

3 The Adaptive Toolbox

In a large world, there is no single best decision rule for all situations. Consider hiring, which involves much uncertainty about job candidates' future performance. Firms typically collect an array of information about the candidates, such as their education, personality, prior work experience, and social media activity, and factor all these elements into their hiring decision. Tesla founder and CEO Elon Musk, however, developed a very different approach. When Tesla was still a small company, Musk is reported to have used a heuristic that considers only a single cue.¹

Musk's hiring rule: If a candidate has exceptional ability, make a job offer; otherwise, do not.

This rule is an instance of a *one-clever-cue heuristic*, a type of heuristic discussed in more detail in this chapter. Musk's rationale was that someone who has shown exceptional ability in the past is likely to show it again. Hiring can also be based on social heuristics, such as *word-of-mouth*. The Korean owner of a Chicago janitorial and cleaning company relied on his own employees to identify good candidates.²

Word-of-mouth: Ask existing employees for recommendations of suitable candidates.

The rationale of word-of-mouth is that employees tend to recommend people who they know are reliable because they feel responsible for the new hire, and their own reputation is at stake.

The set of rules, including heuristics, that an organization or a leader has at their disposal constitutes their *adaptive toolbox*. This toolbox metaphor is in direct contrast to theories that postulate only one general tool for all problems, such as expected utility maximization and Bayesian updating. As the saying goes, to a hammer, everything looks like a nail; to these all-purpose

theories, everything looks like an optimization problem. The multiple tasks that exist in the real world, however, call for a diverse set of tools. The study of the adaptive toolbox thus addresses a *descriptive* question: What's in the toolbox?

To be a good decision maker, one must choose the right tool for the task at hand, similar to a builder who carries a toolbox and knows that a hammer goes with nails and a screwdriver with screws. This is the essence of *ecological rationality*: choose a tool that fits well with the demands of a task. For instance, Musk's hiring heuristic was an excellent choice when the company was small and in need of a vision for growth. At a later stage, with the company expanding, Tesla also needed employees who were good at routine work. Continuing to rely on exceptional ability alone would be counterproductive. Also, if fairness is a major concern in hiring, then the word-of-mouth heuristic used by the Korean owner may not be suitable, as the owner learned the hard way when he was sued for discrimination against non-Koreans. The study of ecological rationality thus addresses a *prescriptive* question: What heuristic should one use for a given task?

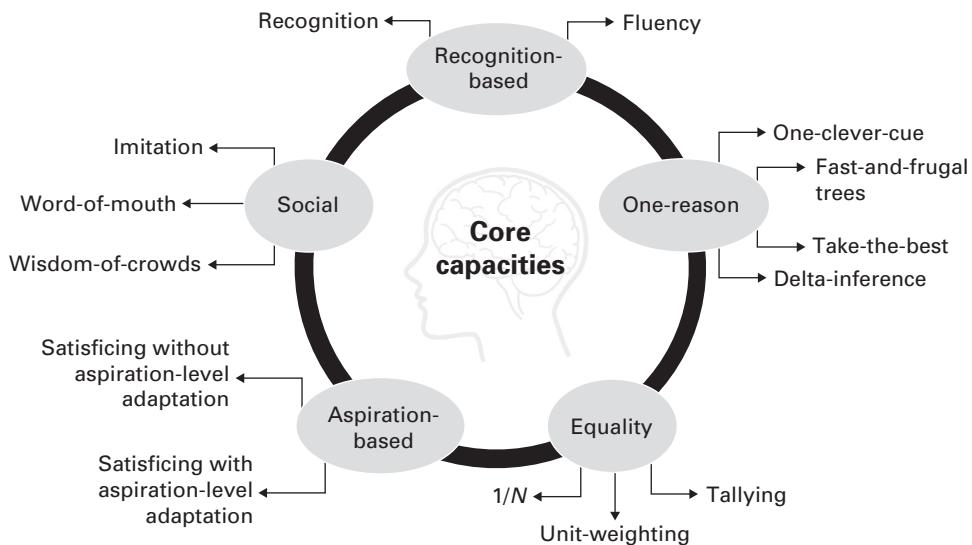
In general, the more experienced decision makers are, the more tools they have in their adaptive toolbox, and, importantly, the better their understanding of the tools' ecological rationality. Indeed, having a large repertoire of tools and being able to use them flexibly are hallmarks of intelligence. Let's now look at the major classes of heuristics in the adaptive toolbox.

Classes of Heuristics

Figure 3.1 lists five major classes of heuristics, along with specific examples in each class. These heuristics exploit core capacities of the human brain and recurrent features of physical and social environments. They have been studied extensively in the fast-and-frugal heuristics research program, and yet they are neither exhaustive nor in everyone's adaptive toolbox.

Recognition-Based Heuristics

Recognition is a core capacity of human memory and occurs with little conscious effort. Even the fact that one does not recognize an object can be informative. This is the rationale of the recognition heuristic.

**Figure 3.1**

Major classes of heuristics in the adaptive toolbox and examples of each.

Recognition Heuristic

Consider a situation that involves choosing between two alternatives. It could be a company choosing one of two possible local banks in a foreign market or a consumer choosing between two brands of shoes.

Recognition heuristic: If one alternative is recognized and the other is not, then choose the recognized alternative.

