

Heuristics, History of

Ralph Hertwig and Thorsten Pachur, Max Planck Institute for Human Development, Center for Adaptive Rationality, Berlin, Germany

© 2015 Elsevier Ltd. All rights reserved.

Abstract

The study of heuristics is an interdisciplinary and time-honored enterprise, with heuristics being examined across a wide range of fields, some focusing on professionals' decision-making. In psychology, two influential research traditions have investigated which cognitive heuristics people use across various tasks. One tradition, which originated in the 1970s, has focused on the circumstances under which heuristics cause deviations from classical norms of rationality. The other, starting in the 1990s, has emphasized that heuristics are neither good nor bad per se, but that their success depends on how they are matched to environmental structures. This view suggests that heuristics, rather than leading to irrationality, enable ecological rationality.

What Is a Heuristic?

A heuristic is a simple but useful method for problem solving, decision-making, and discovery. The origin of the term goes back to the Ancient Greek verb *heuriskein*, which means 'to find out' or 'to discover.' Heuristics are sometimes also referred to as 'mental shortcuts' or 'rules of thumb.' One of their key functions is to reduce the complexity of a problem by ignoring part of the available information or searching only a subset of all possible solutions. Traditionally, heuristics have been regarded as necessary and efficient tools, but ones that produce only second-best solutions. However, more recent psychological research has shown that – under conditions that are ubiquitous in the real world, namely, limited knowledge and uncertainty – heuristics can in fact outperform more complex strategies.

Heuristic methods have been considered in various scientific disciplines (though what exactly is meant by a heuristic varies across fields). Two key treatments of heuristic methods can be distinguished. In philosophy, mathematics, operations research, and artificial intelligence (AI), heuristics have primarily been investigated as prescriptive procedures specifying how a reasonable solution can be found given constraints such as computational intractability and limited time. In biology and psychology, by contrast, heuristic principles have also been used as descriptive models, that is, as models that describe how people and other animals sample information from the external and internal (memory) world, and how they make decisions based on that information. The next section gives a historical overview of the discussion of heuristics within these two contexts.

A Short History of Heuristics

Heuristics as Prescriptive Procedures

Heuristic methods were first developed in philosophy and mathematics as a solution to the problems of algorithmic approaches to complex problems. To illustrate the algorithmic tradition, let us take the mechanical device developed by the Catalan philosopher Raimundus Lullus (1232–1315) in

the thirteenth century. The device was able to automatically generate all combinations of religious and philosophical attributes that could be used in a debate (Figure 1). It consisted of six concentric discs representing the basic classes of arguments; each of these classes in turn had nine further attributes. Rotating the discs against each other produced every possible combination of attributes, thus automatically generating different arguments and potentially producing new ones. The German philosopher Gottfried Wilhelm Leibniz (1646–1716) had a similar, though even grander goal in the seventeenth century, namely to develop an algorithm for solving any conceivable problem, using a universal language that would allow every possible problem to be represented. It became increasingly clear, however, that such algorithmic approaches could easily lead to combinatorial explosion. Subsequent scholars therefore explored heuristics as means to bring

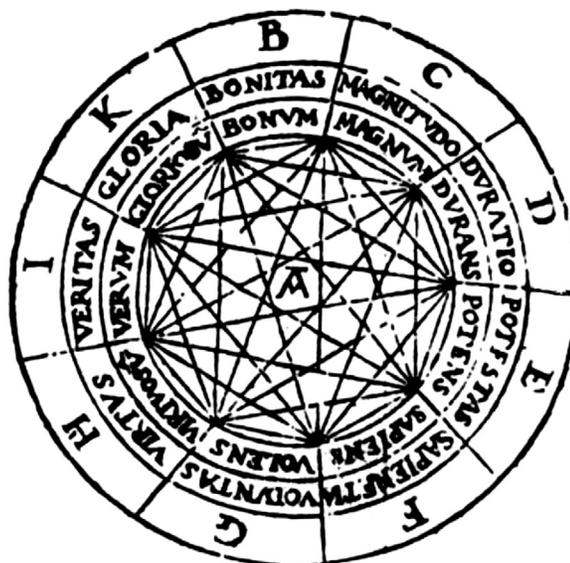


Figure 1 A sketch of one of the discs of Lullus' system for deriving all combinations of arguments. Source: <http://www.medienkunstnetz.de/works/ars-magna/>.

problem solvers to a solution without the need to explore the space of possibilities exhaustively.

