the most extensive recognition memory test ever performed, he presented 10,000 normal pictures for 5 s each. Two days later participants correctly identified them in 8,300 of 10,000 pairs. When more and more pictures were presented, the retention percentage declined slightly, but the absolute number of pictures recognized increased.

Standing's participants must have felt they were making a fair number of guesses. Some research suggests that people may not rely on recognition-based inferences in situations in which they feel their memory is not reliable, such as when presentation times are short or when there are distractions at presentation (Strack & Bless, 1994). However, Standing's results, when adjusted by the usual guessing correction, come out well above chance. With respect to the performance with his "striking" pictures, Standing speculated, "If one million items could be presented under these conditions then 731,400 would be retained" (p. 210). Of course, presenting 1 million items is a feat no experimental psychologist is likely to try.

Recognition memory is expansive, sensitive, and reliable enough to serve in a multitude of inferential tasks. We now discuss how the accuracy of these recognition-based inferences relies not only on the soundness of memory but also on the relationship between recognition and the environment.

Theory

Recognition and the Structure of the Environment

In this article, we investigate recognition as it concerns proper names. Proper name recognition is of particular interest because it constitutes a specialized domain in the cognitive system that can be impaired independently of other language skills (McKenna & Warrington, 1980; Semenza & Sgaramella, 1993; Semenza & Zettin, 1989). Because an individual's knowledge of geography comprises an incomplete set of proper names, it is ideal for recognition studies. In this article, we focus on two situations of limited knowledge: Americans' recognition of German city names and Germans' recognition of city names in the United States. The American students we have tested over the years recognized about

one fifth of the 100 largest German cities, and the German students recognized about one half of the 100 largest U.S. cities. Another reason cities were used to study proper name recognition is because of the strong correlation between city name recognition and population.

It should be clear, however, that the recognition heuristic does not apply everywhere. The recognition heuristic is domain specific in that it works only in environments in which recognition is correlated with the criterion being predicted. Yet, in cases of inference, the criterion is not immediately accessible to the organism. How can correlations between recognition and inaccessible criteria arise? Figure 1 illustrates the ecological rationality of the recognition heuristic. There are "mediators" in the environment that have the dual property of reflecting (but not revealing) the criterion and being accessible to the senses. For example, a person may have no direct information about the endowments of universities, because this information is often confidential. However, the endowment of a university may be reflected in how often it is mentioned in the newspaper. The more often a name occurs in the newspaper, the more likely it is that a person will hear of this name. Because the newspaper serves as a mediator, a person can make an inference about the inaccessible endowment criterion. Three variables reflect the strength of association between the criterion, mediator, and recognition memory: the *ecological correlation,* the *surrogate correlation,* and the *recognition validity.*

The correlation between the criterion and the mediator is called the ecological correlation. In our example, the criterion is the endowment of a university and the mediator variable is the number of times the university is mentioned in the newspaper. The surrogate correlation is that between the mediator (a surrogate for the inaccessible criterion) and the contents of recognition memory, for instance, the number of mentions in the newspaper correlated against recognition of the names mentioned. Surrogate correlations can be measured against the recognition memory of one person (in which case the recognition data will be binary) or against the collective recognition of a group, which we will examine later.
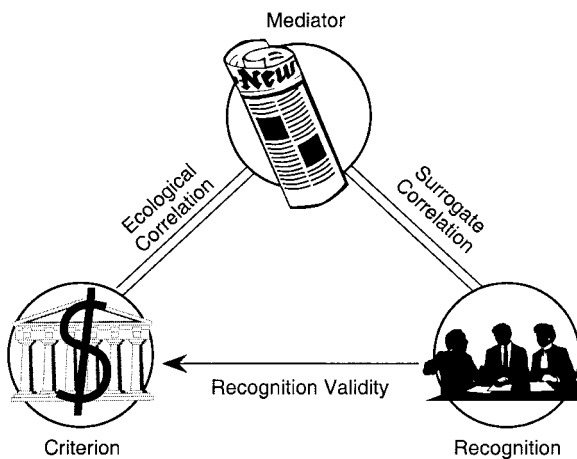

Figure 1. The ecological rationality of the recognition heuristic. An inaccessible criterion (e.g., the endowment of an institution) is reflected by a mediator variable (e.g., the number of times the institution is mentioned in the news), and the mediator influences the probability of recognition. The mind, in turn, uses recognition to infer the criterion.

Figure 1 is reminiscent of Brunswik's lens model (e.g., Hammond, 1996). Recall that the lens model has two parts: environmental and judgmental. Figure 1 can be seen as an elaboration of the environmental part, in which recognition is a proximal cue for the criterion and the recognition validity corresponds to Brunswik's "ecological validity." In contrast to the standard lens model, the pathway from the criterion to recognition is modeled by the introduction of a mediator.

The single most important factor for determining the usefulness of the recognition heuristic is the strength of the relationship between the contents of recognition memory and the criterion. We define the recognition validity as the proportion of times a recognized object has a higher criterion value than an unrecognized object in a reference class. The recognition validity α is thus:

$$\alpha = R/(R + W),$$

where R is the number of correct (right) inferences the recognition heuristic would achieve, computed across all pairs in which one object is recognized and the other is not, and W is the number of incorrect (wrong) inferences under the same circumstances. The recognition validity is essential for computing the accuracy attainable through the recognition heuristic.

Accuracy of the Recognition Heuristic

How accurate is the recognition heuristic? What is the proportion of correct answers one can expect to achieve using the recognition heuristic on two-alternative choice tasks? Let us posit a reference class of N objects and a test whose questions are pairs of objects drawn randomly from the class. Each of the objects is either recognized by the test taker or unrecognized. The test score is the proportion of questions in which the test taker correctly identifies the larger of the two objects.

Consider that a pair of objects must be one of three types: both recognized, both unrecognized, or one recognized and one not. Supposing there are n recognized objects, there are $N - n$ unrecognized objects. This means there are $n(N - n)$ possible pairs in which one object is recognized and the other is not. Similarly, there are $(N - n)(N - n - 1)/2$ pairs in which neither object is recognized. Both objects are recognized in $n(n - 1)/2$ pairs. Dividing each of these terms by the total number of possible pairs, $N(N - 1)/2$, gives the proportion of each type of question in an exhaustive test of all possible pairs.

The proportion correct on an exhaustive test is calculated by multiplying the proportion of occurrence of each type of question by the probability of scoring a correct answer on questions of that type. The recognition validity α, it may be recalled, is the probability of scoring a correct answer when one object is recognized and the other is not. When neither object is recognized, a guess must be made, and the probability of a correct answer is 1/2. Finally, β is the knowledge validity, the probability of getting a correct answer when both objects are recognized. Combining terms in an exhaustive pairing of objects, the expected proportion of correct inferences, $f(n)$, is

$$f(n) =$$

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

(1)

Consider the three parts of the right side of the equation. The term on the left accounts for the correct inferences made by the recognition heuristic. The term in the middle represents the correct inferences resulting from guessing. The right-most term equals the proportion of correct inferences made when knowledge beyond recognition is used. Note that if $n = 0$, that is, no objects are recognized, then all questions will lead to guesses, and the proportion correct will be 1/2. If all objects are recognized ($n = N$), then the left-most two terms become zero and the proportion correct becomes β. The left-most term shows that the recognition heuristic comes into play most under "half ignorance," that is, when the number of recognized objects n equals the number of unrecognized objects $N - n$. (Note, however, that this does not imply that proportion correct will be maximized under these conditions.) Equation 1 specifies the proportion of correct inferences made by using the recognition heuristic whenever possible based on the recognition validity α, the knowledge validity β, and the degree of recognition (n compared with N). Next, we shall see how it leads to a curious state in which less recognition is better than more for making accurate inferences.

