Review Article

Homo Heuristicus: Less-is-More Effects in Adaptive Cognition

Henry Brighton, Gerd Gigerenzer

Submitted: 13 May 2012
Accepted: 29 Jul 2012

Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany

Abstract

Heuristics are efficient cognitive processes that ignore information. In contrast to the widely held view that less processing reduces accuracy, the study of heuristics shows that less information, computation, and time can in fact improve accuracy. We discuss some of the major progress made so far, focusing on the discovery of less-is-more effects and the study of the ecological rationality of heuristics which examines in which environments a given strategy succeeds or fails, and why. Homo heuristicus has a biased mind and ignores part of the available information, yet a biased mind can handle uncertainty more efficiently and robustly than an unbiased mind relying on more resource-intensive and general-purpose processing strategies.

Keywords: *cognition, heuristics, uncertainty*

Introduction

As far as we know, animals have always relied on heuristics to solve adaptive problems, and so have humans. To measure the area of a candidate nest cavity, a narrow crack in a rock, an ant has no yardstick but a rule of thumb: Run around on an irregular path for a fixed period while laying down a pheromone trail, and then leave. Return, move around on a different irregular path, and estimate the size of the cavity by the frequency of encountering the old trail. This heuristic is remarkably precise: Nests half the area of others yielded re-encounter frequencies 1.96 times greater (1). Many evolved rules of thumb are amazingly simple and efficient (2).

The Old Testament says that God created humans in his image and let them dominate all animals, from whom they fundamentally differ (Genesis 1:26). It might not be entirely accidental that in cognitive science some form of omniscience (knowledge of all relevant probabilities and utilities, for instance) and omnipotence (the ability to compute complex functions in a split second) has shaped models of human cognition. Yet humans and animals have common ancestors, related sensory and motor processes, and even share common cognitive heuristics. Consider how a baseball outfielder catches a ball. The view of cognition favoring omniscience and omnipotence suggests that complex problems are solved with complex mental algorithms. Richard Dawkins, for example, argues that "He behaves as if he had solved a set of differential equations in predicting the trajectory of the ball. At some subconscious level, something functionally equivalent to the mathematical calculations is going on" (3, p 96). Dawkins carefully inserts "as if" to indicate that he is not quite sure whether brains actually perform these computations.

And there is indeed no evidence that brains do. Instead, experiments have shown that players rely on several heuristics. The gaze heuristic is the simplest one and works if the ball is already high up in the air: Fix your gaze on the ball, start running, and adjust your running speed so that the angle of gaze remains constant (4). A player who relies on the gaze heuristic can ignore all causal variables necessary to compute the trajectory of the ball–the initial distance, velocity, angle, air resistance, speed and direction of wind, and spin, among others. By paying attention to only one variable, the player will end up where the ball comes down without computing the exact spot. The same heuristic is also used by animal species for catching prey and for intercepting potential mates. In pursuit and predation, bats, birds, and dragonflies maintain a constant optical angle between themselves and their prey, as do dogs when catching a Frisbee (5).

In the 1950s, Herbert Simon proposed that people satisfice rather than maximize (6,7). Maximization means optimization, the process of finding the best solution for a problem, whereas satisficing (a Northumbrian word for "satisfying") means finding a good-enough solution. For

Simon, humans rely on heuristics not simply because their cognitive limitations prevent them from optimizing, but also because of the task environment. For instance, chess has an optimal solution, but no computer or mind, be it Deep Blue or Kasparov, can find this optimal sequence of moves because the sequence is computationally intractable to discover and verify. In the 1970s, the term heuristic acquired a different connotation, undergoing a shift from being regarded as a method that makes computers smart to one that explains why people are not smart.

Daniel Kahneman, Amos Tversky, and their collaborators published a series of experiments in which people's reasoning was interpreted as exhibiting fallacies. "Heuristics and biases" became one phrase. It was repeatedly emphasized that heuristics are sometimes good and sometimes bad, but virtually every experiment was designed to show that people violate a law of logic, probability, or some other standard of rationality. By the end of the 20th century, the use of heuristics became associated with shoddy mental software, generating three widespread misconceptions:

1. Heuristics are always second-best.
2. We use heuristics only because of our cognitive limitations.
3. More time, more information, and more computation would always be better.

These three beliefs are based on the so-called accuracy-effort trade-off, which is considered a general law of cognition: If you invest less effort, the cost is lower accuracy. Effort refers to searching for more information, performing more computation, or taking more time. In fact, these typically go together. Heuristics, on the other hand, allow for fast and frugal decisions; thus it is commonly assumed that they are second-best approximations of more complex "optimal" computations and serve the purpose of trading off accuracy for effort. Contrary to the belief in a general accuracy-effort trade-off, less information and computation can actually lead to higher accuracy, and in these situations the mind does not need to make trade-offs. Here, a less-is-more effect holds. That simple heuristics can be more accurate than complex procedures is one of the major discoveries of the last decades. Heuristics achieve this accuracy by successfully exploiting evolved mental abilities and environmental structures. Since this initial finding a systematic science of heuristics has emerged.