The ingeniousness of this heuristic is that it exploits semi-ignorance, that is, the fact that one has heard of one alternative but not the other. The power of this rationale has been shown, for instance, in predicting winners of matches in the 2003 Wimbledon tennis tournament.³ In the Gentlemen's Singles category, 128 players competed, resulting in 127 matches over seven rounds. To predict the winner of each match, one could use the official Association of Tennis Professionals (ATP) Champions Race Rankings or the ATP Entry Rankings, picking the player who is ranked higher (and who is usually also the higher seed in the tournament). The recognition heuristic, on the other hand, simply picks the player whose name is recognized. Tennis experts could not apply this heuristic because they recognized all the players; in contrast, amateur tennis players recognized only about

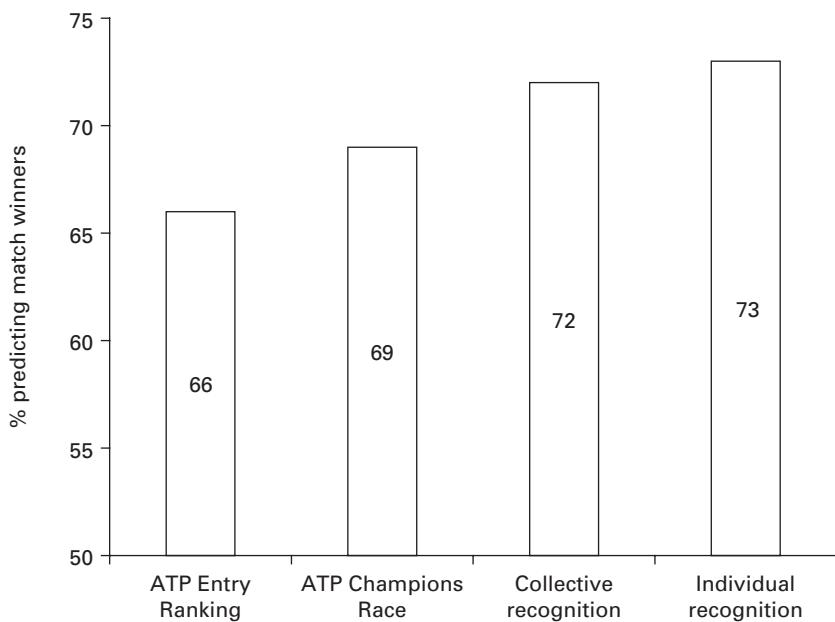


Figure 3.2

Amateur tennis players' recognition of the players predicted match winners in the 2003 Wimbledon tournament better than the official ATP rankings. ATP Entry Ranking and ATP Champions Race are two different rankings of players; collective recognition predicts that the player who ranks higher in name recognition among the amateurs wins the match; individual recognition predicts winners according to the recognition heuristic. Based on Serwe and Frings (2006).

half of the players and could apply the heuristic in about 40 percent of the matches. In those matches, an amateur's recognition on average predicted winners in 73 percent of the matches correctly, higher than the ATP rankings (figure 3.2). Because an individual cannot apply the recognition heuristic in all cases (i.e., when recognizing both or neither of the players), one can alternatively rely on the recognition rates of players among the amateurs. This collective recognition led to a 72 percent predictive accuracy.

The recognition heuristic works here because recognition is highly correlated with the players' performance. Similarly, brand name recognition is typically correlated with the quality of a product, and consumers rely on brand names, preferring those they have heard of, in product selection. When there are more than two products, brand name recognition is often

used to form a consideration set. Companies have tried to exploit consumers' reliance on recognition by investing in brand awareness rather than improving product quality. This tactic decreases the ecological rationality of using the heuristic for consumers because recognition then correlates more with the amount of advertising than with product quality.

Fluency Heuristic

The recognition heuristic relies on whether an alternative is recognized. The fluency heuristic, meanwhile, exploits the speed of recognition, choosing the alternative that is recognized faster. Therefore, it can be applied even if both alternatives are recognized. Fluency can also be used in situations where one has to generate options from memory, as in the case of handball players (see figure 2.1 in chapter 2). The fluency heuristic exploits the evolved ability of the human brain to detect subtle differences in recognition speed. Studies have reported that people can perceive the difference between recognition latencies exceeding 100 milliseconds.⁴ As explained in chapter 2, years of experience make the fluency heuristic ecologically rational: the first option that comes to mind is often the best.

One-Reason Heuristics

Usually, there are multiple reasons for or against the available options. As a result, decision makers can be flustered by the sheer amount of information that they have to deal with. *One-reason heuristics* show that this does not have to be the case.

There are two types of one-reason heuristics. One type looks for a single clever reason and bases its decisions on it: the *one-clever-cue heuristics*. Musk's hiring heuristic is an example. The second may search for more reasons but also bases its decisions on only one reason. These are *sequential search heuristics*.