The Less-Is-More Effect

Equation 1 seems rather straightforward but has some counterintuitive implications. Consider the following thought experiment: Three Parisian sisters receive the bad news that they have to take a test on the 100 largest German cities at their lycée. The test will consist of randomly drawn pairs of cities, and the task will be to choose the more populous city. The youngest sister does not get out much; she has never heard of Germany (nor any of its cities) before. The middle sister is interested in the conversations of grown-ups and often listens in at her parents' cocktail parties. She recognizes half of the 100 largest cities from what she has overheard in the family salon. The elder sister has been furiously studying for her baccalaureate and has heard of all of the 100 largest cities in Germany. The city names the middle sister has overheard belong to rather large cities. In fact, the 50 cities she recognizes are larger than the 50 cities she does not recognize in about 80% of all possible pairs (the recognition validity α is .8). The middle and elder sisters not only recognize the names of cities but also have some knowledge beyond recognition. When they recognize two cities in a pair, they have a 60% probability of correctly choosing the larger one; that is, the knowledge validity β is .6, whereas $\beta = .5$ would mean they have no useful further knowledge. Which of the three sisters will score highest on the test if they all rely on the recognition heuristic whenever they can? Equation 1 predicts their performance as shown in Figure 2.

The youngest sister can do nothing but guess on every question. The oldest sister relies on her knowledge (β) on every question and scores 60% correct. How well does the middle sister, who has half the knowledge of her older sister, do? Surprisingly, she scores the greatest proportion of correct inferences (nearly 68% correct, according to Equation 1). Why? She is the only one who can use the recognition heuristic. Furthermore, she can make the most of her ignorance because she happens to recognize half of the cities, and this allows her to use the recognition heuristic most often. The recognition heuristic leads to a paradoxical situation in which those who know more exhibit lower inferential accuracy than those
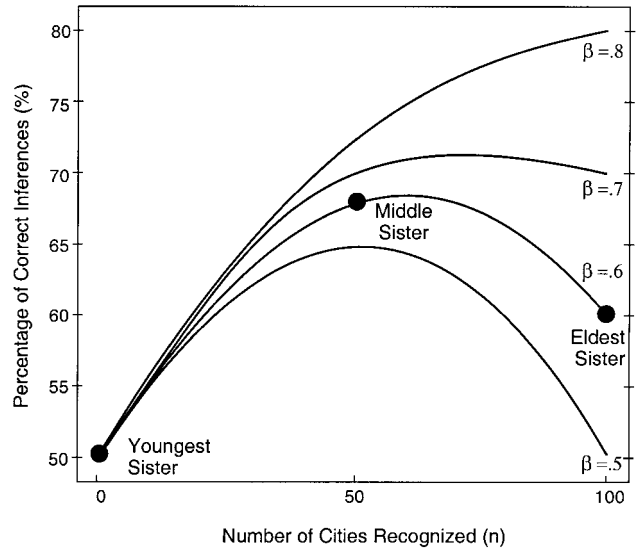

Figure 2. Less-is-more effects illustrated for a recognition validity $\alpha = .8$. When the knowledge validity β is .5, .6, or .7, a less-is-more effect occurs. A β of .5 means that there is no knowledge beyond recognition. When the knowledge validity equals the recognition validity (.8), no less-is-more effect is observed; that is, performance increased with increasing n. The performance of the three sisters is indicated by the three points on the curve for $\beta = .6$. The curve for $\beta = .6$ has its maximum slightly to the right of the middle sister's score.

who know less. We call such situations less-is-more effects. Figure 2 shows how less-is-more effects persist with other values of β.

Forecasting Less-Is-More Effects

In what situations will less-is-more effects arise? We first specify a sufficient condition in idealized environments and then use computer simulation in a real-world environment.

Equation 1 models the accuracy resulting from using the recognition heuristic. The less-is-more effect can be defined as the state of affairs in which the value of $f(n)$ in Equation 1 at some integer from 0 to $N - 1$ (inclusive) is greater than the value at N. For this to happen, the maximum value of $\phi(n)$, the continuous parabola connecting the discrete points, must occur closer to one of the points 0 to $N - 1$ than to point N, that is, between 0 and $N - 1/2$. Solving the equation $\phi'(n) = 0$, when $\phi'(n)$ is simply the first derivative of $\phi(n)$, one locates the maximum of $\phi(n)$ at

$$\frac{-(1 - 2\beta - 2N + 4\alpha N)}{2(1 - 4\alpha + 2\beta)}. \qquad (2)$$

A simple calculation shows that when $\alpha = \beta$, the location of the maximum of $\phi(n)$ is equal to $N - 1/2$: exactly between the $N - 1$st and Nth points. Either increasing α or decreasing β from this point causes the fraction (Equation 2), and thus the location of the maximum, to decrease. From this, we can conclude that there will be a less-is-more effect whenever $\alpha > \beta$, that is, whenever the accuracy of mere recognition is greater than the accuracy achievable when both objects are recognized.

This result allows us to make a general claim. Under the assumption that α and β are constant, any strategy for solving

multiple-choice problems that follows the recognition heuristic will yield a less-is-more effect if the probability α of getting a correct answer merely on the basis of recognition is greater than the probability β of getting a correct answer using more information.

In this derivation, we have supposed that the recognition validity α and knowledge validity β remain constant as the number of cities recognized, n, varies. Figure 2 shows many individuals with different knowledge states but with fixed α and β. In the real world, the recognition and knowledge validities usually vary when one individual learns to recognize more and more objects from experience.[1] Will a less-is-more effect arise in a learning situation in which α and β vary with n, or is the effect limited to situations in which these are constant?

The Less-Is-More Effect: A Computer Simulation

To test whether a less-is-more effect might arise over a natural course of learning, we ran a computer simulation that learned the names of German cities in the order a typical American might come to recognize them. How can one estimate this order? We made the simplifying assumption that the most well-known city would probably be learned about first, then the second-most well-known city, and so on. We judged how well-known cities are by surveying 67 University of Chicago students and asking them to select the cities they recognized from a list. These cities were then ranked by the number of people who recognized them.[2] Munich turned out to be the most well-known city, so our computer program learned to recognize it first. Recognizing only Munich, the program was then given an exhaustive test consisting of all pairs of cities with more than 100,000 inhabitants (the same 83 cities used in Gigerenzer & Goldstein, 1996). Next, it learned to recognize Berlin (the second-most well-known city) to make a total of two recognized cities and was tested on all possible pairs again. It learned and was tested on city after city until it recognized all of them. In one condition, the program learned only the names of cities and made all inferences by the recognition heuristic alone. This result is shown as the bottom curve in Figure 3 labeled *No Cues*. When all objects were unrecognized or all were recognized, performance was at a chance level. Over the course of learning, an inverse ∪ shape, like that in Figure 2, reappears. Here the less-is-more curve is jagged because, as mentioned, the recognition validity was not set to be a constant but was allowed to vary freely as more cities were recognized.

Would the less-is-more effect disappear if the program learned not just the names of cities but information useful for predicting city populations as well? In other words, if there is recall of relevant facts, will this override the less-is-more effect and, therefore, the recognition heuristic? In a series of conditions, the program learned the name of each city, along with one, two, or nine predictive cues for inferring population (the same cues as in Gigerenzer & Goldstein, 1996). In the condition with one cue, the program learned whether each of the cities it recognized was once an exposition site, a predictor with a high ecological validity (.91; see Gigerenzer & Goldstein, 1996). The program then used a decision strategy called Take The Best (Gigerenzer & Goldstein, 1999) to make inferences about which city is larger. Take The Best is a fast and frugal strategy for drawing inferences from cues that uses the recognition heuristic as the first step. It looks up cues in
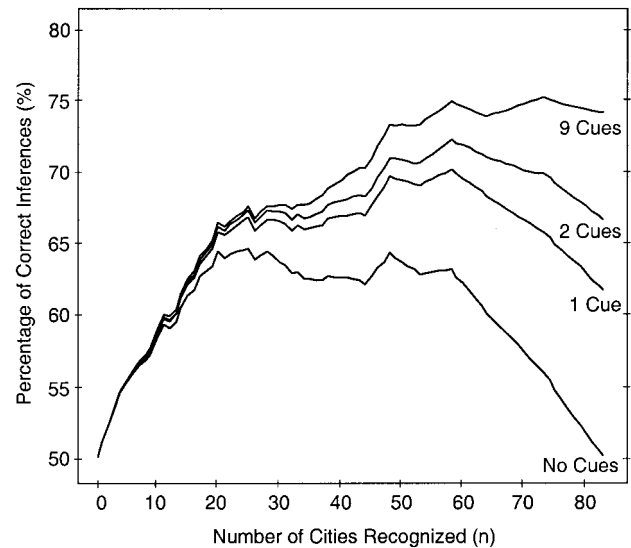

Figure 3. Less-is-more effects when the recognition validity is not held constant but varies empirically, that is, as cities become recognized in order of how well known they are. Inferences are made on recogniton alone (no cues) or with the aid of 1, 2, or 9 predictive cues and Take The Best (see Goldstein & Gigerenzer, 1999).

the order of their validity and stops search as soon as positive evidence for one object (e.g., the city was an exposition site) but not for the other object is found. It is about as accurate as multiple regression for this task (Gigerenzer & Goldstein, 1999).