René Descartes (1596–1650), the French polymath of the seventeenth century, had formulated simple rules that were supposed to guide the problem solver toward the relevant aspects, rather than to all parts of a problem. In the nineteenth century, the mathematician and philosopher Bernard Bolzano (1781–1848), who proposed various heuristics for problem solving in his renowned *Theory of Science* (*Wissenschaftslehre*), further developed this approach for epistemic agents (e.g., the method of attempting to find truths by means of something that is not yet known to be true, as opposed to deducing truth from known truths). These scholars' heuristics consisted of rather general procedures that were described in relatively vague terms. Their main goals were to offer a road to knowledge beyond proper deduction, to avoid an arbitrary solution process, and to foster creative thinking. For instance, many heuristics aimed at finding useful problem representations by way of analogy or metaphor.

Heuristic methods for problem solving and discovery received wider recognition in modern mathematics through the work of the mathematician George Pólya (1887–1985). Pólya's (1945) procedures consisted of simple rules, such as dividing the process toward a solution into simple steps by, for instance, finding an analogy to a problem, finding a more specialized problem, or decomposing and recombining the problem. Procedures inspired by Pólya later also informed the new field of AI. Relatedly, heuristics proved to be of great practical value in operations research, an applied field of mathematics. Here, they have been implemented as computer-based tools for planning and aiding decision-making in industry.

Methods for simplifying decision-making have also been discussed in economics. Many decisions are made under uncertainty, with key aspects of the problem remaining unknown. In the absence of complete knowledge of the option space and the probability distributions, normative principles of rational choice are difficult or even impossible to apply. Heuristics for making decisions under uncertainty have therefore been proposed (Savage, 1954). The maximin rule (Coombs et al., 1970), for example, selects from a set of options the one that would yield the most attractive outcome under the worst-case scenario, ignoring both the outcomes to be expected under better conditions and the (unknown) probability distribution. The satisficing principle is a heuristic for sequential decision-making (Simon, 1956). Instead of aspiring to identify the globally best option, this principle evaluates each option according to whether it meets a certain minimum aspiration level, and the first option encountered that satisfies that requirement is chosen.

Another research field of economics in which simple, heuristic decision strategies have received attention is game theory. One of the most frequently studied strategies is tit-for-tat. This strategy, which can be applied in repeated games (e.g., the iterated prisoner's dilemma), determines whether or not a player cooperates with his/her opponent. At the first encounter with the other player, tit-for-tat cooperates; in subsequent encounters, it simply copies the other's behavior (cooperate vs defect) in the previous encounter. In a classical computer simulation tournament (Axelrod, 1984), tit-for-tat proved to be the most successful – though simplest – of the

strategies submitted. As well as informing attempts to improve conflict resolution, the study of simple strategies for cooperation within game theory offers insights into how cooperative societies might have evolved in the first place.

The advent of the computer as a computational tool and metaphor of the mind in the 1950s spawned attempts to simulate intelligent behavior in machines. The goal of the new field of AI has been to develop computer programs that can perform tasks such as playing chess, proving logical theorems, or understanding language – and a common challenge is, again, to find ways of limiting boundless and therefore prohibitively expensive search of the problem space. In contrast to previous treatments of heuristics, which often portrayed them in rather vague terms, heuristic rules in AI research have been formulated precisely, often in terms of computational models. A prominent heuristic method for limiting search is means–ends analysis, which was developed in the context of Alan Newell and Herbert Simon's (1972) General Problem Solver system. To move from a current state in the problem space toward a goal state (representing the solution), means–ends analysis reduces the distance between the two by first addressing the most important dimension on which the two differ, followed by the second most important dimension, and so on.