One-Clever-Cue Heuristics

A *clever cue* is one that is so powerful that considering other cues (or reasons) does not improve performance but rather slows decision making or even decreases performance. Consider the problem of how baseball outfielders catch a fly ball. One possible solution is that they calculate the trajectory of the ball and run to the point where it will hit the ground:

$$z(x) = x \left(\tan \alpha_0 + \frac{mg}{\beta v_0 \cos \alpha_0} \right) + \frac{m^2 g}{\beta^2} \ln \left(1 - \frac{\beta}{m} \frac{x}{v_0 \cos \alpha_0} \right)$$

To calculate point $z(x)=0$ where the ball hits the ground, the player would have to estimate both the initial angle α_0 of the ball's direction relative to the ground and the initial speed v_0 of the ball, know the ball's mass m and friction β , set the gravity acceleration g as 9.81 m/s^2 , and be able to calculate tangent and cosine. Even then, the formula is overly simplified, in that it ignores wind and spin. Importantly, the true challenge is not computing the equation, but estimating its parameters, such as the initial angle and the initial speed.

Experienced players rely instead on simple heuristics. If the ball is high in the air, the gaze heuristic guides players to the ball.

Gaze heuristic: Fixate your eyes on the ball, run, and adjust your speed so that the angle of gaze remains constant.

Figure 3.3 shows that by keeping the angle constant, the player arrives at the location where the ball lands. The angle of the gaze is a clever cue. Players who rely on it need not estimate the trajectory of the ball; in fact, they can safely ignore all the factors necessary for calculating the trajectory.

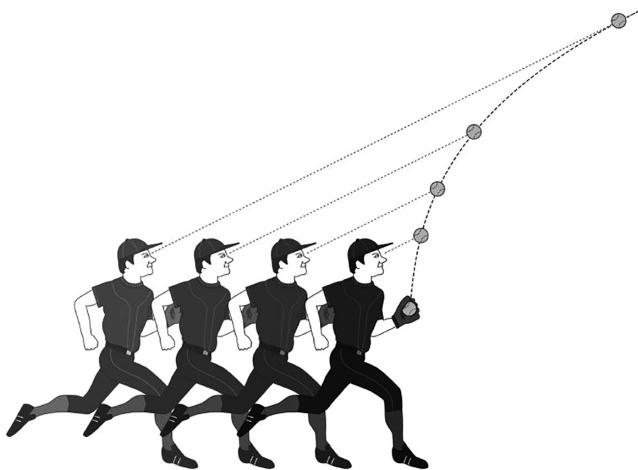


Figure 3.3

The gaze heuristic, a one-clever-cue heuristic, enables baseball players to catch a fly ball. To do so, the player adjusts the running speed so that the angle of the gaze remains constant. Various animals also use the heuristic to intercept prey and find mates. Source: Gigerenzer (2007).

The gaze heuristic was not invented by baseball outfielders. Bats, birds, fish, and other animals use it for hunting prey and finding mates.⁵ The heuristic was also built into an extremely successful autonomous guided weapon: the AIM-9 Sidewinder short-range air-to-air missile.⁶ The missile is an inexpensive but robust interception system whose “gaze” is directed at a point source of heat, which is the target. Although it was first employed in the 1950s, the AIM-9 Sidewinder is still in service in many countries, and new developments appear to be based on the same heuristic that maintains a constant angle of approach.

In the world of management, quite a few one-clever-cue heuristics can be found. Often, they are used to reject alternatives or to narrow down choices. Warren Buffett’s famous rule, “Never invest in a business you cannot understand,” specifies a single reason that is enough to exclude investing. Apple’s strategic rule, “Only enter markets where we can be the best,” is another case in point.

In their book *Simple Rules: How to Thrive in a Complex World*, the organizational scholars Donald Sull and Kathleen Eisenhardt described over 100 simple rules that people use in business strategy.⁷ Many of the rules are of the one-clever-cue kind. For example, after the collapse of the Soviet Union, a Russian private equity firm used strategy rules when making its investment decisions, including “work only with executives who know criminals but aren’t criminals themselves” and “invest in companies offering products a typical Russian family might purchase if they had an extra \$100 to spend each month.” The extent to which these rules are ecologically rational in other countries and times is open to investigation.

In cases where one clever cue is insufficient, a number of cues may be searched *sequentially*. Yet only one cue (reason) is used to make a decision. Fast-and-frugal trees, take-the-best, and delta-inference are examples of sequential search heuristics.

Fast-and-Frugal Trees

Emergency physicians must determine whether a patient needs immediate treatment or can be treated later; checkpoint soldiers must determine whether an approaching vehicle is friendly or carries a suicide bomb; and managers need to decide whether an employee should be promoted or not. Fast-and-frugal trees are tools for making such classification decisions. Unlike complex decision trees, a fast-and-frugal tree checks only a few cues or questions and tries to make a decision after each.

Fast-and-frugal tree: A simple decision tree with n cues and n + 1 exits.

It has three building blocks:

Search rule: Search through cues in a predetermined order.

Stopping rule: Stop the search if a cue leads to an exit.

Decision rule: Act according to what the exit specifies.