Does adding predictive information about exposition sites wash out the less-is-more effect? With one cue, the peak of the curve shifted slightly to the right, but the basic shape persisted. When the program recognized more than 58 cities, including information about exposition sites, the accuracy still went down. In the condition with two cues, the program learned the exposition site information and whether each city had a soccer team in the major league (another cue with a high validity, .87). The less-is-more effect was lessened, as is to be expected when adding recall knowledge to recognition, but still pronounced. Recognizing all

---

[1] When there are many different individuals who recognize different numbers of objects, as with the three sisters, it is possible that each individual has roughly the same recognition validity. For instance, for the University of Chicago students we surveyed (Gigerenzer & Goldstein, 1996), recognition validity in the domain of German cities was about 80% regardless of how many cities each individual recognized. However, when one individual comes to recognize more and more objects, the recognition validity can change with each new object recognized. Assume a person recognizes n of N objects. When she learns to recognize object $n + 1$, this will change the number of pairs for which one object is recognized and the other is not from $n(N - n)$ to $(n + 1)(N - n - 1)$. This number is the denominator of the recognition validity (the number $R + W$ of correct plus incorrect inferences), which consequently changes the recognition validity itself. Learning to recognize new objects also changes the numerator (the number R of correct inferences); in general, if the new object is large, the recognition validity will go down, and if it is small, it will go up.

[2] If several cities tied on this measure, they were ordered randomly for each run of the simulation.

cities and knowing all the information contained in two cues (the far right-hand point) resulted in fewer correct inferences than recognizing only 23 cities. Finally, we tested a condition in which all nine cues were available to the program, more information for predicting German city populations than perhaps any German citizen knows. We see the less-is-more effect finally flattening out; however, it does not go away completely. Even when all 83 cue values are known for each of nine cues and all cities are recognized, the point on the far right is still lower than 24 other points on that curve. A beneficial amount of ignorance can enable even higher accuracy than extensive knowledge can.

To summarize, the simplifying assumption that the recognition validity α and knowledge validity β remain constant is not necessary for the less-is-more effect to arise: It held as α and β varied in a realistic way. Moreover, the counterintuitive effect even appeared when complete knowledge about several predictors was present. The effect appears rather robust in theory and simulation. Will the recognition heuristic and, thus, the less-is-more effect emerge in human behavior?

Does the Recognition Heuristic Predict People's Inferences?

It could be that evolution has overlooked the inferential ease and accuracy the recognition heuristic affords. Do human judgments actually accord with the recognition heuristic? We test this question in an experiment in which people make inferences from limited knowledge.

Method

The participants were 22 students from the University of Chicago who were native speakers of English and who had lived in the United States for the last 10 years. They were given all the pairs of cities drawn from the 25 ($n = 6$) or 30 ($n = 16$) largest cities in Germany, which resulted in 300 or 435 questions for each participant, respectively. The task was to choose the larger city in each pair. Furthermore, each participant was asked to check the names of the cities he or she recognized off of a list. Half of the participants took this recognition test before the experiment and half after. (Order turned out to have no effect.) From this recognition information, we calculated how often each participant had an opportunity to choose in accordance with the recognition heuristic and compared this number with how often they actually did. If people use the recognition heuristic, they should predominantly choose recognized cities over unrecognized ones. If they use a compensatory strategy that takes more information into account, this additional information may often suggest not choosing the recognized city. After all, a city can be recognized for reasons that have nothing to do with population.

Results

Figure 4 shows the results for the 22 individual participants. For each participant, one bar is shown. The bar shows the proportion of judgments that agreed with the recognition heuristic among all cases in which it could be applied. For example, participant A had 156 opportunities to choose according to the recognition heuristic and did so every time, participant B did so 216 of 221 times, and so on. The proportions of recognition heuristic adherence ranged between 73% and 100%. The mean proportion of inferences in accordance with the recognition heuristic was 90% (median 93%).
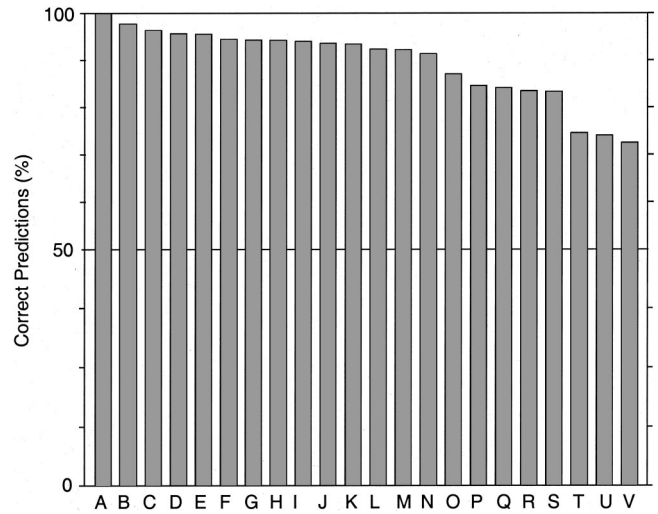
Figure 4. Recognition heuristic accordance. For each of 22 participants, the bars show the percentage of inferences consistent with the recognition heuristic. The individuals are ordered from left to right according to how often their judgments agreed with the recognition heuristic.

This simple test of the recognition heuristic showed that it captured the vast majority of inferences.

Test Size Influences Performance

Equation 1 allows for a novel prediction: The number of correct inferences depends in a nonmonotonic but systematic way on the number of cities N included in the test. That is, for constant recognition validity α and knowledge validity β the test size N (and n that depends on it) predicts various proportions of correct answers. Specifically, Equation 1 predicts when the deletion or addition of one object to the test set should decrease or increase performance.

Method

Using Equation 1, we modeled how the accuracy of the participants in the preceding experiment would change if they were tested on various numbers of cities and test these predictions against the participants' demonstrated accuracy. These predictions were made using nothing but information about which cities people recognize; no parameters are fit. With the data from the previous experiment, we looked at the participants' accuracy when tested on the 30 largest cities, then on just the 29 largest cities, and so on, down to the 2 largest. (This was done by successively eliminating questions and rescoring.) For each participant, at each test size, we could compute the number of objects they recognized and the recognition validity α from the recognition test. Assuming, for simplicity, that the knowledge validity β always was a dummy value of .5, we used Equation 1 to predict the change in the proportion of correct inferences when the number of cities N on the test varied.

Results

The average predictions for and the average actual performance of the individuals who took the exhaustive test on the 30 largest cities are shown in Figure 5. There are 28 predicted changes (up or down), and 26 of these 28 predictions match the data. Despite the
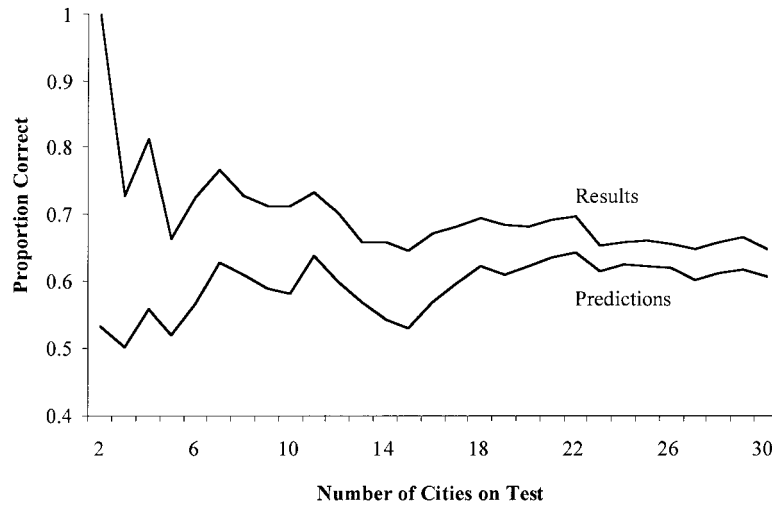
Figure 5. The use of the recognition heuristic implies that accuracy depends on test size (*N*) in an irregular but predictable way (Equation 1). The predictions were made with recognition information alone, and no parameters were fit. The knowledge validity used was a dummy value that assumes people will guess when both cities are recognized. The predictions mirror the fluctuations in accuracy in 26 of the 28 cases.

apparent irregularity of the actual changes, Equation 1 predicts them with great precision. Note that there are no free parameters used to fit the empirical curve.