The discovery of less-is-more

Many theories of cognition–from exemplar models to prospect theory to Bayesian models of cognition–assume that all pieces of information should be combined in the final judgment. The classical critique of these models is that in the real world, search for information costs time or money, so there is a point where the costs of further search are no longer justified. This has led to optimization-under-constraints theories in which search in the world (9) or in memory (10) is terminated when the expected costs exceed the benefits. Note that in this "rational analysis of cognition," more information is still considered better, apart from its costs. Similarly, the seminal analysis of the adaptive decision maker (11) rests on the assumption that the rationale for heuristics is a trade-off between accuracy and effort, where effort is a function of the amount of information and computation consumed:

> *Accuracy-effort trade-off:* Information and computation costs time and effort; therefore minds rely on simple heuristics that are less accurate than strategies that use more information and computation.

Here is the first important discovery: Heuristics can lead to more accurate inferences than strategies that use more information and computation (see below). Thus, the accuracy-effort trade-off does not generally hold; there are situations where one attains higher accuracy with less effort. Even when information and computation is entirely free, there is typically a point where less is more:

> *Less-is-more effects:* More information or computation can decrease accuracy; therefore, minds rely on simple heuristics in order to be more accurate than strategies that use more information and time.

To justify the use of heuristics by accuracy-effort trade-offs means that it is not worth the effort to rely on more complex estimations and computations. A less-is-more effect, however, means that minds would not gain anything from relying on complex strategies, even if direct costs and opportunity costs were zero. Accuracy-effort trade-offs are the conventional justification for why the cognitive system relies on heuristics (12,13), which refrains from any normative implications. Less-is-more effects are a second justification with normative consequences: They challenge the classical definition of rational decision-making as the process of weighting and adding all information. Note that the term less-is-more does not mean that the less information one uses, the better the performance. Rather, it refers to the existence of a point at which more information or computation becomes detrimental, independent of costs.

Ignoring information can lead to more accurate predictions

In the 1970s, researchers discovered that equal (or random) weights can predict almost as accurately as, and sometimes better than, multiple linear regression (14–17). Weighting equally is also termed *tallying*, reminiscent of the tally sticks for counting, which can be traced back some 30 000 years in human history. These results came as a surprise to the scientific community. When Robin Dawes presented the results at professional conferences, distinguished attendees told him that they were "impossible," his paper with Corrigan was first rejected and deemed "pre-mature". A sample of recent textbooks in econometrics revealed that none referred to the findings of Dawes and Corrigan (18). In these original demonstrations, there was a slight imbalance: Multiple regression was tested by cross-validation (that is, the model was fitted to one half of the data and tested on the other half) but tallying was not. Czerlinski, Gigerenzer, and Goldstein conducted 20 studies in which both tallying and multiple regression were tested by cross-validation, correcting for this imbalance (19). All tasks were paired comparisons. For instance, estimating which of two Chicago high schools will have a higher drop-out rate, based on cues such as writing score and proportion of Hispanic students. Ten of the 20 data sets were taken from a textbook on applied multiple regression (20). Averaged across all data sets, tallying achieved a higher predictive accuracy than multiple regression (Figure 1). Regression tended to overfit the data, as can be seen by the cross-over of lines: It had a higher fit than tallying, but a lower predictive accuracy.

The point here is not that tallying leads to more accurate predictions than multiple regression. The real and new question is in which environments simple tallying is more accurate than multiple regression, and in which environments it is not. This is the question of the *ecological rationality of tallying*. Tallying avoids precise computation of cue weights. Next, we consider less-is-more effects which arise by ignoring cues. The take-the-best heuristic is a model of how people infer which of two objects has a higher value on a criterion, based on binary cue values retrieved from memory. For convenience, the cue value that signals a higher criterion value is 1, and the other cue value is 0. Take-the-best consists of 3 building blocks:

1. Search rule: Search through cues in order of their validity.
2. Stopping rule: Stop on finding the first cue that discriminates between the objects (i.e., cue values are 1 and 0).
3. Decision rule: Infer that the object with the positive cue value (1) has the higher criterion value.

Take-the-best is a member of the one good-reason family of heuristics because of its stopping rule: Search is stopped after finding the first cue that enables an inference to be made. Take-the-best simplifies decision making by both stopping after the first cue and by ordering cues unconditionally by validity, which for it cue is given by:

v_i = number of correct inferences using cue i / number of possible inferences using cue i.

Both these simplifications have been observed in the behavior of humans and other animals, but routinely interpreted as signs of irrationality rather than adaptive behavior. In the late 1990s, our research group tested how accurately this simple heuristic predicts which of two cities has the larger population, using real-world cities and binary cues, such as whether the city has a soccer team in the major league (21,22). The unexpected result was that inferences relying on one good reason were more accurate than both multiple regression and tallying. We obtained the same result, on average, for 20 studies (Figure 1). This result came as a surprise to both us and the rest of the scientific community. But there were more surprises to come. Chater et al., (23) used the city population problem and tested take-the-best against heavy-weight non linear strategies: A three-layer feedforward connectionist network, trained using the backpropagation algorithm (24), 2 exemplar-based models the nearest-neighbor classifier (25), and Nosofsky's generalised context model (26), and the decision tree induction algorithm C4.5 (27). The predictive accuracy of the four complex strategies was rather similar, but the performance of take-the-best differed considerably. When the percentage of training examples (the sample size) was small or moderate (up to 40% of all objects), take-the-best out performed or matched all the competitors, but when the sample size was larger, more information and computation seemed to be better. This was the first time that relying on one good reason was shown to be as accurate as nonlinear methods, such as a neural network. Yet, as Brighton (28) showed in a re-analysis, Chater et al.'s method of fitting the models on the learning sample and then testing these models on the entire sample (including the learning sample), favored those models that overfit the data, especially at high sample sizes. When cross-validation was used, there was a new surprise: The predictive accuracy of take-the-best exceeded that of all rival models over the entire range of sample sizes (Figure 2). Cross-validation provides a far more reliable model selection criterion and is standard practice for assessing the relative performance of models of inductive inference (29,30).