In an experiment, we asked managers to decide whether to keep or lay off a salesperson on the basis of their weekly sales performance.⁸ The mean, trend, and variation of the performance were visible in a chart summarizing the sales record. A rule that many managers adopted is the fast-and-frugal tree shown in figure 3.4. First, consider whether the person's mean sales performance is above average. If yes, the person is not laid off and no other

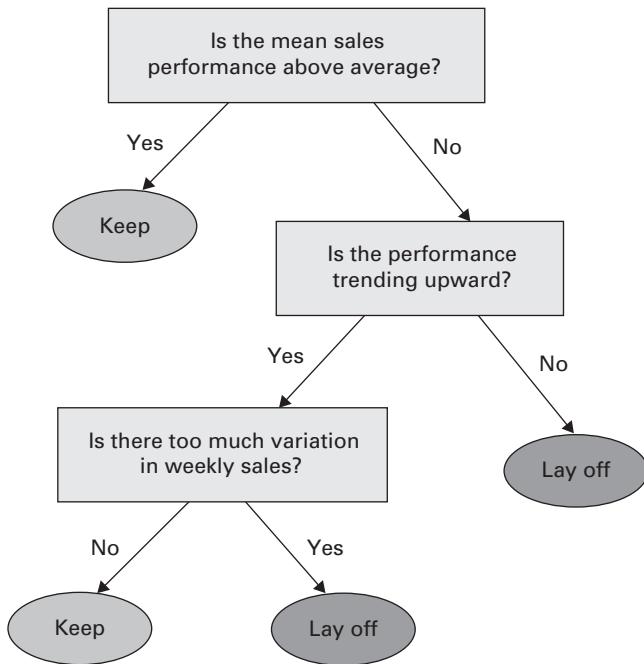


Figure 3.4

A fast-and-frugal tree used by managers to decide whether to keep or lay off a salesperson. If the mean sales performance is above average, the salesperson is kept. Otherwise, a second question about performance trend is asked, which may or may not lead to a decision. A third question about performance variation is asked in cases where the first two questions do not lead to a decision.

questions are asked. If the performance is below average, then the next question asked is whether the performance trend is upward. If not, the person is laid off; otherwise, a final question about the variation of sales is asked and a decision is then made. Unlike in a full decision tree, the order of cues is important in fast-and-frugal trees. The first cue can immediately lead to a decision, and the other cues cannot overturn that decision. For instance, a person with an above-average performance is kept even if the trend is downward and the weekly sales fluctuate widely.

Take-the-Best and Delta-Inference

Fast-and-frugal trees are heuristics for deciding on a single target (e.g., whether to fire an employee), whereas take-the-best and delta-inference are heuristics for choosing between two alternatives. Their logic and building blocks are otherwise similar to those of fast-and-frugal trees. The difference between the two is that take-the-best typically processes binary cues (e.g., whether a job candidate has a college degree), whereas delta-inference can process all types of cues, continuous, categorical, and binary (e.g., the candidates' IQ scores and their education levels). The *delta* in *delta-inference* refers to a threshold value above which the alternatives are judged to differ enough on a cue; this is when the search stops and a decision is made.

Consider the National Football League (NFL), the league for professional American football. In the US, it is the most popular sports league in terms of revenue generated, and NFL games are watched by millions every week throughout its playing season. The journalist Gregg Easterbrook used to write a football column called "Tuesday Morning Quarterback" for ESPN. In 2007, two readers wrote to him independently proposing a simple prediction model: the team with the better record wins; if the records are equal, then the home team wins.⁹ In essence, this model is an example of delta-inference, in which the first cue is the teams' win-loss record and the second is home team or away (see figure 3.5). The delta in the "win-loss record" cue is set at 0 (i.e., any difference will lead to a prediction), and the home-team cue is binary.

This simple heuristic beat all but one of the dozens of experts whose records were tracked by Easterbrook in the 2007–2008 season. It achieved almost the same feat in the 2008–2009 season, beating all but two experts.¹⁰ At times, Easterbrook questioned the picks made by the heuristic and replaced them with his own favorites. By doing so, the predictive accuracy was *lowered!* Using this heuristic, one does not need to have any insider information, spend time

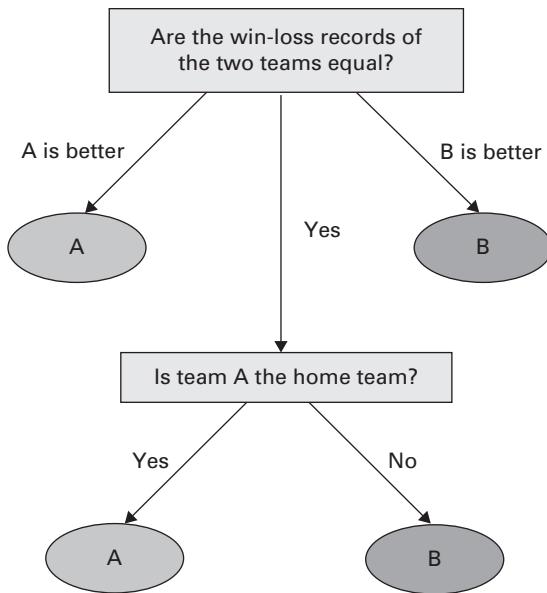


Figure 3.5

The delta-inference heuristic applied to predict winners of NFL games. Between two teams, the team with the better record is predicted to win the upcoming matchup. If the two teams' records are equal, then the home team is predicted to win.

reading reports and conducting sophisticated analyses, know the histories of the competing teams, or even understand the rules of American football. All the information needed can be found easily on any website publishing NFL game information.

One may attempt to improve the accuracy of delta-inference by trying to find the “optimal” deltas (i.e., deltas best fitted to past data). In a study of thirty-nine real-world problems—such as predicting which of two high schools would have a higher dropout rate and which diamond of a pair would sell for a higher price—we found that simply setting the deltas at 0 was just as accurate as using optimally fitted thresholds. The heuristic is also as good as complex models, such as Bayesian linear regression.¹¹

Equality Heuristics

One-reason heuristics work well when there is a powerful cue. However, in situations where the cues are similarly informative, *equality heuristics* are the

better choice. They integrate cues in a simple manner, such as by summing the reasons pro and con. This sets equality heuristics apart from optimization models that estimate weights for different reasons and take interdependencies and interactions among cues into consideration.