An interesting feature of Figure 5 is the vertical gap between the two curves. This difference reflects the impact of the knowledge validity β, which was set at the dummy value of .5 for demonstrative purposes. The reason this difference decreases with increasing *N* is that, as the tests begin to include smaller and smaller cities, the true knowledge validity β of our American participants tended toward .5. That is, it is safe to assume that when Americans are tested on the 30 (or more) largest German cities, their probability of correctly inferring which of two recognized cities is larger is only somewhat better than chance.

To summarize, the recognition heuristic predicts how performance changes with increasing or decreasing test size. The empirical data showed a strong, nonmonotonic influence on performance, explained almost entirely by the recognition heuristic.

Noncompensatory Inferences: Will Inference Follow the Recognition Heuristic Despite Conflicting Evidence?

People often look up only one or two relevant cues, avoid searching for conflicting evidence, and use noncompensatory strategies (e.g., Einhorn, 1970; Einhorn & Hogarth, 1981, p. 71; Fishburn, 1974; Hogarth, 1987; Payne, Beltman, & Johnson, 1993; Shepard, 1967a). The recognition heuristic is a noncompensatory strategy: If one object is recognized and the other is not, then the inference is determined; no other information about the recognized object is searched for and, therefore, no other information can reverse the choice determined by recognition. Noncompensatory judgments are a challenge to traditional ideals of rationality because they dispense with the idea of compensation by integration. For instance, when Keeney and Raiffa (1993) discussed lexicographic strategies—the prototype of noncompensatory rules—they

repeatedly inserted warnings that such a strategy "is more widely adopted in practice than it deserves to be" because "it is naively simple" and "will rarely pass a test of 'reasonableness'"(pp. 77–78). The term *lexicographic* means that criteria are looked up in a fixed order of validity, like the alphabetic order used to arrange words in a dictionary. Another example of a lexicographic structure is the Arabic (base 10) numeral system. To decide which of two numbers (with an equal number of digits) is larger, one looks at the first digit. If the first digit of one number is larger, then the whole number is larger. If the first digits are equal, then one looks at the second digit, and so on. This simple method is not possible for Roman numerals, which are not lexicographic. Lexicographic strategies are noncompensatory because the decision made on the basis of a cue higher up in the order cannot be reversed by the cues lower in the order. The recognition heuristic is possibly the simplest of all noncompensatory strategies: It only relies on subjective recognition and not on objective cues. In this section, we are not concerned with the "reasonableness" of its noncompensatory nature (which we have analyzed in earlier sections in terms of α, β, *n*, and *N*), but with the descriptive validity of this property. Would the people following the recognition heuristic still follow it if they were taught information that they could use to contradict the choice dictated by recognition?

In this experiment, we used the same task as before (inferring which of two cities has the larger population) but taught participants additional, useful information that offered an alternative to following the recognition heuristic, in particular, knowledge about which cities have soccer teams in the major league (the German "Bundesliga"). German cities with such teams tend to be quite large, so the presence of a major league soccer team indicates a large population. Because of this relationship, we can test the challenging postulate of whether the recognition heuristic is used in a noncompensatory way. Which would participants choose as larger: an unrecognized city or a recognized city that they learned has *no* soccer team?

Method

Participants were 21 students from the University of Chicago. All were native English speakers who had lived in the United States for at least the last 10 years. The experiment consisted of a training session and a test session. At the beginning of the training session, participants were instructed to write down everything they were taught, and they were informed that after training they would be given a test consisting of pairs of cities drawn from the 30 largest in Germany. During the training session, participants were taught that 9 of the 30 largest cities in Germany have soccer teams and that the 9 cities with teams are larger than the 21 cities without teams in 78% of all possible pairs. They were also taught the names of 4 well-known cities that have soccer teams as well as the names of 4 well-known cities that do not. When they learned about these 8 cities, they believed they had drawn them at random from all 30 cities; in actuality, the computer program administering the experiment was rigged to present the same information to all participants. After the training, participants were asked to recall everything they had been taught without error. Those who could not do so had to repeat training until they could.

After participants passed the training phase, they were presented pairs of cities and asked to choose the larger city in each pair. Throughout this test, they could refer to their notes about which cities do and do not have soccer teams. To motivate them to take the task seriously, they were offered a chance of winning $15 if they scored more than 80% correct. To reiterate, the point of the experiment was to see which city the participants would choose as the larger one: a city they had never heard of before, or one that they had recognized beforehand but had just learned had no soccer team. From the information presented in the training session (which did not make any mention of recognition), one would expect the participants to choose the unrecognized city in these cases. The reason for this is as follows. An unrecognized city either does or does not have a soccer team. If it does (a 5 in 22 chance from the information presented), then there is a 78% chance that it is larger. If it does not, then soccer team information is useless and a guess must be made. Any chance of the unrecognized city having a soccer team suggests that it is probably larger. Participants who do not put any value on recognition should always choose the unrecognized city.

Note that the role of recognition or the recognition heuristic was never mentioned in the experiment. All instruction concerned soccer teams. The demand characteristics in this experiment would suggest that, after passing the training session requirements, the participants would use the soccer team information for making the inferences.

Results

The test consisted of 66 pairs of cities. Of these, we were only interested in 16 critical pairs that contained one unrecognized city and one recognized city that does not have a soccer team. Before or after this task, we tested which cities each participant recognized (order had no effect). In those cases in which our assumptions about which cities the participants recognized were contradicted by the recognition test, items were eliminated from the analysis, resulting in fewer than 16 critical pairs.[3]

Figure 6 reads the same as Figure 4. Twelve of 21 participants made choices in accordance with the recognition heuristic without exception, and most others deviated on only one or two items. All in all, inferences accorded with the recognition heuristic in 273 of the 296 total critical pairs. Despite the presence of conflicting knowledge, the mean proportion of inferences agreeing with the heuristic was 92% (median 100%). These numbers are even a bit higher than in the previous study, which, interestingly, did not involve the teaching of contradictory information.

It appears that the additional information about soccer teams was not integrated into the inferences, consistent with the recog-
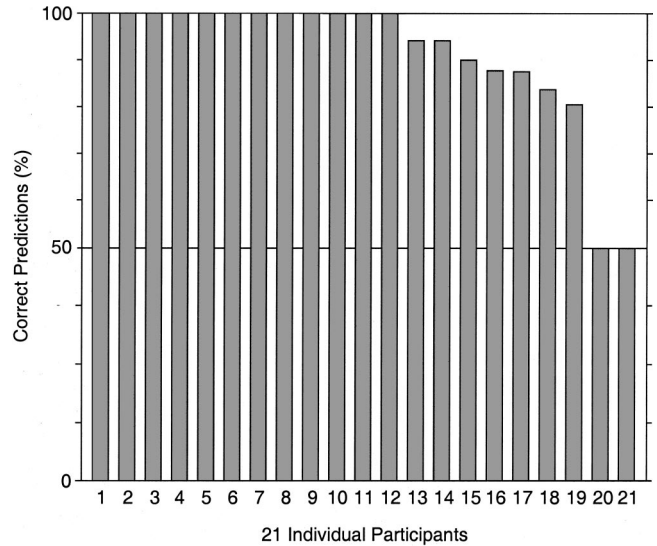

Figure 6. Recognition heuristic adherence despite training to encourage the use of information other than recognition. The bars show the percentage of inferences consistent with the recognition heuristic. The individuals are ordered from left to right by recognition heuristic accordance.

nition heuristic. This result supports the hypothesis that the recognition heuristic was applied in a noncompensatory way.