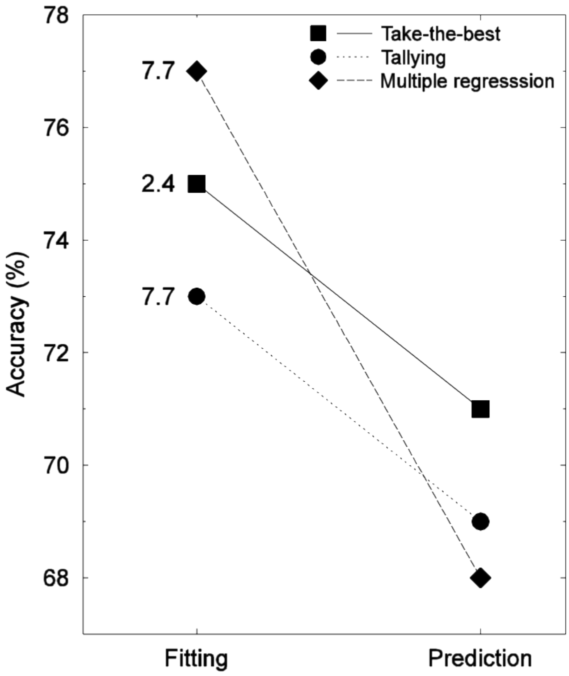

Figure 1: Less-is-more effects. Both tallying and take-the-best predict more accurately than multiple regression, despite using less information and computation. Note that multiple regression excels in data fitting ("hindsight"), that is fitting its parameters to data that is already known, but performs relatively poorly in prediction ("foresight," as in cross-validation). Take-the-best is the most frugal, that is, it looks up, on average, only 2.4 cues when making inferences. In contrast, both multiple regression and tallying look up 7.7 cues on average. The results shown are averaged across 20 studies, including psychological, biological, sociological, and economic inference tasks (19).

Once again, another less-if-more effect was discovered, and a new question emerged: In which environments does relying on one good reason result in better performance than when relying on a neural network or on other linear and non-linear inference strategies? The success of take-the-best seems to be due to the fact that it ignores dependencies between cues in what turns out to be an adaptive processing policy when observations are sparse. Whereas all the competitors in Figure 2 attempt to estimate the dependencies between cues in order to make better inferences, take-the-best ignores them by ordering the cues by validity. In fact, when one alters the search rule of take-the-best by carrying out the more resource-intensive process of ordering cues by conditional validity, performance drops to the level of the more resource-intensive algorithms (Figure 2a). Conditional validity takes into account the fact that when one cue appears before another in the cue order, this first cue is likely to affect the validity of the second cue and all subsequent ones.

These two results are instances of a broader class of less-is-more effects found in the last decades, both analytically and experimentally. We use *less-is-more* here as a generic term for the class of phenomena in which the accuracy-effort trade-off does not hold, although the individual phenomena differ in their nature and explanation. Findings that show how less can be more have often been regarded as curiosities rather than as opportunities to re-think how the mind works. We turn now to the second step of progress made: the development of an understanding of *why* and *when* heuristics are more accurate than strategies that use more information and computation. The answer is not in the heuristic alone, but in the match between a heuristic and its environment. The rationality of heuristics is therefore ecological, not logical.

Ecological rationality

All inductive processes, including heuristics, make bets. This is why a heuristic is not inherently good or bad, or accurate or inaccurate, as is sometimes believed. Its accuracy is always relative to the structure of the environment. The study of the ecological rationality asks the following question: In which environments will a given heuristic succeed, and in which will it fail? Understanding when a heuristic succeeds is often made easier by first asking why it succeeds. As we have shown, when analysing the success of heuristics, we often find that they avoid overfitting the observations. The statistical concept of overfitting is part of the explanation for why heuristics succeed, but to gain a clearer understanding of how and when heuristics exploit the structure of the environment, this issue can be examined more closely.