Tallying

Tallying is based on humans' core ability to count and compare numbers. It is a tool used to make classification decisions, and it works the same as fast-and-frugal trees but is based on the opposite logic. Instead of ordering cues and searching them sequentially, tallying treats all cues equally. Consider a task with n binary cues, where a *positive cue value* indicates category X and k ($1 < k \leq n$) is a classification threshold.

Tallying: Set a number k. If a target has k positive cue values or more, classify it as in category X; otherwise, do not.

In essence, tallying embodies democratic voting among cues. It is simple and transparent and can lead to highly accurate classifications. For instance, the avalanche researchers Ian McCammon and Pascal Hägeli designed a tallying rule called the "obvious clues method" to evaluate avalanche risk: The situation is classified as dangerous if more than three of seven clues are present.¹² These clues, such as whether there has been an avalanche in the past forty-eight hours and whether there is liquid water on the snow surface as a result of recent sudden warming, were derived from years of observations and are indicative of avalanche risk. When tested against eight more complex methods, the obvious clues method achieved the highest prevention rate (i.e., accidents that would have been prevented). Allan Lichtman's *Keys to the White House* model, which predicts which candidate will win the popular vote in the US presidential election, is another example.¹³ Since its first prediction in 1984, this tallying model has correctly picked all the winners with the exception of 2016, where it predicted that Donald Trump would win the popular vote (Trump won the presidency but not the popular vote).

Unit-Weighting

Organizations often use multiple linear regression to predict the values of a continuous variable, such as sales of a product. These models estimate the weights of cues to reflect their relative contributions. *Unit-weighting*, in contrast, weights all cues equally to reduce estimation error. At first glance,

unit-weighting seems to be a good example of the effort–accuracy trade-off: by dispensing with the effort of estimating cue weights one ends up with lower judgment accuracy. However, a landmark study by the psychologists Robyn Dawes and Bernard Corrigan showed that this is not the case. In three of four tasks they examined, including predictions of college students' grade point averages, graduate students' academic success, and patients' psychiatric diagnoses, unit-weighting was more accurate than multiple linear regression. In light of this finding, Dawes and Corrigan proclaimed that to make good judgments, "the whole trick is to decide which variables to look at and then to know how to add."¹⁴ Knowing the exact cue weights is of little value.

In assessing the personalities and attitudes of their employees, organizations often survey their potential or current employees using multiple-item scales that are weighted equally to form a composite score. Does it mean that answers to each item truly matter equally to the assessed underlying construct? Probably not. But there are two main reasons why unit-weighting is a good rule. First, the exact weighting scheme has little impact on the rankings of people being assessed. Second, the more items that are used, the larger the number of weights and correlations between items that need to be estimated, and the higher the estimation error. To avoid overfitting, it is thus reasonable to just weight the items equally.

1/N

Now consider a different type of problem: how to allocate limited resources to N alternatives, such as a limited budget to different divisions of a company or a limited amount of savings to different investment products. Once again, there are two visions for how to solve this problem. One is to get as much data as possible from the past, use the data to estimate weights for each alternative, and allocate resources according to the weights: that is, allocate more resources to those alternatives with greater weights. The other vision is for situations of uncertainty where the future is unlikely to be like the past. Here, one needs to simplify to avoid estimation error: that is, overfitting on the past. The 1/ N heuristic allocates equal amounts to all alternatives and uses the principle of diversification in the same spirit as tallying and unit-weighting. As mentioned in chapter 1, Harry Markowitz's mean–variance model reflects the first vision, whereas the 1/ N heuristic that he used for his own investments is in the spirit of the latter. Here, 1/ N has been shown

to perform on par with or better than the mean–variance and other highly sophisticated investment models, with considerably less time and effort.¹⁵

Beyond monetary allocations, the $1/N$ rule is also considered a fair way for parents or supervisors to distribute attention among children or employees. Interestingly, parents' use of the fair $1/N$ rule may wind up causing the *middle-born-child effect*: Growing up, middle-born children (e.g., the second child in a three-child family) receive fewer resources from their parents.¹⁶ Assuming that parents allocate resources equally among their children at any given time, the firstborn will receive all the resources before others are born, and the child born last will get all the resources after the older ones become more independent or leave home. The middle ones never have such opportunities and must share the resources all the time. Therefore, cumulatively, they receive fewer resources, despite their parents' intent to be fair. This counter-intuitive result exemplifies that the outcome of a heuristic depends on its environment (here, the number of siblings): if there are two children, the parents' goal of fairness is achieved, but not otherwise.

Aspiration-Based Heuristics

Kurt Lewin is often credited as the founder of social psychology. Among the countless discoveries that he made, one is the concept of aspiration, a goal that people are motivated to achieve. The concept was later borrowed by Herbert Simon and became the key ingredient of his well-known satisficing heuristic.