Heuristics as Descriptive Models of the Mind

On the assumption that heuristics, though not perfect, are often effective tools for dealing with a complex and uncertain world, it seems reasonable to suppose that the mind employs them naturally and spontaneously. Indeed, the psychologist Karl Duncker (1903–40) argued that human problem-solving strategies rest on heuristic principles (Duncker, 1935). Similarly, Gestalt psychologists conceived of perception in terms of heuristics. Max Wertheimer (1880–1943), for instance, identified a set of simple principles that organize sensory input to yield object perception – such as the 'laws' of proximity, closure, and similarity (Wertheimer, 1923/1938). As in other contexts, these principles of perceptual organization represent best guesses that usually can be trusted, although they do not work all of the time. They are useful because they are adapted to certain regularities in the environment; for example, elements belonging to the same object are typically in close vicinity to each other (the law or heuristic of proximity; see Figure 2).

Presently, cognitive heuristics are most commonly studied in research on how people make decisions. Inspired by Herbert Simon's (1916–2001) seminal work on 'bounded rationality' – the study of how people reason and make decisions with limited computational and informational resources, and when

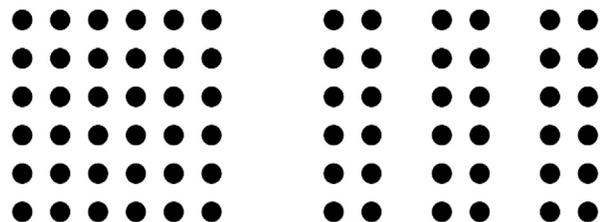


Figure 2 The Gestalt 'heuristic' of proximity: objects positioned close together are perceived as belonging together.

the conditions for rationality postulated by the model of neoclassical economics are not met (Simon, 1956, 1989, 1990) – several research programs have set out to identify and test simple decision strategies. Daniel Kahneman (1934 to present) and Amos Tversky (1937–96) showed that people assessing the frequency or probability of an event often use simple mental shortcuts, such as the availability heuristic or the representativeness heuristic, according to which a judgment is based on ease of retrieval (availability) or on the similarity of an event to a given class of events (representativeness) (Kahneman et al., 1982). What these heuristics have in common is that the cognitive system, when required to make a computationally complex judgment (that necessitates application of the laws of logic or probability theory), substitutes a less complex heuristic attribute, such as similarity (Griffin et al., 2002). Empirical support for the use of these heuristics was provided by creating conditions in which the use of heuristics resulted in systematic errors or lapses of reasoning ('fallacies,' 'cognitive illusions'; Griffin et al., 2002), suggesting human irrationality. From the 1970s, heuristics therefore came to be seen in cognitive and social psychology, and subsequently in behavioral economics, as fallible cognitive shortcuts that people – constrained by cognitive limitations – fall back on in situations in which they would be better off using logic and probability theory. The price of relying on these shortcuts is the risk of committing systematic and severe fallacies of reasoning.

The heuristics examined in Kahneman and Tversky's heuristics-and-biases program were typically vaguely defined verbal descriptions of mental processes. In the research program on the 'adaptive decision-maker,' in contrast, Payne et al. (1993) proposed heuristics for multi-attribute preference tasks that were described in terms of process models. Using computer simulations, Payne et al. tested the accuracy of various preference strategies, both simple and complex, under conditions of time pressure or when the inputs to the strategies, the options' attributes, were highly correlated. The results were then compared with people's actual strategy selection under the same conditions. One key insight was that people are adaptive decision-makers, switching between strategies based on a trade-off between a strategy's prospective accuracy and the time and cognitive effort required to implement it.

From the 1990s, a new research program on cognitive heuristics emerged that challenged both the idea that heuristics are error-prone mental tools (Kahneman et al., 1982) and the notion of a general accuracy–effort trade-off (Payne et al., 1993). Gerd Gigerenzer et al.'s (1999, 2011) research program on 'ecological rationality' proposed that simple heuristics represent prescriptive models of the human mind, thereby connecting the science of heuristics in psychology to earlier research (see Heuristics as Prescriptive Procedures). This approach also suggested that heuristics take advantage of environmental structures, and thus picked up on ideas first introduced by Gestalt psychologists (see above).

The new approach to heuristics thus emphasizes that the real world imposes circumstances on the cognitive system that render optimization difficult or impossible (e.g., computational intractability; incomplete knowledge of alternatives, consequences, and probabilities; and changing preference structures). People therefore rely on heuristics not primarily

because of their cognitive limitations – although it is indisputable that the human cognitive system is subject to constraints – but because of inescapable properties of real-world problems.