Heuristics and bias

The study of heuristics is often associated with the term bias. The heuristics and biases program

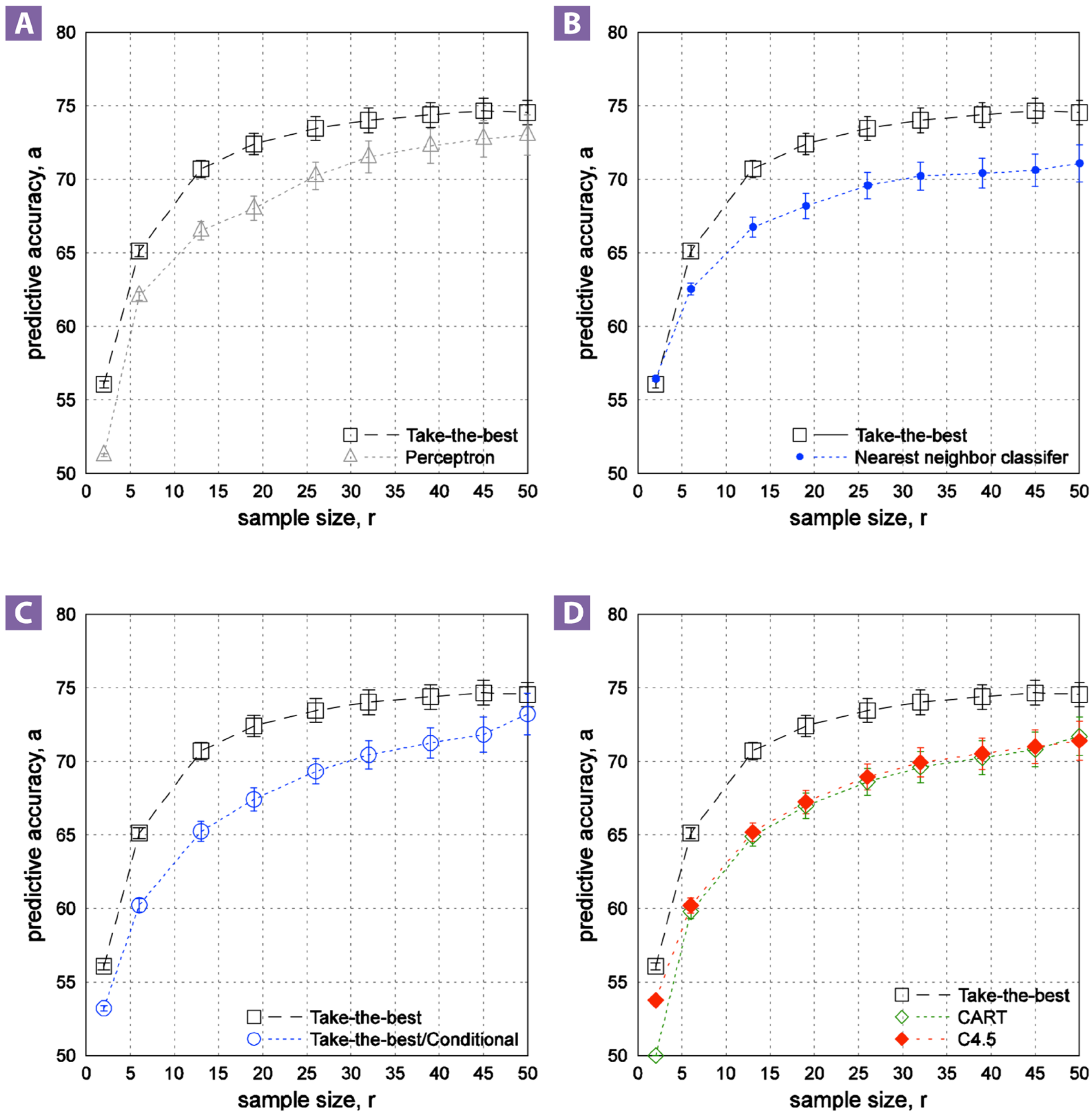

Figure 2: For the city population task, the performance of take-the-best compared to five alternative models. Each panel plots the predictive accuracy of take-the-best and a rival model as a function of the number of objects used to train the model. Take-the-best outperforms; (A) A linear perceptron (essentially logistic regression), (B) the nearest neighbor classifier, (C) a variant of take-the-best that uses a more resource-intensive search rule that orders cues by conditional validity, and (D) two tree induction algorithms, C4.5, and CART (classification and regression trees).

of Kahneman and Tversky used the term with a negative connotation: Reasoning errors reveal human biases that, if overcome, would result in better decisions. In this view, a bias is defined as the difference between human judgment and a "rational" norm, often taken as a law of logic or probability, such as statistical independence as in the gamblers' fallacy. In contrast to this negative use of bias, simple heuristics are perhaps best understood from the perspective of pattern recognition and machine learning, where there are many examples of how a biased induction algorithm can predict more accurately than an unbiased one (29). Findings such as these can be explained by decomposing prediction error into the sum of three components, only one of which is bias:

$$\text{Total error} = (\text{bias})^2 + \text{variance} + \text{noise}.$$

The derivation of this expression can be found in many machine learning and statistical inference textbooks (30–33) (29), but is perhaps most thoroughly set out and discussed in a landmark article by Geman et al., (34). The concepts of bias and variance can be understood by first imagining an underlying (true) function that some induction algorithm is attempting to learn. The algorithm attempts to learn the function from only a (potentially noisy) data sample generated by this function. Averaged across all possible data samples of a given size, the bias of the algorithm is defined as the difference between the underlying function and the mean function induced by the algorithm from these data samples. Thus, zero bias is achieved if this mean function is precisely the underlying function. Variance captures how sensitive the induction algorithm is to the contents of these individual samples and is defined as the sum squared difference between the mean function, mentioned above, and the individual functions induced from each of the samples.