Will a Less-Is-More Effect Occur Between Domains?

A less-is-more effect can emerge in at least three different situations. First, it can occur between two groups of people, when a more knowledgeable group makes systematically worse inferences than a less knowledgeable group in a given domain. An example was the performance of the American and German students on the question of whether San Diego or San Antonio is larger. Second, a less-is-more effect can occur between domains, that is, when the same group of people achieves higher accuracy in a domain in which they know little than in a domain in which they know a lot. Third, a less-is-more effect can occur during knowledge acquisition, that is, when the same group or individual makes more erroneous inferences as a result of learning. In this and the following experiment, we attempt to demonstrate less-is-more effects of the latter two types, starting with a less-is-more effect between domains.

The mathematical and simulation results presented previously show that less-is-more effects emerge under the conditions specified. However, the curious phenomenon of a less-is-more effect is harder to demonstrate with real people than by mathematical proof or computer simulation. The reason is that real people do not always need to make inferences under uncertainty; they sometimes

---

[3] Another precaution we took concerned the fact that unrecognized cities are often smaller than recognized ones, and this could work to the advantage of the recognition heuristic in this experiment. How can one tell whether people are following the recognition heuristic or choosing correctly by some other means? To prevent this confusion, the critical test items were designed so that the unrecognized cities were larger than the recognized cities in half of the pairs.

have definite knowledge and can make deductions (e.g., if they know for certain that New York is the largest American city, they will conclude that every other city is smaller). For this reason, even if there is a between-domains less-is-more effect for all items about which an inference must be made, this effect may be hidden by the presence of additional, definite knowledge.

In the following experiment, American participants were tested on their ability to infer the same criterion, population, in two different domains: German cities and American cities. Naturally, we expected the Americans to have considerably more knowledge about their own country than about Germany. Common sense (and all theories of knowledge of which we are aware) predicts that participants will make more correct inferences in the domain about which they know more. The recognition heuristic, however, could pull performance in the opposite direction, although its effect will be counteracted by the presence of certain knowledge that the Americans have about cities in the United States. Could the test scores on the foreign cities nevertheless be nearly as high as those on the domestic ones?

Method

Fifty-two University of Chicago students took two tests each: one on the 22 largest cities in the United States, and one on the 22 largest cities in Germany. The participants were native English speakers who had lived the preceding 10 years in the United States. Each test consisted of 100 pairs of randomly drawn cities, and the task was to infer which city is the larger in each pair. Half the subjects were tested on the American cities first and half on the German cities first. (Order had no effect.)

As mentioned, the curious phenomenon of a less-is-more effect is harder to demonstrate with real people than on paper because of definite knowledge. For instance, many Americans, and nearly all of the University of Chicago students, can name the three largest American cities in order. Knowing only the top three cities and guessing on questions that do not involve them led to 63% correct answers without making any inferences, only deductions. Knowing the top five in order yields 71% correct. No comparable knowledge can be expected for German cities. (Many of our participants believed that Bonn is the largest German city; it is 23rd.) Thus, the demonstration of a less-is-more effect is particularly difficult in this situation because the recognition heuristic only makes predictions about uncertain inference, not about the kinds of definite knowledge the Americans had.

Results

The American participants scored a mean 71.1% (median 71.0%) correct on their own cities. On the German cities, the mean accuracy was 71.4% (median 73.0%) correct. Despite the presence of definite knowledge about the American cities, the recognition heuristic still caused a slight less-is-more effect. For half of the participants, we kept track of which cities they recognized, as in previous experiments. For this group, the mean proportion of inferences in accordance with the recognition heuristic was 88.5% (median 90.5%). The recognition test showed that participants recognized a mean of 12 German cities, roughly half of the total, which indicates that they were able to apply the recognition heuristic nearly as often as theoretically possible (see Equation 1).

In a study that is somewhat the reverse of this one, a less-is-more effect was demonstrated with German students who scored higher when tested on American cities than on German ones (Hoffrage, 1995; see also Gigerenzer, 1993).

Despite all the knowledge—including certain knowledge—the Americans had about their own cities, and despite their limited knowledge about Germany, they could not make more accurate inferences about American cities than about German ones. Faced with German cities, the participants could apply the recognition heuristic. Faced with American cities, they had to rely on knowledge beyond recognition. The fast and frugal recognition heuristic exploited the information inherent in a lack of knowledge to make inferences that were slightly more accurate than those achieved from more complete knowledge.

Will a Less-Is-More Effect Occur as Recognition Knowledge Is Acquired?

"A little learning is a dangerous thing," warned Alexander Pope. The recognition heuristic predicts cases in which increases in knowledge can lead to decreases in inferential accuracy. Equation 1 predicts that if $\alpha > \beta$, the proportion of accurate inferences will increase up to a certain point when a person's knowledge increases but thereafter decrease because of the diminishing applicability of the recognition heuristic. This study aims to demonstrate that less-is-more effects can emerge over the course of time as ignorance is replaced with recognition knowledge.

The design of the experiment was as follows. German participants came to the laboratory four times and were tested on American cities. As they were tested repeatedly, they may have gained what we lightheartedly call an "experimentally induced" sense of recognition for the names of cities they had not recognized before the experiment. This induced recognition is similar to that generated in the "overnight fame" experiments by Jacoby, Kelley, Brown, and Jasechko (1989), in which mere exposure caused nonfamous names to be judged as famous. Can mere exposure to city names cause people to infer that formerly unrecognized cities are large? If so, this should cause accuracy on certain questions to drop. For instance, in the first session, a German who has heard of Dallas but not Indianapolis, would correctly infer that Dallas is larger. However, over the course of repeated testing, this person may develop an experimentally induced sense of recognition for Indianapolis without realizing it. Recognizing both cities, she becomes unable to use the recognition heuristic and may have to guess.

It is difficult to produce the counterintuitive effect that accuracy will decrease as city names are learned because it is contingent on several assumptions. The first is that recognition will be experimentally induced, and there is evidence for this phenomenon in the work by Jacoby, Kelley, et al. (1989) and Jacoby, Woloshyn, and Kelley (1989). The second assumption is that people use the recognition heuristic, and there is evidence for this in the experimental work reported here. The third assumption is that, for these participants, the recognition validity α is larger than the knowledge validity β: a necessary condition for a less-is-more effect. The simulation depicted in Figure 3 indicates that this condition might hold for certain people.

Method

Participants were 16 residents of Munich, Germany, who were paid for their participation. (They were not paid, however, for the correctness of their answers because this would have encouraged them to do research on

populations between sessions). In the first session, after a practice test to get used to the computer, they were shown the names of the 75 largest American cities in random order. (The first three, New York, Los Angeles, and Chicago, were excluded because many Germans know they are the three largest). For each city, participants indicated whether they had heard of the city before the experiment, and then, to encourage the encoding of the city name in memory, they were asked to write the name of each city as it would be spelled in German. Participants were not informed that the cities were among the largest in the country, only that they were American cities. They were then given a test consisting of 300 pairs of cities, randomly drawn for each participant, and asked to choose the larger city in each pair.

About 1 week later, participants returned to the lab and took another test of 300 pairs of American cities randomly drawn for each participant. The third week was a repetition of the second week. The fourth week was the critical test. This time, participants were given 200 carefully selected questions. There were two sets of 100 questions each that were used to test two predictions. The first set of 100 was composed of questions taken from the first week's test. This set was generated by listing all questions from the first session's test in which one city was recognized and the other not (according to each participant's recognition in the first week) and randomly drawing (with replacement) 100 times. We looked at these repeated questions to test the prediction that accuracy will decrease as recognition knowledge is acquired.

A second set of 100 questions consisted of pairs with one "experimentally induced" city (i.e., a city that was unrecognized before the first session but may have become recognized over the course of repeated testing) and a new, unrecognized city introduced for the first time in the fourth session. All new cities were drawn from the next 50 largest cities in the United States, and a posttest recognition survey was used to verify whether they were novel to the participants. (Participants, however, did not know from which source any of the cities in the experiment were drawn.)

Which would people choose: an experimentally induced city that they had learned to recognize in the experiment or a city they had never heard of before? If people use the recognition heuristic, this choice should not be random but should show a systematic preference for the experimentally induced cities. This set of questions was introduced to test the hypothesis that recognition information acquired during the experiment would be used as a surrogate for genuine recognition information.