How Well Do Heuristics Work?

The new conception of heuristics in terms of ecologically rational strategies (Gigerenzer et al., 2011) has also challenged the deeply entrenched beliefs that heuristics are doomed to produce second-best results and that optimization is always better. It has been shown that, under some circumstances, heuristics perform as well as or even better than complex strategies involving extensive estimations and computations. For instance, consider heuristics for inferring which of two (or more) alternatives (e.g., cities, companies, sports teams) has a higher value on a quantitative criterion (e.g., homelessness rate, revenue, final score) on the basis of known pieces of information (cues). The normative approach involves weighting and combining all the available cues using complex algorithms such as multiple regression, neural networks, Bayesian networks, or classification and regression trees (see Gigerenzer and Brighton, 2009). Although each of these algorithms processes cues in very different ways, they are all based on the shared premise that all available information should be taken into account.

Models of cognitive heuristics dispense with this premise. They can shortcut the demanding chore of processing and integrating all available information in at least three ways. First, rather than assigning an 'optimal' weight to each cue individually, heuristics can simply tally cues for or against each alternative (tallying heuristic), with each cue receiving the same weight (Dawes, 1979). Second, instead of using all available cues, they can consider just a subset of cues and ignore dependencies between cues. Such 'noncompensatory' heuristics use the first cue that permits a decision to be made and ignore all other cues. This class of heuristics assumes some kind of cue ordering and a stopping rule that specifies when search for information is to be terminated. Cue ordering can, for instance, be based on some measure of cue diagnosticity, such as validity (i.e., the probability that the cue correctly predicts the target variable). The take-the-best heuristic assumes that cues are processed in order of validity, and it compares both alternatives on a single cue, one at a time, until a cue is found that distinguishes between the alternatives (Gigerenzer et al., 2011). Third, some heuristics radically winnow down the number of cues inspected to just one. The recognition heuristic, for instance, only ever considers the 'recognition' cue: when one alternative is recognized and the other is not, it selects the recognized one (Goldstein and Gigerenzer, 2002). The fluency heuristic makes an inference based on the 'fluency' cue: when both alternatives are recognized but one is recognized faster, it selects the one that is recognized faster (Hertwig et al., 2008).

How well do heuristics that do not consider all information perform? And when and why can they perform well? Katsikopoulos et al. (2010) examined the accuracy of the take-the-best, tallying, and minimalist heuristics across real-world data sets drawn from domains such as biology, economics, and sociology. For example, one inference problem required

the strategies to predict which of two cities (e.g., Chicago or Los Angeles) has a higher homelessness rate (criterion value) on the basis of cues such as the cities' average temperature. Two more complex strategies were included as benchmarks: first, linear regression, which weights and sums cues and estimates dependencies between cues; second, naive Bayes, which selects the alternative with the higher probability of having the higher criterion value, given the alternatives' entire cue profile. To compare the accuracy of each of these models, the authors split each data set into two parts and estimated the parameters of each model (e.g., cue validity, cue weight) on one part, the 'training set.' These parameter estimates were used, for each model, to make inferences on the other part, the 'test set.' This is called cross validation, a standard procedure for assessing the predictive performance of models of inductive inference. The size of the training set ranged from just 2–10 objects in the population to 50% of the objects.

Figure 3 plots mean predictive accuracy, defined as a strategy's proportion of correct inferences in the test set, as a function of the training set size. Strategies are differentially influenced by the information that small training sets provide. The accuracy of linear regression and naive Bayes is compromised, presumably because very small training sets provide too few data points to reliably estimate the cue weights. The heuristics, in contrast, are capable of managing with very limited information: Tallying outperforms the other strategies when the training set contains only two objects; the take-the-best heuristic is the most accurate model for training sets of between 3 and 10 objects. The results in Figure 3 appear to be the rule rather than the exception. Simple heuristics such as take-the-best have also been shown to perform well in cross validation when compared with such complex and resource-intensive strategies as linear perception and tree-induction algorithms (Gigerenzer and Brighton, 2009).