Notice that an unbiased algorithm may suffer from high variance, because the mean function may be precisely the underlying function but the individual functions may suffer from excess variance and hence high error. An algorithm's susceptibility to bias and variance will always depend on the underlying function and how many observations of this function are available. Our cognitive systems are confronted with the bias-variance dilemma whenever they attempt to make inferences about the world. What can this tell us about the cognitive processes used to make these inferences? First of all, cognitive science is increasingly stressing the senses in which the cognitive system performs remarkably well when generalizing from few observations, so much so that human performance is often characterized as optimal (35,36). These findings place considerable constraints on the range of potential processing models capable of explaining human performance. From the perspective of the bias-variance dilemma, the ability of the cognitive system to make accurate predictions despite sparse exposure to the environment strongly indicates that the variance component of error is successfully being kept within acceptable limits. Although variance is likely to be the dominant source of error when observations are sparse, it is nevertheless controllable. This analysis has important implications for the possibility of general-purpose models. To control variance, one must abandon the ideal of general-purpose inductive inference and instead, consider to one degree or another, specialisation (34). Put simply, the bias-variance dilemma shows formally why a mind can be better off with an adaptive toolbox of biased, specialised heuristics. A single, general-purpose tool with many adjustable parameters is likely to be unstable and incur greater prediction error as a result of high variance.

Biased minds for making better predictions

The relationship between mind and environment is often viewed from the perspective of bias, following the "mirror view" of adaptive cognition (37). In this view, a good mental model or processing strategy is assumed to be one that mirrors the properties of the world as closely as possible, preferably with no systematic bias, just as a linear model is assumed to be appropriate if the world is also linear. A cognitive system with a systematic bias, in contrast, is seen as a source of error and the cause of cognitive illusions. If this were true, how can cognitive heuristics that rely only on one good reason and ignore the rest make more accurate inferences than strategies that use more information and computation do, as illustrated in Figure 2? We have identified three reasons:

1. The advantage of simplicity is not because the world is similarly simple, as suggested by the mirror view. This is illustrated by the apparent paradox that although natural environments exhibit dependencies between cues (such as the environment considered in Figure 2, where correlations between cues range between -0.25 and 0.54), take-the-best can make accurate predictions by ignoring

those dependencies, so much so that it can outperform strategies that explicitly set out to model these dependencies. Superior performance is achieved by betting on lower variance, not lower bias.

2. As a consequence, if observations are sparse, simple heuristics like take-the-best are likely to outperform more general, flexible strategies. It is under these conditions that variance will be the most dominant component of error.
3. Similarly, the more noise in the observations, the more likely a simple heuristic like take-the-best will outperform more flexible strategies. The greater the degree of noise, the more dominant the variance component of error is likely to be.

This argument is supported by a diverse set of related findings. First, consider how a retail marketing executive might distinguish between active and nonactive customers. Experienced managers tend to rely on a simple hiatus heuristic: Customers who have not made a purchase for 9 months are considered inactive. Yet there are more sophisticated methods, such as the Pareto/Negative Binomial Distribution (NBD) model, which considers more information and relies on more complex computations. But when tested, these methods turned out to be less accurate in predicting inactive customers than the hiatus rule (38). Second, consider the problem of searching literature databases, where the task is to order a large number of articles so that the most relevant ones appear at the top of the list. In this task, a "one-reason" heuristic (inspired by take-the-best) using limited search outperformed both a "rational" Bayesian model that considered all of the available information and PsychINFO (39). Third, consider the problem of investing money into N funds. Harry Markowitz received the Noble Prize in economics for finding the optimal solution, the mean-variance portfolio. When he made his own retirement investments, however, he did not use his optimizing strategy, but instead relied on a simple heuristic: 1/*N*, that is, allocate your money equally to each of N alternatives (see Table 1 below). Was his intuition correct? Taking 7 investment problems, a study compared the 1/*N* rule with 14 optimizing models, including the mean-variance portfolio and Bayesian and non-Bayesian models (40). The optimizing strategies had 10 years of stock data to estimate their parameters and on that basis had to predict the next month's performance. Next, the 10-year window was moved 1 month ahead, and the next month had to be predicted and so on until the data ran out. 1/*N*, in contrast, does not need any past information. In spite (or because) of this, 1/*N* ranked first (out of 15) on certainty equivalent returns, second on turnover, and fifth on the Sharpe ratio, respectively.

Unpacking the adaptive toolbox

The adaptive toolbox is a metaphor used to conceptualize the stock of strategies available to the organism. Research on the adaptive toolbox attempts to formulate a deeper understanding of the heuristics that humans and other animals use, the building blocks of heuristics that can be used to generate new ones, and the evolved capacities that these building blocks exploit (41). Table 1 shows ten heuristics in the adaptive toolbox of humans. But how does the mind select a heuristic that is reasonable for the task at hand? Although far from a complete understanding of this mostly unconscious process, we know there are at least three selection principles. The first is that memory constrains the choice set of heuristics and thereby creates specific cognitive niches for different heuristics (42). Consider the choice between the first three heuristics in Table 1: (1) the recognition heuristic, (2) the fluency heuristic, and (3) take-the-best. Assume it is 2003, and a visitor has been invited to the third round of the Wimbledon Gentlemen's tennis tournament and encouraged to place a bet on who will win. The two players are Andy Roddick and Tommy Robredo. First, assume that the visitor is fairly ignorant about tennis and has heard of Roddick but not of Robredo. This state of memory restricts the choice set to the recognition heuristic:

> If you have heard of one player but not the other, predict that the recognized player will win the game.