Results

If the assumptions just specified hold, one should observe that the percentage of times experimentally induced cities are inferred to be larger than recognized cities should increase and, subsequently, cause accuracy to decrease.

In the first week, participants chose unrecognized cities over recognized ones in 9.6% of all applicable cases, consistent with the proportions reported in Studies 1 and 3. By the fourth week, the experimentally induced cities were chosen over those that were recognized before the first session 17.2% of the time (Table 1). Participants' accuracy on the 100 repeated questions dropped from a mean of 74.8% correct (median 76%) in the first week to 71.3% correct (median 74%) in the fourth week, $t(15) = 1.82$, $p = .04$, one-tailed. To summarize, as unrecognized city names were presented over 4 weeks, participants became more likely to infer that these cities were larger than recognized cities. As a consequence, accuracy dropped during this month-long experiment. Surprisingly, this occurred despite the participants having ample time to think about American cities and their populations, to recall information from memory, to ask friends or look in reference books for correct answers, or to notice stories in the media that could inform their inferences.

Table 1
A Less-Is-More Effect Resulting From Learning

Session	Mean % correct	Median % correct	Inferences (%)[a]
1	74.8	76	9.6
4	71.3	74	17.2

Note. The percentage of correct answers drops from the first to the fourth sessions. As German participants saw the same novel American city names over and over again in repeated testing, they began to choose them over cities that they recognized from the real world (column 4).
[a] Represents inferences in which an experimentally induced city was chosen over one recognized from before the experiment.

How did participants make inferences in the second set of 100 questions, in which each question consisted of two unrecognized cities: one new to the fourth session and one experimentally induced? If the repeated presentation of city names had no effect on inferences, participants would be expected to choose both types of cities equally often (about 50% of the time) in the fourth session. However, in the fourth session, participants chose the experimentally induced cities over the new ones 74.3% of the time (median 77%). Recognition induced in the laboratory had a marked effect on the direction of people's inferences.

The Ecological Rationality of Name Recognition

What is the origin of the recognition heuristic as a strategy? In the case of avoiding food poisoning, organisms seem as if they are genetically prepared to act in accordance with the recognition heuristic. Wild Norway rats do not need to be taught to prefer recognized foods over novel ones; food neophobia is instinctual (Barnett, 1963). Having such a strategy as an instinct makes adaptive sense: If an event is life threatening, organisms needing to learn to follow the recognition heuristic will most likely die before they get the chance. Learning the correlation between name recognition and city size is not an adaptive task. How did the association between recognition and city population develop in the minds of the participants we investigated?

The recognition validity—the strength of the association between recognition and the criterion—can be explained as a function of the ecological and the surrogate correlations that connect an unknown environment with the mind by means of a mediator (see Figure 1). If the media are responsible for the set of proper names we recognize, the number of times a city is mentioned in the newspapers should correlate with the proportion of readers who recognize the city; that is, there should be a substantial surrogate correlation. Furthermore, there should be a strong ecological correlation (larger cities should be mentioned in the news more often). To test this postulated ecological structure, we analyzed two newspapers with large readerships: the *Chicago Tribune* in the United States and *Die Zeit* in Germany.

Method

Using an Internet search tool, we counted the number of articles published in the *Chicago Tribune* between January 1985 and July 1997 in which the words *Berlin* and *Germany* were mentioned together. There were 3,484. We did the same for all cities in Germany with more than

100,000 inhabitants (there are 83 of them) and checked under many possible spellings and misspellings. Sixty-seven Chicago residents were given a recognition test in which they indicated whether they had heard of each of the German cities. The proportion of participants who recognized a given city was the city's recognition rate.

The same analysis was performed with a major German language newspaper, *Die Zeit*. For each American city with more than 100,000 inhabitants, the number of articles was counted in which it was mentioned. The analysis covered the period from May 1995 to July 1997. Recognition tests for the American cities were administered to 30 University of Salzburg students (Hoffrage, 1995).

Results

Three measures were obtained for each city: the actual population, the number of mentions in the newspaper, and the recognition rate. Figure 7 illustrates the underlying structure and shows the correlations between these three measures. For the class of all German cities with more than 100,000 inhabitants, the ecological correlation (i.e., the correlation between the number of newspaper articles in the *Chicago Tribune* mentioning a city and its actual population) was .70; the number of times a city is mentioned in the newspaper is a good indicator of the city's population. The surrogate correlation, that is, the correlation between the number of newspaper articles about a city and the number of people recognizing it was .79; the recognition rates are more closely associated with the number of newspaper articles than with the actual populations. This effect is illustrated by large German and American cities that receive little newspaper coverage, such as Essen, Dortmund, and San Jose. Their recognition rates tend to follow the low frequency of newspaper citations rather than their actual population. Finally, the correlation between the number of people recognizing a city and its population—the recognition validity expressed as a correlation—was .60. Recognition is a good predictor of actual population but not as strong as the ecological and surrogate correlations.
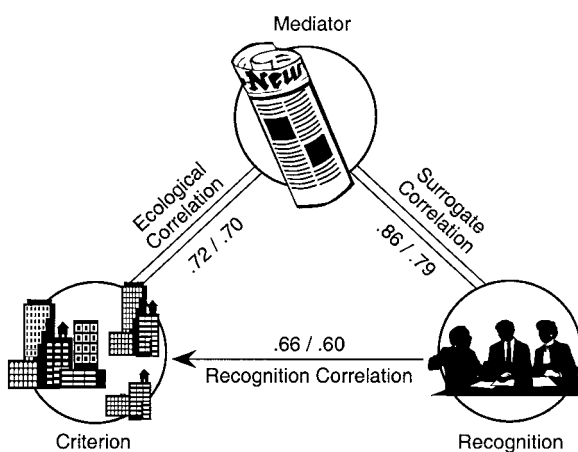

Figure 7. Ecological correlation, surrogate correlation, and recognition correlation. The first value is for American cities and the German newspaper *Die Zeit* as mediator, and the second value is for German cities and the *Chicago Tribune* as mediator. Note that the recognition validity is expressed, for comparability, as a correlation (between the number of people who recognize the name of a city and its population).

Do these results stand up in a different culture? For the class of all American cities with more than 100,000 inhabitants, the ecological correlation was .72. The surrogate correlation was .86, and the correlation between recognition and the rank order of cities was .66. These results are consistent with those from the American data, with slightly higher correlations.

This study illustrates how to analyze the ecological rationality of the recognition heuristic. The magnitude of the recognition validity, together with n and N (Equation 1), specifies the expected accuracy of the heuristic but does not explain why recognition is informative. The ecological and surrogate correlations (see Figure 1) allow one to model the network of mediators that could explain why and when a lack of recognition is informative, that is, when missing knowledge is systematically rather than randomly distributed.

Institutions, Firms, and Name Recognition

For both the *Chicago Tribune* and *Die Zeit*, the surrogate correlation was the strongest association, which suggests that individual recognition was more in tune with the media than with the actual environment. Because of this, to improve the perceived quality of a product, a firm may opt to manipulate the product's name recognition through advertising instead of investing in the research and development necessary to improve the product's quality. Advertising manipulates recognition rates directly and is one way in which institutions exploit recognition-based inference.

One way advertisers achieve name recognition is by associating the name of their products with strong visual images. If the real purpose of the ad is to convey name recognition and not communicate product information, then only the attention-getting quality of the images, and not the content of the images, should matter. Advertiser Oliviero Toscani bet his career on this strategy. In his campaign for Bennetton, he produced a series of advertisements that conveyed nothing about actual Bennetton products but sought to induce name recognition by association with shocking images, such as a corpse lying in a pool of blood or a dying AIDS patient. Would associating the name of a clothing manufacturer with bloody images cause people to learn the name? Toscani (1997) reported that the campaign was a smashing success and that it vaulted Bennetton's name recognition into the top five in the world, even above that of Chanel. In social domains, recognition is often correlated with quality, resources, wealth, and power. Why? Perhaps because individuals with these characteristics tend to be the subject of news reporting as well as gossip. As a result, advertisers pay great sums for a place in the recognition memory of the general public, and the lesser known people, organizations, institutions, and nations of the world go on crusades for name recognition. They all operate on the principle that if we do not recognize them, then we will not favor them.