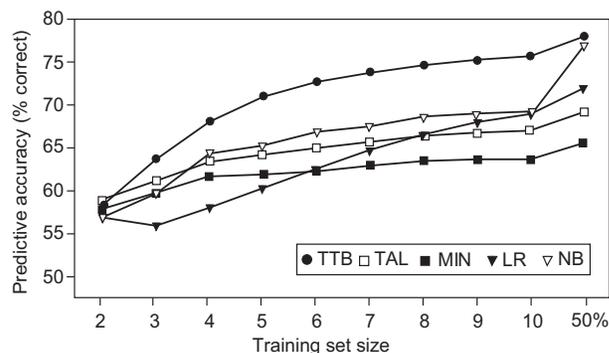


Figure 3 Mean predictive accuracy (averaged across 19 data sets) of three heuristics (take-the-best (TTB), tallying (TAL), and minimalist (MIN)) and two benchmark models (linear regression (LR) and naive Bayes (NB)) as a function of the size of the training set (2–10 objects and 50% of the objects in the population). For training samples as small as 3–10 alternatives, the take-the-best heuristic is more accurate than are the computationally complex models (or the other heuristics; the minimalist heuristic randomly selects one cue until it finds one that discriminates); here, the strategies' best performing variants are plotted. Adapted from Katsikopoulos, K.V., Schooler, L.J., Hertwig, R., 2010. The robust beauty of ordinary information. *Psychological Review* 117, 1259–1266.

These results suggest that more information, more computation, and more time need not always result in better performance, and that the accuracy–effort trade-off does not generally hold. A better understanding has recently emerged as to why and when heuristics are more (or less) accurate than strategies that require more information and computation.

Why Heuristics Work

There is more than one reason for the surprising success of heuristics. Two of the most important are their ability to exploit sophisticated evolved capacities, and the bias–variance dilemma.

Evolved Capacities

Heuristics can exploit evolved human capacities. For instance, the recognition heuristic and the fluency heuristic (see above for definitions) take advantage of the evolved capacities for recognition memory and systematic forgetting (i.e., pushing aside information that is unlikely to be needed again; Schooler and Hertwig, 2005). The simple tit-for-tat heuristic for making decisions in social exchange situations – cooperate first and then imitate your counterpart's behavior (Axelrod, 1984) – takes advantage of a suite of evolved capacities, including memory (players must remember the exchanged value and the accrued debt), numerical discrimination (to evaluate whether an exchange was equitable), and temporal discounting (as reciprocity involves paying an immediate cost for future benefits, those future benefits must be time-discounted appropriately; Stevens and Hauser, 2004).

Ecological Rationality

According to Simon's adaptive view of 'bounded rationality,' human rational behavior "is shaped by a scissors whose two blades are the structure of task environments and the computational capabilities of the actor" (Simon, 1990: 7). This view suggests that any cognitive strategy, heuristic in nature or otherwise, is not inherently good or bad, successful or unsuccessful (Gigerenzer et al., 1999). Rather Simon's scissors metaphor raises the following question: In which environment will a given strategy succeed, and when and where will it fail? Investigations that aim to uncover the environmental structures that foster or are detrimental to a specific heuristic's success have been referred to as studying 'ecological rationality' (Todd et al., 2012). Ecological rationality brings a prescriptive dimension to the psychological investigation of heuristics. One key insight stems from the bias–variance dilemma (Gigerenzer and Brighton, 2009).

The bias–variance dilemma is a well-known phenomenon in machine learning (Hastie et al., 2001). The term 'bias' does not refer to the kinds of deviations from classical norms of rationality that have been documented in the heuristics-and-biases research tradition (Kahneman, 2011). Rather, it refers to the theoretical bet that all inductive strategies, including heuristics, make about the structure of the environment. Bias is one of three components of the

prediction error produced by any inductive strategy. The total amount of error is the sum of these three components:

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

The bias and variance components of the total prediction error can be understood by assuming an underlying (true) function that a given induction algorithm aims to learn. The algorithm's window on the true function is a (potentially noisy) data sample generated by this function. Averaged across all possible data samples of a given size (see the size of a training set in Figure 3), the algorithm's bias equals the difference between the true function and the mean function inferred by the algorithm on the basis of various data samples. If the mean function hits the mark perfectly (i.e., the true function), bias will be zero. Variance, in contrast, quantifies the algorithm's responsiveness to the specific contents of the samples. It is the sum-squared difference between the mean function and the individual functions extracted from each of the samples.