As it happened, Roddick won the match. In fact, this correct inference is not an exception. This simple heuristic predicted the matches of Wimbledon 2003 and 2005 with equal or higher accuracy than the ATP rankings and the seeding of the Wimbledon experts did (56,57). Now assume that the visitor has heard of both players, but recalls nothing else about them. That state of memory limits the choice set to the fluency heuristic:

> If you have heard of both players, but the name of one came faster to your mind than the other, predict that this player will win the game.

Table 1: Ten well-studied heuristics for which there is evidence that they are in the adaptive toolbox of humans. Each heuristic can be used to solve problems in social and nonsocial environments. See the references given for more information regarding their ecological rationality, and the surprising predictions they entail

Heuristic	Definition[1]	Ecologically rational, if	Surprising findings (examples)
Recognition heuristic (43,44)	If one of two alternatives is recognized, infer that it has the higher value on the criterion	Recognition validity > 0.5	Less-is-more effect if $\alpha > \beta$; systematic forgetting can be beneficial (45)
Fluency heuristic (46)	If both alternatives are recognized but one is recognized faster, infer that it has the higher value on the criterion	Fluency validity > 0.5	Less-is-more effect; systematic forgetting can be beneficial (45)
Take-the-best (21)	To infer which of two alternatives has the higher value: (1) search through cues in order of validity (2) stop search as soon as a cue discriminates (3) choose the alternative this cue favors	See Table 1	Often predicts more accurately than multiple regression (19,28)
Tallying (15)	To estimate a criterion, do not estimate weights but simply count the number of positive cues	Cue validities vary little, low redundancy (47,48)	Often predict equally or more accurately than multiple regression (19)
Satisficing (6,49)	Search through alternatives and choose the first one that exceeds your aspiration level	Number of alternatives decreases rapidly over time, such as in seasonal mating pools (50)	Aspiration levels can lead to significantly better choices than chance, even if they are arbitrary (51,52)
1/N; equality heuristic (40)	Allocate resources equally to each of N alternatives	High unpredictability, small learning sample, and large N	Can outperform optimal asset allocation portfolios
Default heuristic (53)	If there is a default, do nothing	Values of those who set defaults match those of the decision maker, when the consequences of a choice are hard to foresee	Explains why mass mailing has little effect on organ donor registration; predicts behavior when trait and preference theories fail
Tit-for-tat (54)	Cooperate first and then imitate your partner's last behavior	The other players also play tit-for-tat; the rules of the game allow for defection or cooperation but not divorce	Can lead to a higher payoff than optimization (backward induction)
Imitate the majority (55)	Consider the majority of people in your peer group and imitate their behavior	Environment is stable or only changes slowly; info search is costly or time-consuming	A driving force in bonding, group identification, and moral behavior
Imitate the successful (55)	Consider the most successful person and imitate his or her behavior	Individual learning is slow; information search is costly or time-consuming	A driving force in cultural evolution

[1] For formal definitions, see references.

Finally, assume that the visitor is more knowledgeable and can recall various facts about both players. That again eliminates the recognition heuristic and leaves a choice between the fluency heuristic and take-the-best. According to the experimental evidence, the majority of subjects switch to knowledge-based heuristics such as take-the-best when the values of both alternatives on relevant cues can be recalled (8), consistent with an analysis of the relative ecological rationality of the two heuristics in this situation. The general point is that memory "selects" heuristics in a way that makes it easier and faster to apply a heuristic when it is likely to yield accurate decisions (42). In the extreme case where the visitor has not heard of any of the players, none of the heuristics can be used. In this event, the visitor can resort to social heuristics, such as imitate the majority: Bet on the player on whom most others bet.

The second known selection principle, after memory, is feedback. Strategy selection theory (58) provides a quantitative model that can be understood as a reinforcement theory where the unit of reinforcement is not a behavior, but a heuristic. This model allows predictions about the probability that a person selects one strategy within a defined set of strategies. The third selection principle relies on the structure of the environment, as analyzed in the study of ecological rationality. For instance, the recognition heuristic is likely to lead to fast and accurate judgments if the recognition validity is high, that is, a strong correlation between recognition and the criterion exists, as is the case for tennis and other sports tournaments. There is experimental evidence that people tend to rely on this heuristic if the recognition validity is high but less so if the recognition validity α is low or at chance level ($\alpha = 0.5$). For instance, name recognition of Swiss cities is a valid predictor for their population ($\alpha = 0.86$), but not for their distance from the center of Switzerland, the city of Interlaken ($\alpha = 0.51$). Pohl (59) reported that 89% of participants relied on the recognition heuristic in judgments of population, but only 54% in judgments of distance to Interlaken. Thus, the use of the recognition heuristic involves two processes: first, recognition in order to see whether the heuristic can be applied, and second, evaluation in order to judge whether it should be applied.