Satisficing

The heuristics introduced so far help one choose between two or a few alternatives. The *satisficing heuristic* can handle a large number of options, even in situations where one does not know how many alternatives exist. In its basic form, when evaluating options on just one attribute, such as price or expected profit, it has three steps:

Step 1: Set an aspiration level α and examine the options one by one.

Step 2: Choose the first option that satisfies α .

Step 3: If no option has satisfied α after time β , then change α by an amount γ and continue until a satisfying option is found.

If only the first two steps are used, the procedure is called *satisficing without aspiration-level adaptation*; if all three steps are used, it is *satisficing with*

aspiration-level adaptation. In business, satisficing is used for pricing commodities. An analysis of over 600 German used car dealers showed that 97 percent of them relied on satisficing, with or without aspiration-level adaptation. The most frequent strategy was to begin with an average price, lower the price after about four weeks, and repeat the procedure until the car was sold.¹⁷

The basic form of satisficing can be easily generalized to more than one attribute by setting an aspiration level for each attribute. Suppose that a venture capital firm wants to invest in a start-up in an emerging field and is concerned with three attributes: the excellence of the company's five-year vision, the proportion of engineers among all employees, and the charisma of its founders. Using the satisficing heuristic, the firm sets an aspiration level on each attribute, starts the search, and settles on the first start-up that meets all the aspirations.

There may be better alternatives out there. But two factors besides uncertainty make satisficing a good rule: the cost of making a search and market competition. When search is a necessary part of the decision-making process, it generally imposes a cost, as most people who have bought a house can testify. The satisficing heuristic effectively sets a stopping rule on the search and prevents the search cost from getting out of control. Moreover, good things are desired by many; to get them usually involves competition. If one keeps searching without making a decision, good opportunities will likely be gone, picked up by others. Monetary investments, house buying, and mate choices are all like that. Therefore, it is imperative for one to know what one wants and act fast when a good alternative is obtainable.

The so-called secretary problem bears a resemblance to the two-step satisficing heuristic, but it assumes a small world where the number of options n is known (and n is not very large, to avoid endless search). In this problem, a company aims to find the best secretary by interviewing candidates one by one and deciding whether to give a candidate an offer immediately after the interview. Once a candidate has been rejected, the company cannot recall that candidate at a later point. When the total number of candidates is known, the solution that maximizes the probability of getting the best secretary is to interview the first 37 percent of the candidates without making an offer and then keep interviewing until a candidate with a higher quality is found. However, if the number of candidates is unknown and the goal is to select an excellent instead of the best secretary (e.g., top 10 percent), then a simpler solution works better. It is called "Try a dozen," in which 37 percent

is replaced by a fixed number, 12. This heuristic has a higher chance of finding a suitable secretary with a substantially reduced search time.¹⁸ Interestingly, after his first wife died, the astronomer Johannes Kepler is recorded to have considered eleven women as possible replacements before making his final choice. He had a productive second marriage with seven children, during which he wrote four more major works.

Social Heuristics

All the heuristics introduced so far in this chapter can be used to solve social and nonsocial problems. For instance, satisficing can be used to choose a house to buy, but also to choose a partner to marry. Yet there is another class of heuristics that are genuinely social, in that they rely on social information only. We introduce here three kinds of social heuristics: *imitation*, *word-of-mouth*, and *wisdom-of-crowds*.

Imitation

Imitation is an enabler of human culture. No other species copies the behavior of others so generally and precisely as humans. From a very young age, children can already imitate the actions of others and understand their intentions, emulate the behaviors of adults and peers as a means to learn and affiliate with in-group members, and conform to majority behavior and social rules. Chimpanzees also imitate, but only occasionally and much less skillfully.¹⁹ Learning by imitation not only helps children survive in an unfamiliar, uncertain, and possibly dangerous world but also instills stability in human groups and facilitates the passing of knowledge and social norms over generations.

Companies also frequently engage in the same kind of social learning by imitating other companies' successful products and technologies. For example, Amazon released Echo, an Internet-based home assistant device, in 2015. It was a huge success despite concerns over privacy. One year later, Google released a remarkably similar product called Google Home, and Apple did the same with its own Home Pod in 2017. Imitation provides a quick and relatively safe way for companies to enter markets. Instead of investing millions or billions in a cutting-edge technology whose commercial success is uncertain, companies can simply copy and improve on a market-tested idea, reducing the probability of failure. That said, blatant imitation without at

least some attempt at differentiation (e.g., a lower price or new features) can hurt a copycat by putting it in a disadvantaged “late-mover” position as well as the entire market by preventing the invention of better technologies and products.

Word-of-Mouth

With the *word-of-mouth heuristic*, one’s decisions are based on others’ recommendations, as stated at the beginning of this chapter. Companies use it to find good hires and reliable business partners, job seekers to narrow down potential employers, and consumers to decide where to dine and what to buy. To be successful, word-of-mouth requires a relation of trust and long-term dependence between the person who asks and the one who recommends. It stops working when that trust is abused, especially when recommenders have other goals in mind than providing the most truthful information or suitable alternatives.

Wisdom-of-Crowds

In a short paper published in the journal *Nature*, Sir Francis Galton reported the first documented case of the wisdom-of-crowds.²⁰ Around 800 people bet on the weight of a dressed ox at a country fair in Plymouth, England. Galton gathered all the betting tickets and found that the mean of the estimates was only one pound off the actual weight.

Wisdom-of-crowds: Estimate a quantity by averaging the independent judgments of many people.