Ignoring noise, the total error of any induction algorithm depends on these two components, and on how the algorithm trades them off. Even an unbiased algorithm may suffer from substantial errors to the extent that the individual functions, possibly derived from few observations (see Figure 3), involve a great deal of variance. Therefore, seeking low bias will not always be adaptive for an organism. By means of illustration, Figure 4(a) plots the mean daily temperature for London in the year 2000, as well as two polynomial models aiming to capture the true temperature function underlying the data: a relatively simple degree-3 and a more complex degree-12 polynomial model. Figure 4(b) plots the mean error resulting from inducing the true function based on random samples of 30 daily observations (training set) as a function of the degree of the polynomial model. A simple relationship emerged: The higher the degree of the polynomial model, the better the model 'fits,' that is, the better it describes the mean function from the training set. Here, more is better. In predicting the observations not included in the training set, however, more is not always better. Instead, the relationship between the degree of the polynomial model and the models' predictive accuracy is U-shaped, with high-degree (>9) and low-degree (1) polynomials suffering from large errors (due to variance). The best predicting polynomial model has degree 4.

The bias–variance dilemma is one framework that provides insights into when and why simple heuristics are successful in some environments, but not in others (Todd et al., 2012). The notion of ecological rationality and Simon's scissors metaphor emphasizes the importance of the match between the structures of simple heuristics and the structures of the environment. Heuristics can be surprisingly accurate relative to computationally more complex strategies when their structure matches that of the environment, and the bias component of error is thus successfully kept at bay; the bias component will not, however, be within acceptable limits when heuristics and environment mismatch. The variance component is at least as important for the success of heuristics. Variance is likely to be the leading source of error when the available observations about the environment are sparse. Controlling this source of error is therefore crucial for

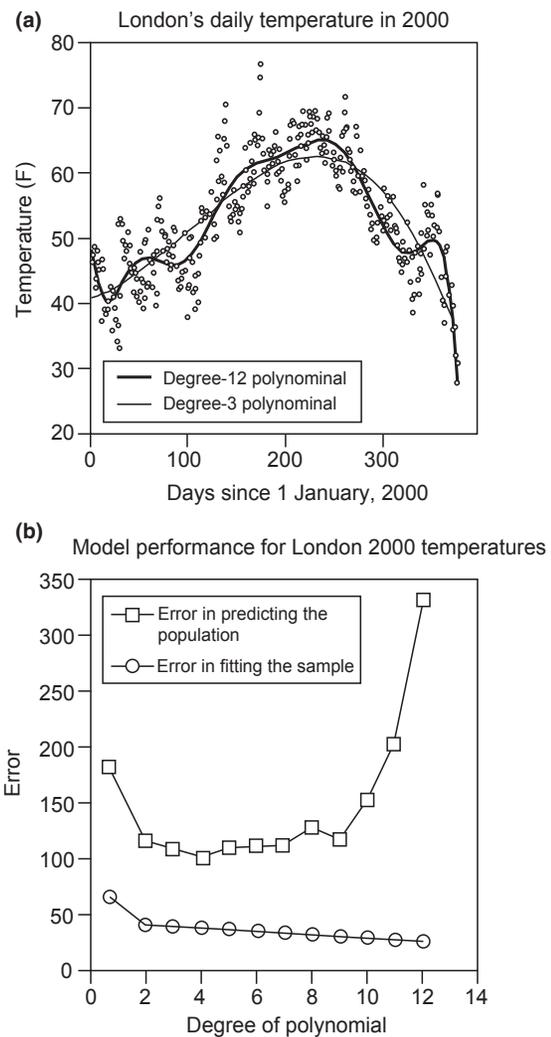


Figure 4 (a) London's mean daily temperature in the year 2000, along with a degree-3 polynomial and a degree-12 polynomial. The two polynomial models were fitted using the least squares method. (b) The mean error in fitting samples of 30 observations and the mean prediction error of the same polynomial models, as a function of degree of polynomial. Adapted from Gigerenzer, G., Brighton, H., 2009. Homo heuristics: why biased minds make better inferences. *Top Cognitive Science* 1,107–143.

any inductive algorithm. An algorithm with many adjustable parameters – and thus equipped with great flexibility to accommodate myriad true underlying functions and keep the bias component at bay – is likely to incur substantial prediction error when the window on the environment (the data sample) is fogged up. With sparse observations, flexibility via parameters can become a burden and cause 'overfitting' of the function suggested by the small data samples. In contrast, by forgoing high flexibility and not extracting as much information as is possible from a limited sample of data, simple heuristics control the variance component.