Homo heuristicus

In this article, we summarized a vision of human nature based on an adaptive toolbox of heuristics rather than on traits, attitudes, preferences, and similar internal explanations. We discussed the progress made in developing a science of heuristics, beginning with the discovery of less-is-more effects that contradict the prevailing explanation in terms of accuracy-effort trade-offs. Instead, we argue that the answer to the question "Why heuristics?" lies in their ecological rationality, that is, in the environmental structures to which a given heuristic is adapted. Appealing to the bias-variance dilemma, we proposed how the ecological rationality of heuristics can be formally studied, focusing on uncertain criteria and small samples that constitute environmental structures which fast and frugal heuristics can exploit. Homo heuristicus can rely on heuristics because they are accurate, not because they require less effort at the cost of some accuracy (60). We hope to have raised our readers curiosity about the emerging science of heuristics, and also hope that some might be inspired to solve some of the open questions, such as whether there is a system of building blocks of heuristics, similar to the elements in chemistry, and how a vocabulary for describing relevant environmental structures can be found.

Correspondence

Dr Henry Brighton
PhD Cognitive Science (UK)
Center for Adaptive Behavior and Cognition
Max Planck Institute for Human Development
Lentzeallee 94
14195 Berlin, Germany
Tel: +4930-8240 6339
Fax: +4930-8240 6394
Email: hbrighton@mpib-berlin.mpg.de

References

1. Mugford, S. T., Mallon, E. B., Franks, N. R. The accuracy of Buffon's needle: A rule of thumb used by ants to estimate area. *Behavioral Ecology*. 2001;**12(6)**:655–658.
2. Hutchinson, J. M. C., Gigerenzer, G. Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet. *Behavioural Processes*. 2005;**69**:97–124.
3. Dawkins, R. *The selfish gene*. 2nd ed. Oxford (UK): Oxford University Press; 1989.
4. Gigerenzer, G. *Gut feelings: The intelligence of the unconscious*. New York (NY): Viking; 2007.
5. Shaffer, D. N., Krauchunas, S. M., Eddy, M., McBeath, M. K. How dogs navigate to catch Frisbees. *Psychological Science*. 2004;**15**:437–441.
6. Simon, H. A. A behavioral model of rational choice. *Quarterly Journal of Economics*. 1955;**69**:99–118.

7. Simon, H. A. *Models of my life*. New York (NY): Basic Books; 1991.
8. Marewski, J. N., Gaissmaier, W., Schooler, L. J., Goldstein, D. G., Gigerenzer, G. *Models of recognition-based multi-alternative inference*. Manuscript submitted for publication; 2008.
9. Stigler, G. J. The economics of information. *Journal of Political Economy*. 1961;**69**:213–225.
10. Anderson, J. R. The adaptive nature of human categorization. *Psychological Review*. 1991;**98**: 409–429.
11. Payne, J. W., Bettman, J. R., & Johnson, E. J. *The adaptive decision maker*. New York (NY): Cambridge University Press; 1993.
12. Beach, L. R., Mitchell, T. R. A contingency model for the selection of decision strategies. *Academy of Management Review*. 1978;**3**:439–449.
13. Shah, A. K., Oppenheimer, D.M. Heuristics made easy: An effort-reduction framework. *Psychological Bulletin*. 2008;**137**:207–222.
14. Dawes, R. M., Corrigan, B. Linear models in decision making. *Psychological Bulletin*. 1974;**81**:95–106.
15. Dawes, R. M. The robust beauty of improper linear models in decision making. *American Psychologist*. 1979;**34**:571–582.
16. Einhorn, H. J., Hogarth, R. M. Unit weighting schemes for decision making. *Organizational Behavior and Human Performance*. 1975;**13**:171–192.
17. Schmidt, F. L. The relative efficiency of regression and simple unit weighting predictor weights in applied differential psychology. *Educational and Psychological Measurement*. 1971;**31**:699–714.
18. Hogarth, R. M. *When simple is hard to accept*. In: P. M. Todd, G. Gigerenzer, ABC Research Group, editors. *Ecological rationality: Intelligence in the world*. United Kingdom (UK): Oxford University Press; 2012.
19. Czerlinski, J., Gigerenzer, G., Goldstein, D. G. How good are simple heuristics? In: G. Gigerenzer, P.M. Todd, ABC Research Group, editors. *Simple heuristics that make us smart*. New York (NY): Oxford University Press; 1999. p. 97–118.
20. Weisberg, S. *Applied linear regression*. New York (NY): Wiley; 1985.
21. Gigerenzer G., Goldstein, D. G. Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*. 1996;**103**:650–669.
22. Gigerenzer, G., Goldstein, D. G. Betting on one good reason: The Take The Best heuristic. In: G. Gigerenzer, P. M. Todd, ABC Research Group, Editors. *Simple heuristics that make us smart*. New York (NY): Oxford University Press; 1999. p. 75–95.
23. Chater, N., Oaksford, M., Nakisa, R., Redington, M. Fast, frugal and rational: How rational norms explain behavior. *Organizational Behavior and Human Decision Processes*. 2003;**90**:63–86.
24. Rumelhart, D. E., Hinton, G. E., Williams, R. J. Learning internal representations by error propagation. In: D. E. Rumelhart, J. L. McClelland, Editors. *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge (MA): MIT Press. 1986;**1**:318–362.
25. Cover, T., Hart, P. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*. 1967;**13**:21–27.
26. Nosofsky, R. M. Relations between exemplar similarity and likelihood models of classification. *Journal of Mathematical Psychology*. 1990;**34**:393–418.
27. Quinlan, J. R. C4.5: *Programs for machine learning*. *San Mateo*, Canada (CA): Morgan Kaufmann; 1993.
28. Brighton, H. Robust inference with simple cognitive models. In: C. Lebiere, R. Wray, editors. *AAAI Spring Symposium: Cognitive Science Principles Meet AI-Hard Problems*. Canada (CA): American Association for Artificial Intelligence; 2006. p. 17–22.
29. Hastie, T., Tibshirani, R., Friedman, J. *The elements of statistical learning: Data mining, inference, and prediction*. New York (NY): Springer; 2001.
30. Stone, M. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society B*. 1974;**36**:111–147.
31. Alpaydin, E. *Introduction to machine learning*. Cambridge (MA): MIT Press; 2004.
32. Bishop, C. M. *Neural networks for pattern recognition*. Oxford (UK): Oxford University Press; 1995.
33. Bishop, C. M. *Pattern recognition and machine learning*. New York (NY): Springer; 2006.
34. Geman, S., Bienenstock, E., Doursat, R. Neural networks and the bias/variance dilemma. *Neural Computation*. 1992;**4**:1–58.
35. Oaksford, M., Chater, N. Rational models of cognition. United Kingdom (UK): Oxford University Press; 1998.
36. Griffiths, T. L., Tenenbaum, J. B. Optimal predictions in everyday cognition. *Psychological Science*. 2006;**17(9)**:767–773.
37. Brighton, H. Gigerenzer, G. How heuristics exploit uncertainty. In: P. M. Todd, G. Gigerenzer, ABC Research Group, editors. *Ecological rationality: Intelligence in the world*. New York (NY): Oxford University Press; 2012.
38. Wübben, M., Wangenheim, F v. Instant customer base analysis: Managerial heuristics often “get it right”. *Journal of Marketing*. 2008;**72**:82–93.
39. Lee, M. D., Loughlin, N., Lundberg, I. B. Applying one reason decision-making: The prioritization of literature searches. *Australian Journal of Psychology*. 2002;**54**:137–143.
40. DeMiguel, V., Garlappi, L., Uppal, R. *Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?* Review of Financial Studies: Oxford University Press; 2007.