The foundation for the wisdom-of-crowds is the law of large numbers in statistics: the larger a sample, the closer the mean of the sample to the true value. A key condition for the mean judgment to be accurate is that individual estimates are independent. If they are influenced by others—for example, a vocal leader—the estimates will not be independent and the mean can be biased, as in groupthink. In business, leaders too often utter their own opinions first, which affects what subordinates say (or even think) and makes wisdom-of-crowds no longer an ecologically rational heuristic. To avoid this pitfall, another heuristic can be useful: *first listen, then speak*. This heuristic is for the leaders, not the subordinates. It makes collecting the fruits of wisdom-of-crowds possible.

In the age of the Internet and social media, people increasingly rely on user ratings of restaurants, books, and many other products to make choices,

hoping to harness the wisdom of crowds. If these ratings are made independently and without bias, then they will be good guidelines. Yet these conditions are not always in place. A 2021 report shows that among all the online reviews posted in 2020, 31 percent were estimated to be fake.²¹ One source of the fake reviews is “bot farms,” which manipulate rankings, stars, likes, and hearts for a fee.

Ecological Rationality of Heuristics

Could it be that Elon Musk makes better choices on the basis of a single reason than using many reasons, or an entire assessment center? The study of the ecological rationality of one-reason heuristics gives an answer to this question—and it is yes. One can prove that there are conditions under which relying on one reason is as good as or better than considering more information. The dominant-cue condition (discussed next) is one. Yet the study of ecological rationality also prescribes when other classes of heuristics are expected to be successful. We have already mentioned some of these conditions. Here, we focus on two general results. The first shows that the distribution of cue weights provides a guideline for choosing heuristics from the adaptive toolbox, and the second explains why simple heuristics can predict better than complex models in situations of uncertainty.

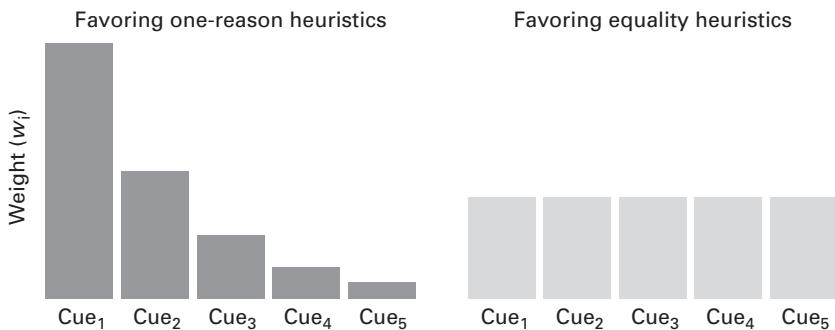
Dominant and Equal Cues

Cues drive both the absolute and the relative performance of heuristics. Generally, one-reason heuristics are ecologically rational in conditions where a dominant cue exists, whereas equality heuristics are ecologically rational when cues are equally valid. To understand why, let us consider the situation in which n binary cues are available to make a binary decision, such as hire or do not hire.

A linear model that weights and adds all cues has the form

$$\gamma = w_1x_1 + w_2x_2 + \dots + w_nx_n$$

where γ is the criterion variable, x_i is the value of a cue i ($i = 1, \dots, n$), and w_i is the decision weight of a cue that is ordered and reflects a cue's relative contribution after a cue or cues of higher ranks are considered. To simplify, all weights are positive. The model prescribes “hire” if γ is positive; otherwise, “do not hire.”

**Figure 3.6**

Distributions of cue weights (w_i) under which one-reason heuristics and equality heuristics are ecologically rational, respectively. Left: The dominant-cue condition that favors one-reason heuristics. Right: The equal-cue condition that favors equality heuristics. Based on Gigerenzer et al. (2022).

This linear model cannot make decisions more accurately than a one-clever-cue heuristic that bases its decisions solely on the most valid cue (i.e., Cue₁), if the sum of the weights of all other cues is smaller than the weight of Cue₁—thus, it is impossible for other cues to overturn the decisions made by Cue₁.²² This is referred to as the *dominant-cue condition*, in which values of cue weights are such that

$$w_1 > \sum_{i=2}^n w_i$$

The left side of figure 3.6 shows an example of such a condition, in which the five cue weights are 1, 1/2, 1/4, 1/8, and 1/16. It is also an example of a stronger version of the dominant-cue condition where the weight of any cue is larger than the sum of the weights of subsequent cues. In this condition, it is guaranteed that one-reason sequential heuristics, such as take-the-best and fast-and-frugal trees, can never be outperformed by a linear model either.²³

When all the cue weights are equal, as illustrated on the right side of figure 3.6, it is clear that one-reason heuristics cannot work better than equality heuristics such as tallying. In this equal-cue condition, no cue is better than any other; therefore, to make a good decision, one needs to take all cues into account. This is also the condition where no linear models that weight cues differentially can outperform tallying.

When cues are highly correlated, the dominant-cue condition is more likely to hold, as information added by other cues besides the most valid one is

limited. In the aforementioned study that examined delta-inference in thirty-nine real-world problems, the top three cues in each problem tended to be highly correlated, and the dominant-cue condition held for most cases. That is the main reason why delta-inference with delta at 0, which decides almost exclusively on the basis of the most valid cue, did as well as linear regression across all problems. On the other side, when cues are independent, the equal-cue condition is more likely to hold. Although it is rare for cue weights to be precisely equal, equality heuristics can still be ecologically rational when cue weights do not differ much, or are difficult to estimate, because of the instability and uncertainty of the environment, the insufficiency of data, or both.