These analyses have suggested a new answer to the question of why humans have learned to use simple heuristics that ignore part of the information and forgo complex processing: In

uncertain worlds that often afford organisms only limited and costly chances to make observations, heuristics appear to achieve robust inferences.

Taxonomies of Cognitive Heuristics

There are various ways to map the mind's inventory of heuristics, the building blocks of which they are composed, and the evolved capacities that they co-opt. One taxonomic dimension is the specific task a heuristic is designed to solve. Across different research traditions in psychology, cognitive heuristics have been proposed for tasks as diverse as probability and frequency judgments in situations of uncertainty (Griffin et al., 2002), preferential choices under risk and uncertainty (Payne et al., 1993), inferential choices and estimates (Gigerenzer et al., 1999), sequential search (Gigerenzer et al., 1999), moral judgments and behavior (Hertwig et al., 2013), and resource allocations in social games (Hertwig et al., 2013), as well as for a wide variety of professional tasks, ranging from predicting criminal offenders' whereabouts to customer loyalty (Gigerenzer et al., 2011).

Another way to carve up the mind's inventory of heuristics is in terms of the domains in which they can operate. Hertwig et al. (2013) proposed that two broad domains be distinguished: 'games against nature' and 'social games.' Games against nature refer to situations in which a person needs to predict or outwit nature (e.g., identifying a good enough location to set up camp or predicting tomorrow's weather). The outcome a person experiences in these games is determined jointly by his/her decision and the true state of nature. In social games, in contrast, how well a person performs depends not on a dispassionate other such as nature, but on the decisions of other players who pursue their own interests. Hertwig et al. (2013) distinguished heuristics according to whether they are bound to games against nature (e.g., the ancient rule of thumb, "Red sky at night, sailors' delight; red sky in the morning, sailors' warning"), to social games (e.g., tit-for-tat), or are able to travel across social and nonsocial worlds (e.g., take-the-best and the N^{-1} heuristic).

Still another taxonomic dimension is the route through which a heuristic becomes part of a person's repertoire of cognitive strategies. Heuristics can be learned individually (through trial and error), socially (e.g., from parents and peers), and through evolution (e.g., possibly sequential search heuristics in the context of mate choice; Gigerenzer et al., 1999). They can also be explicitly designed as cognitive tools to help professionals make better decisions in the real world. For instance, fast-and-frugal trees that follow the template of heuristics such as take-the-best (Gigerenzer et al., 1999) can be constructed to foster accurate medical diagnoses (Jenny et al., 2013). Due to their simplicity, such trees can be easily memorized and understood and could be taught in medical schools.

Further Directions

The study of heuristics is an interdisciplinary and time-honored enterprise, with heuristics being examined in a wide range of domains, including applied fields such as operations research

and marketing. In psychology, two influential research traditions have investigated the cognitive heuristics that people use to make inferences, choices, and judgments in social and physical environments. One tradition has focused on the circumstances under which heuristics cause people to deviate from classical norms of rationality (Kahneman et al., 1982). Another tradition has emphasized that heuristics are neither good nor bad per se, nor are they always inferior to more complex strategies. Instead, the performance of a heuristic depends on its match with the structure of the environment ('ecological rationality'; Gigerenzer et al., 1999) and the amount of knowledge a decision-maker has about that environment (bias-variance dilemma). Future research should continue to explore the structures of decision environments and choice architectures that foster the usefulness of heuristics. Relatedly, the psychological issue of how decision-makers choose between different heuristics needs to be better understood. Finally, the insights into the interplay between the design of environments and heuristics should be implemented to help professionals in the real world – who often operate under conditions of time pressure and incomplete information – to make better decisions.