Those who try to become recognized through advertising may not be wasting their resources. As mentioned, the "overnight fame" experiments by Jacoby, Kelley, et al. (1989) demonstrate that people may have trouble distinguishing between names they learned to recognize in the laboratory and those they learned in the real world. In a series of classic studies, participants read famous and nonfamous names, waited overnight, and then made fame judgments. Occasionally, people would mistakenly infer that a nonfamous name, learned in the laboratory, was the name of a

famous person. The recognition heuristic can be fooled, and the previously anonymous can inherit the appearance of being skilled, famous, important, or powerful.

Can Compensatory Strategies Mimic the Recognition Heuristic?

So far we have dealt only with recognition and not with information recalled from memory. The recognition heuristic can act as a subroutine in heuristics that process knowledge beyond recognition. Examples are the Take The Last and Take The Best heuristics, which feature the recognition heuristic as the first step (Gigerenzer & Goldstein, 1996). These heuristics are also noncompensatory and can benefit from the information implicit in a lack of recognition but are also able to function with complete information. However, in cases in which one object is recognized and the other is not, they make the same inference as the recognition heuristic.

Do compensatory strategies exist that do not have the recognition heuristic as an explicit step but that, unwittingly, make inferences as if they did? It seems that only noncompensatory cognitive strategies (such as lexicographic models) could embody the noncompensatory recognition heuristic. However, it turns out that this is not so. For instance, consider a simple tallying strategy that counts the pieces of positive evidence for each of two objects and chooses the object with more positive evidence (see Gigerenzer & Goldstein, 1996). Such a strategy would be a relative of a strategy that searches for confirming evidence and ignores disconfirming evidence (e.g., Klayman & Ha, 1987). Interestingly, because there can be no positive evidence for an unrecognized object, such a strategy will never choose an unrecognized object over a recognized one. If this strategy, however, were, in addition, to pay attention to negative evidence and choose the object with the larger sum of positive and negative values, it would no longer make the same choice as the recognition heuristic. This can be seen from the fact that a recognized object may carry more negative than positive evidence, whereas an unrecognized object carries neither. As a consequence, compensatory strategies such as tallying that mimic the recognition heuristic will also mimic a less-is-more effect (Gigerenzer & Goldstein, 1999).

Domain Specificity

The recognition heuristic is not a general purpose strategy for reasoning. It is a domain-specific heuristic. Formally, its domain specificity can be defined by two characteristics. First, some objects must be unrecognized ($n < N$). Second, the recognition validity must be higher than chance ($\alpha > .5$). However, what domains, in terms of content, have these characteristics? We begin with examples of recognition validities $\alpha > .5$, and then propose several candidate domains, without claims to an exhaustive list.

Table 2 lists 10 topics of general knowledge, ranging from the world's highest mountains to the largest American banks. Would the recognition heuristic help to infer correctly which of, for example, two mountains is higher? We surveyed 20 residents of Berlin, Germany, about which objects they recognized for each of the 10 topics and then computed each participant's recognition validity for each topic. Table 2 shows the average recognition

Table 2

A Sample of Recognition Validities

Topic	Recognition validity α
10 Largest Indian cities	0.95
20 Largest French cities	0.87
10 Highest mountains	0.85
15 Largest Italian provinces	0.81
10 Largest deserts	0.80
10 Tallest buildings	0.79
10 Largest islands	0.79
10 Longest rivers	0.69
10 Largest U.S. banks	0.68
10 Largest seas	0.64

validities. These validities are often high, and the topics in Table 2 can be extended to include many others.

Domains

Alliances and competition. Deciding whom to befriend and whom to compete against is an adaptive problem for humans and other social animals (e.g., Cosmides & Tooby, 1992; de Waal, 1982). Organisms need a way to assess quickly who or what in the world has influence and resources. Recognition of a person, group, firm, institution, city, or nation signals its social, political, and economic importance to others of its kind that may be contemplating a possible alliance or competition. For instance, in the highly competitive stock market, the recognition heuristic has, on average, matched or outperformed major mutual funds, the market, randomly picked stocks, and the less recognized stocks. This result has been obtained both for the American and the German stock markets (Borges, Goldstein, Ortmann, & Gigerenzer, 1999).

Risk avoidance. A second domain concerns risky behavior when the risks associated with objects are too dangerous to be learned by experience or too rare to be learned in a lifetime (Cosmides & Tooby, 1994). Food avoidance is an example; the recognition heuristic helps organisms avoid toxic foods (Galef, 1987). If an organism had to learn from individual experience which risks to take (or foods to eat), it could be fatal. In many wild environments, novelty carries risk, and recognition can be a simple guide to a safe choice.

Social bonding. Social species (i.e., species that have evolved cognitive adaptations such as imprinting on parents by their children, coalition forming, and reciprocal altruism) are equipped with a powerful perceptual machinery for the recognition of individual conspecifics. Face recognition and voice recognition are examples. This machinery is the basis for the use of the recognition heuristic as a guide to social choices. For instance, one author's daughter always preferred a babysitter whom she recognized to one of whom she had never heard, even if she was not enthusiastic about the familiar babysitter. Recognition alone cannot tell us whom, among recognized individuals, to trust; however, it can suggest that we not trust unrecognized individuals.

Fast and Frugal Heuristics

The current work on the recognition heuristic is part of a research program to study the architecture and performance of fast

and frugal heuristics (Gigerenzer & Goldstein, 1996; Gigerenzer & Selten, 2001; Gigerenzer, Todd, & the ABC Research Group, 1999; Goldstein & Gigerenzer, 1999; Todd & Gigerenzer, 2000). The recognition heuristic is a prototype of these adaptive tools. It uses recognition, a capacity that evolution has shaped over millions of years that allows organisms to benefit from their own ignorance. The heuristic works with limited knowledge and limited time and even requires a certain amount of missing information. Other research programs have pointed out that one's ignorance can be leveraged to make inferences (Glucksberg & McCloskey, 1981); however, they refer to an ignorance of deeper knowledge, not ignorance in the sense of a mere lack of recognition. The program of studying fast and frugal heuristics is based on the following theoretical principles.

Rules for Search, Stopping, and Decision

The common defining characteristics of these heuristics are fast and frugal rules for (a) search, that is, where to search for cues; (b) stopping, that is, when to stop searching without attempting to compute an optimal stopping point at which the costs of further search exceed the benefits; and (c) decision, that is, how to make an inference or decision after search is stopped. The recognition heuristic follows particularly simple rules. Search extends only to recognition information, not to recall. Search is stopped whenever one object is recognized and the other is not; no further information is looked up about the recognized object. The simple decision rule is to choose the recognized object. Limited search and nonoptimizing stopping rules are the key processes in Simon's (e.g., 1955) and Selten's (e.g., 1998) models of bounded rationality. However, only a few theories of cognitive processing specify models of search and stopping; this even holds for those who attach the label "bounded rationality" to their ideas. Most theories only specify rules for decision, such as how information should be combined in some linear or nonlinear fashion.

Transparency

Fast and frugal heuristics are formulated in a way that is highly transparent: It is easy to discern and understand just how they function in making decisions. Because they involve few, if any, free parameters and a minimum of computation, each step of the algorithm is open to scrutiny. These simple heuristics stand in sharp contrast to more complex and computationally involved models of mental processes that may generate good approximations to human behavior but are also rather opaque. For instance, the resurgence of connectionism in the 1980s brought forth a crop of neural networks that were respectable models for a variety of psychological phenomena but whose inner workings remained mysterious even to their creators (for alternatives see Regier, 1996; Rumelhart & Todd, 1993). Whereas these models are precisely formulated but nontransparent, there also exist imprecisely formulated and nontransparent models of heuristics, such as "representativeness," "availability," and other one-word explanations (see Gigerenzer, 1996, 1998; Lopes, 1991). Transparent fast and frugal heuristics use no (or only a few) free parameters, allow one to make quantifiable and testable predictions, and avoid possible misunderstanding (or mystification) of the processes involved, even if they do sacrifice some of the allure of the unknown.