The Bias–Variance Dilemma

Take a look at figure 3.7. Two players threw darts at the board. Which player did better? Most would say player A. Yet this player has a clear bias: the darts are all to the lower-right side of the bull's-eye. Player B has no bias, as the average position of the darts is in the bull's-eye; however, the darts are all over the place and far from the target. This analogy helps explain why and when heuristics predict better than more complex models.

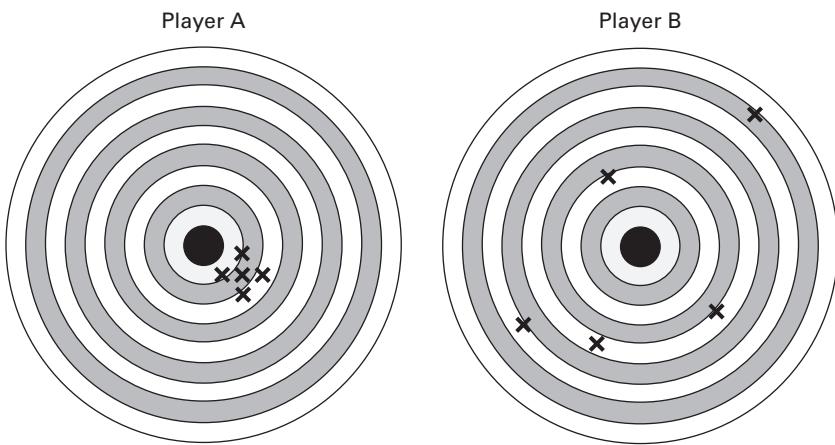


Figure 3.7

A dartboard illustration of the bias–variance dilemma. Player A's darts show a clear bias but only small variance, as the darts are all to the lower-right of the bull's-eye but close to each other. Player B's darts show no bias but considerable variance, as the average position of the darts is in the bull's-eye, but each dart is quite far from the others. Based on Gigerenzer et al. (2022).

The prediction error of a model comprises three components:

$$\text{Prediction error} = \text{bias}^2 + \text{variance} + \epsilon$$

where *bias* is the systematic difference between a model's mean predictions and a true value, *variance* reflects the sensitivity of a model to sampling error, and ϵ is the irreducible error caused by random noise.²⁴ In predicting the sales of a product, for example, a model makes a prediction x_1 based on one random sample of observations, x_2 on another sample, and x_s on sample s. The difference between the mean of these predictions \bar{x} and the true sales values μ is bias, and the variability of these predictions around \bar{x} is variance.

In a stable world and with an ample amount of data, one may find a model that has both a small bias and a small variance. In an uncertain world and with limited observations, however, a bias–variance dilemma is usually present: Models with fewer free parameters tend to have smaller variance but larger bias than models with more free parameters, analogous to the contrast between the two dart players. Heuristics such as $1/N$, one-clever-cue, and take-the-best have none, one, or only a few parameters to estimate. Therefore, benefiting from the smaller variances, they often have lower prediction errors than highly parameterized models, such as multiple regression and Bayesian models. The advantage is even stronger in conditions where heuristics have the same bias as complex models, such as the dominant-cue condition for the one-clever-cue heuristics.

Ecological rationality goes hand in hand with the adaptive toolbox: good decision makers need to have both a repertoire of decision tools and the ability to pick a tool that can handle a task well. In the next part of this book, we use this insight to take a closer look at the heuristics in the adaptive toolbox of organizations and leaders, as well as their ecological rationality.

Smart Management

How Simple Heuristics Help Leaders Make Good Decisions in an Uncertain World

By: Jochen Reb, Shenghua Luan, Gerd Gigerenzer

Citation:

Smart Management: How Simple Heuristics Help Leaders Make Good Decisions in an Uncertain World

By: Jochen Reb, Shenghua Luan, Gerd Gigerenzer

DOI: 10.7551/mitpress/14939.001.0001

ISBN (electronic): 9780262378574

Publisher: The MIT Press

Published: 2024



The MIT Press

© 2024 Massachusetts Institute of Technology

All rights reserved. No part of this book may be used to train artificial intelligence systems or reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Reb, Jochen, 1973– author. | Luan, Shenghua, author. | Gigerenzer, Gerd, author.

Title: Smart management : how simple heuristics help leaders make good decisions in an uncertain world / Jochen Reb, Shenghua Luan, Gerd Gigerenzer.

Description: Cambridge, Massachusetts : The MIT Press, [2024] | Includes bibliographical references and index.

Identifiers: LCCN 2023027153 (print) | LCCN 2023027154 (ebook) | ISBN 9780262548014 (paperback) | ISBN 9780262378567 (epub) | ISBN 9780262378574 (pdf)

Subjects: LCSH: Decision making. | Management. | Leadership.

Classification: LCC HD30.23 .R42 2024 (print) | LCC HD30.23 (ebook) |

DDC 658.4/03—dc23/eng/20230925

LC record available at <https://lccn.loc.gov/2023027153>

LC ebook record available at <https://lccn.loc.gov/2023027154>