See also: Decision Making, Psychology of; Decision Making: Nonrational Theories; Decision and Choice: Bounded Rationality; Decision and Choice: Heuristics; Heuristics in Social Cognition; Moral Reasoning in Psychology; Risk and Uncertainty in the Social Sciences: Implications for Social Work Theory and Practice.

Bibliography

- Axelrod, R., 1984. *The Evolution of Cooperation*. Basic Books, New York.
- Coombs, C.H., Dawes, R.M., Tversky, A., 1970. *Mathematical Psychology: An Elementary Introduction*. Prentice-Hall, Englewood Cliffs.
- Dawes, R.M., 1979. The robust beauty of improper linear models in decision making. *American Psychologist* 34, 571–582.
- Duncker, K., 1935. *Zur Psychologie des produktiven Denkens* [The Psychology of Productive Thought]. Springer, Berlin.
- Gigerenzer, G., Brighton, H., 2009. Homo heuristicus: why biased minds make better inferences. *Top Cognitive Science* 1, 107–143.
- Gigerenzer, G., Hertwig, R., Pachur, T. (Eds.), 2011. *Heuristics: The Foundations of Adaptive Behavior*. Oxford University Press, Oxford.
- Gigerenzer, G., Todd, P.M., ABC Research Group, 1999. *Simple Heuristics That Make Us Smart*. Oxford University Press, New York.
- Goldstein, D.G., Gigerenzer, G., 2002. Models of ecological rationality: the recognition heuristic. *Psychological Review* 109, 75–90.
- Griffin, D.W., Gilovich, T., Kahneman, D., 2002. *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge University Press, Cambridge.
- Hastie, T., Tibshirani, R., Friedman, J., 2001. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York.
- Hertwig, R., Herzog, S.M., Schooler, L.J., Reimer, T., 2008. Fluency heuristic: a model of how the mind exploits a by-product of information retrieval. *Journal of Experimental Psychology, Learning* 34, 1191–1206.
- Hertwig, R., Hoffrage, U., The ABC Research Group, 2013. *Simple Heuristics in a Social World*. Oxford University Press, New York.
- Jenny, M.A., Pachur, T., Williams, S.L., Becker, E., Margraf, J., 2013. Simple rules for detecting depression. *Journal of Applied Research in Memory and Cognition* 2, 149–157.
- Kahneman, D., 2011. *Thinking, Fast and Slow*. Farrar, Straus & Giroux, New York.
- Kahneman, D., Slovic, P., Tversky, A. (Eds.), 1982. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge University Press, New York.

- Katsikopoulos, K.V., Schooler, L.J., Hertwig, R., 2010. The robust beauty of ordinary information. *Psychological Review* 117, 1259–1266.
- Newell, A., Simon, H.A., 1972. *Human Problem Solving*. Prentice-Hall, Englewood Cliffs.
- Payne, J.W., Bettman, J.R., Johnson, E.J., 1993. *The Adaptive Decision Maker*. Cambridge University Press, Cambridge.
- Pólya, G., 1945. *How to Solve It*. Princeton University Press, Princeton.
- Savage, L.J., 1954. *The Foundations of Statistics*. John Wiley & Sons, New York.
- Schooler, L.J., Hertwig, R., 2005. How forgetting aids heuristic inference. *Psychological Review* 112, 610–628.
- Simon, H.A., 1956. Rational choice and the structure of the environment. *Psychological Review* 63, 129–138.
- Simon, H.A., 1989. The scientist as problem solver. In: Klahr, D., Kotovsky, K. (Eds.), *Complex Information Processing: The Impact of Herbert A. Simon*. Erlbaum, Hillsdale, pp. 373–398.
- Simon, H.A., 1990. Invariants of human behavior. *Annual Review of Psychology* 41, 1–19.
- Stevens, J.R., Hauser, M.D., 2004. Why be nice? Psychological constraints on the evolution of cooperation. *Trends in Cognitive Sciences* 8, 60–65.
- Todd, P.M., Gigerenzer, G., The ABC Research Group, 2012. *Ecological Rationality: Intelligence in the World*. Oxford University Press, New York.
- Wertheimer, M., 1923/1938. *Untersuchungen zur Lehre von der Gestalt II [Studies on the theory of Gestalt II]*. *Psychologische Forschung* 4, 301–350.