Ecological Rationality

Models of cognitive algorithms—from models of risky choice (e.g., Lopes, 1994) to multiple regression lens models (e.g., Hammond, Hursch, & Todd, 1964) to neural networks (e.g., Regier, 1996)—typically rely on predictor variables that have objective correspondents in the outside world. In this article, we proposed and studied a heuristic that makes inferences without relying on objective predictors but solely on an internal lack of recognition.

We defined the conditions in which the recognition heuristic is applicable and in which the less-is-more effect emerges. In the experimental studies reported, inferences accorded with the recognition heuristic about 90% of the time or more. The recognition heuristic, in the form of Equation 1, predicts how changes in test size affect accuracy, and these predicted changes were experimentally confirmed in 26 of 28 cases. The heuristic also predicts noncompensatory judgments, and this prediction was obtained in a training experiment in which participants were taught predictive information that suggested contradicting recognition. We demonstrated two less-is-more effects. In the first, a group achieved slightly higher accuracy in a foreign domain in which they knew little than in a familiar domain where they knew far more. In the second, we saw how experimentally induced recognition significantly affects the inferences people make.

The recognition heuristic is a cognitive adaptation. In cases of extremely limited knowledge, it is perhaps the only strategy an organism can follow. However, it is also adaptive in the sense that there are situations, called less-is-more effects, in which the recognition heuristic results in more accurate inferences than a considerable amount of knowledge can achieve. It is the model of an ecologically rational heuristic, exploiting patterns of information in the environment to make accurate inferences in a fast and frugal way.

References

Ayton, P., & Önkal, D. (1997). *Forecasting football fixtures: Confidence and judged proportion correct.* Unpublished manuscript, Department of Psychology, The City University of London.

Barnett, S. A. (1963). *The rat: A study in behavior.* Chicago: Aldine.

Birnbaum, M. H. (1983). Base rates in Bayesian inference: Signal detection analysis of the cab problem. *American Journal of Psychology, 96,* 85–94.

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 (Eds.), *Simple heuristics that make us smart* (pp. 59–72). New York: Oxford University Press.

Brown, J., Lewis, V. J., & Monk, A. F. (1977). Memorability, word frequency and negative recognition. *Quarterly Journal of Experimental Psychology, 29,* 461–473.

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.

Cosmides, L., & Tooby, J. (1994). Beyond intuition and instinct blindness: Toward an evolutionarily rigorous cognitive science. *Cognition, 50,* 41–77.

Craik, F. I. M., & McDowd, M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 474–479.

Delbecq-Derousńe, J., Beauvois, J. F., & Shallice, T. (1990). Preserved recall versus impaired recognition. *Brain, 113,* 1045–1074.

de Waal, F. (1982). *Chimpanzee politics.* London: Jonathan Cape.

Duncker, K. (1945). *On problem solving* (L. S. Lees, Trans.). *Psychological Monographs, 58* (5, Whole No. 270). (Original work published 1935)

Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment* (pp. 17–52). New York: Wiley.

Einhorn, H. J. (1970). The use of nonlinear, noncompensatory models in decision making. *Psychological Bulletin, 73,* 221–230.

Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology, 32,* 53–88.

Fishburn, P. C. (1974). Lexicographic orders, utilities and decision rules: A survey. *Management Science, 20,* 1442–1471.

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

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 Behavior, 18,* 199–205.

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review, 98,* 254–267.

Gigerenzer, G. (1993). The bounded rationality of probabilistic mental models. In K. I. Manktelow & D. E. Over (Eds.), *Rationality: Psychological and philosophical perspectives* (pp. 284–313), London: Routledge.

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

Gigerenzer, G. (1998). Surrogates for theories. *Theory & Psychology, 8,* 195–204.

Gigerenzer, G. (2000). *Adaptive thinking: Rationality in the real world.* New York: Oxford University Press

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 104,* 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 (Eds.), *Simple heuristics that make us smart* (pp. 75–95). New York: Oxford University Press.

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

Gigerenzer, G., & Murray, D. J. (1987). *Cognition as intuitive statistics.* Hillsdale, NJ: Erlbaum.

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.

Glucksberg, S., & McCloskey, M. (1981). Decisions about ignorance: Knowing that you don't know. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 311–325.

Goldstein, D. G. (1997). Models of bounded rationality for inference. *Dissertation Abstracts International, 58* (01), 435B. (UMI No. AAT 9720040)

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

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology, 73,* 407–420.

Groner, M., Groner, R., & Bischof, W. F. (1983). Approaches to heuristics: A historical review. In R. Groner, M. Groner, & W. F. Bischof (Eds.), *Methods of heuristics* (pp. 1–18). Hillsdale, NJ: Erlbaum.

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

Hammond, K. R., Hursch, C. J., & Todd, F. J. (1964). Analyzing the components of clinical inference. *Psychological Review, 71,* 438–456.

Hoffrage, U. (1995). *The adequacy of subjective confidence judgments: Studies concerning the theory of probabilistic mental models.* Unpublished doctoral dissertation, University of Salzburg, Austria.

Hogarth, R. M. (1987). *Judgement and choice: The psychology of decision.* Chichester, United Kingdom: Wiley.

Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56,* 326–338.

Jacoby, L. J., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology, 118,* 115–125.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103,* 582–591.

Keeney, R. L., & Raiffa, H. (1993). *Decisions with multiple objectives.* Cambridge, United Kingdom: Cambridge University Press.

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

Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94,* 211–228.

Lopes, L. L. (1991). The rhetoric of irrationality. *Theory and Psychology, 1,* 65–82.

Lopes, L. L. (1994). Psychology and economics: Perspectives on risk, cooperation, and the marketplace. *Annual Review of Psychology, 45,* 197–227.

McKenna, P., & Warrington, E. K. (1980). Testing for nominal dysphasia. *Journal of Neurology, Neurosurgery and Psychiatry, 43,* 781–788.

Mellers, B. A., Schwartz, A., & Cooke, A. D. J. (1998). Judgment and decision making. *Annual Review of Psychology, 49,* 447–477.

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

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

Polya, G. (1954). *Mathematics and plausible reasoning. Vol. 1: Induction and analogy in mathematics.* Princeton, NJ: Princeton University Press.

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

Rumelhart, D. E., & Todd, P. M. (1993). Learning and connectionist representations. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV* (pp. 3–30). Cambridge, MA: MIT Press/Bradford Books.

Schacter, D. L., & Tulving, E. (1994). What are the memory systems of 1994? In D. L. Schacter & E. Tulving (Eds.), *Memory systems* (pp. 1–38). Cambridge, MA: MIT Press.

Schonfield, D., & Robertson, B. (1966). *Memory storage and aging. Canadian Journal of Psychology, 20,* 228–236.

Selten, R. (1998). Aspiration adaptation theory. *Journal of Mathematical Psychology, 42,* 191–214.

Semenza, C., & Sgaramella, T. M. (1993). Production of proper names: A clinical case study of the effects of phonemic cueing. In G. Cohen & D. M. Burke (Eds.), *Memory for proper names* (pp. 265–280). Hove, United Kingdom: Erlbaum.

Semenza, C., & Zettin, M. (1989). Evidence from aphasia for the role of proper names as pure referring expressions. *Nature, 342,* 678–679.

Shepard, R. N. (1967a). On subjectively optimum selections among multiattribute alternatives. In W. Edwards & A. Tversky (Eds.), *Decision making* (pp. 257–283) Harmondsworth: Penguin Books.

Shepard, R. N. (1967b). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior, 6,* 156–163.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69,* 99–118.

Squire, L. R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. *Annual Review of Psychology, 44,* 453–495.

Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology, 25,* 207–222.

Strack, F., & Bless, H. (1994). Memory for nonoccurrences: Metacognitive and presuppositional strategies. *Journal of Memory and Language, 33,* 203–217.

Todd, P. N., & Gigerenzer, G. (2000). Précis of simple heuristics that make us smart. *Behavioral and Brain Sciences, 23,* 727–780.

Toscani, O. (1997). *Die Werbung ist ein lächelndes Aas* [Advertisement is a beast that smiles at us]. Frankfurt, Germany: Fischer.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Warrington, E. K., & McCarthy, R. A. (1988). The fractionation of retrograde amnesia. *Brain and Cognition, 7,* 184–200.