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

42. Marewski, J. N., Schooler, L. J. How memory aids strategy selection. The Brunswik Society Newsletter. 2008:**23**;27-28.

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

44. Goldstein, D. G., Gigerenzer, G. Models of ecological rationality: The recognition heuristic. *Psychological Review*. 2002;**109**:75–90.

45. Schooler, L. J., Hertwig, R. How forgetting aids heuristic inference. *Psychological Review*. 2005;**112**:610–628.

46. Jacoby, L. L., Dallas, M. On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*. 1981;**110**:306-340.

47. Hogarth, R. M., Karelaia, N. Ignoring information in binary choice with continuous variables: When is less "more"? *Journal of Mathematical Psychology*. 2005;**49**:115–124.

48. Hogarth, R. M., Karelaia, N. "Take-the-best" and other simple strategies: Why and when the work "well" with binary cues. *Theory and Decision*. 2006;**61**:205–249.

49. Todd, P. M., Miller, G. F. From pride and prejudice to persuasion: Realistic heuristics for mate search. In: G. Gigerenzer, P. M. Todd, ABC Research Group, editors. *Simple heuristics that make us smart*. New York (NY): Oxford University Press; 1999. p. 287–308.

50. Dudey, T., Todd, P. M. Making good decisions with minimal information: Simultaneous and sequential choice. *Journal of Bioeconomics*. 2002;**3**:195–215.

51. Gilbert, J. P., Mosteller, F. Recognizing the maximum of a sequence. *American Statistical Association Journal*. 1966;**61**:35–73.

52. Bruss, F. T. Der Ungewissheit ein Schnippchen schlagen. *Spektrum der Wissenschaft*. 2000;**6**:106.

53. Johnson, E. J., Goldstein, D. G. *Do defaults save lives? Science*. 2003;**302**:1338–1339.

54. Axelrod R. *The evolution of cooperation*. New York (NY): Basic Books; 1984.

55. Boyd, R., Richerson, P. J. *The origin and evolution of cultures*. New York (NY): Oxford University Press; 2005.

56. Scheibehenne, B., Brˆder, A. Predicting Wimbledon 2005 tennis results by mere player name recognition. *International Journal of Forecasting*. 2007;**3**: 415–426.

57. Serwe, S., Frings, C. Who will win Wimbledon? The recognition heuristic in predicting sports events. *Journal of Behavioral Decision Making*. 2006;**19**:321–332.

58. Rieskamp, J., Otto, P. E. SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology General*. 2006;**135**:207–236.

59. Pohl, R. F. Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making*. 2006;**19**:251–271.

60. Gigerenzer G., Brighton, H. Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*. 2009;**1**:107–143.