



Preprints of the
Max Planck Institute
for Research on Collective Goods
Bonn
2004/15

Is there an optimization in bounded rationality?
The ratio of aspiration levels

Martin Beckenkamp



Is there an optimization in bounded rationality? The ratio of aspiration levels

Martin Beckenkamp

November 2004

Is there an optimization in bounded rationality?

The ratio of aspiration levels

Martin Beckenkamp

Abstract

Simon's (1955) famous paper was one of the first to cast doubt on the validity of rational choice theory; it has been supplemented by many more papers in the last three and a half decades. Nevertheless, rational choice theory plays a crucial role in classical and neoclassical economic theory, which presumes a completely rational agent. The central points characterizing such an agent are: (1) The agent uses all the information that is given to him. (2) The agent has clear preferences with respect to the results of different actions. (3) The agent has adequate competences to optimize his decisions. As an alternative to this conception, Simon (1955) himself suggests the concept of "bounded rationality". In this context, Simon (1956) discusses a principle, which he names the "satisficing principle" (for explanations with respect to this notion cf. Gigerenzer & Todd 1999, p. 13). It assumes that, instead of searching for an optimal action, the search for an action terminates if an alternative has been found that satisfies a given "aspiration level". It will be demonstrated that although the satisficing principle is nothing but a heuristic, there is a mathematical optimization at work when aspiration levels are used in this kind of problems. The question about the optimal aspiration level can be posed. Optimization within the framework of bounded rationality is possible. However, the way in which such an optimization can be achieved is very simple: Optimal thresholds in binary sequential decisions rest with the median.

∇ Helpful comments by Thomas Gaube, Hendrik Hakenes and Christoph Engel and the linguistic support by Darrell Arnold are gratefully acknowledged.

Rational Choice and Bounded Rationality

Rational Choice theory, an integral tool of classical and neoclassical economics, assumes that humans make rational decisions in accordance with the principles of *homo oeconomicus*. In a set of different alternatives of actions A or a set of considered alternatives of actions $\check{A} \subseteq A$, a rational agent searches for the alternative, $a \in A$ or $a \in \check{A}$ respectively, yielding the best possible outcome.¹ In Rubinstein's attempt to "uncover some of the assumptions buried in" what he calls "the rational-man procedure" (Rubinstein 1988, p8f.), he mentions various characteristics of *homo oeconomicus*: He is *completely informed about the problem* he faces and has knowledge about the set of the possible states of affairs S and about the set of states of affairs $S_a \subseteq S$ that could result from his actions. He has *clear preferences*. The preferences of the agent $V(s)$ with respect to different concrete states of affairs $s \in S$ are completely ordered. He has the *ability to optimize*. The agent must be able to choose the alternative that ensures him the result he prefers most. This implies that his ability to calculate is *unlimited*, and he does not make mistakes. Finally he is indifferent to logically equivalent descriptions of alternatives and choice sets. These assumptions had largely already been discovered and clarified, at least in rough outlines, by Simon (1955) who characterizes economic man as follows:

Traditional economic theory postulates an "economic man," who, in the course of being "economic" is also rational. This man is assumed to have knowledge of the relevant aspects of his environment which, if not absolutely complete, is at least impressively clear and voluminous. He is assumed to have a well-organized and stable system of preferences, and a skill in computation that enables him to calculate, for the alternative courses of action that are available to him, which of these will permit him to reach the highest attainable point on his preference scale.
(Simon 1955, p. 99).

He was one of the first to become suspicious of the usefulness of these assumption for the theoretical foundations of economics. In Rubin's words:

The refutation of the rational man paradigm by experimental data is not new. As early as 1955 Simon asserted "Recent developments in economics ... have raised great doubts as to whether this schematized model of economic man provides a suitable foundation on which to erect a theory – whether it be a theory of how firms *do* behave or how they 'should' rationally behave." Since then, a great deal of additional experimental evidence has been accumulated, mainly by psychologists.
(Rubinstein, 1998, S. 16; citing Simon 1955, p. 99).

Instead of bearing up the assumptions of rational choice, Herbert Simon (1955) introduced the satisficing-principle in order to accommodate constraints with regard to rational choice. The satisficing-principle is based on two essential ideas. The first idea is a radical simplification of the evaluation function $V(s)$ with respect to the range of possible results. The only valid result of the evaluation function $V(s)$ are either the values 1, and -1 or the values 1, 0 and -1. In the first case, Simon interprets these values as satisfactory or unsatisfactory, in the latter case as a win, a draw or a loss, like in a chess game. As an example using the dichotomous evaluation function,

1 The notation chosen here follows Simon 1955, which is different from Rubinstein (1998). This is why, in the following citations of Rubinstein, his formal explanations have been omitted.

Simon describes a situation where an individual wants to sell a house. In this example the role of aspiration levels within the satisficing principle as one conceptualization of bounded rationality is rather clear:

As an example ... let S represent possible prices for a house an individual is selling. He may regard \$15,000 as an “acceptable” price, anything over this amount as “satisfactory”, anything less as “unsatisfactory.” In psychological theory we would fix the boundary at the “aspiration level”; in economic theory we would fix the boundary at the price which evokes indifference between selling and not selling (an opportunity cost concept).
(Simon 1955, p. 104f.).

Another feature of this example is also important, and Simon discusses it when he replies to some anticipated caveats. We are often in situations where decisions have to be made and we don’t have simultaneous access to the different alternatives at our disposal. Rather, we have sequential access to the different options. This feature will be decisive for the argument from probability theory which follows below. This argument, from Simon, demonstrates that the assumption of sequential decision situations is not artificial at all:

The objection may be raised that, although \$16,000 and \$25,000 are both “very satisfactory” prices for the house, a rational individual would prefer to sell at the higher price, and hence, that the simple pay-off function is an inadequate representation of the choice situation. The objection may be answered in several different ways, each answer corresponding to a class of situations in which the simple function might be appropriate.

First, the individual may not be confronted simultaneously with a number of buyers offering to purchase the house at different prices, but may receive a sequence of offers, and may have to decide to accept or reject each one before he receives the next. (Or more generally, he may receive a sequence of pairs or triplets or n -tuples of offers, and may have to decide whether to accept the highest of an n -tuple before the next n -tuple is received.) Then, if the elements S correspond to n -tuples of offers, $V(s)$ would be 1 whenever the highest offer in the n -tuple exceed the “acceptance price” the seller had determined upon that time. We can then raise the further question of what would be a rational process for determining the acceptance price.
(Simon 1955, p.105).

In the following, it will be demonstrated that although the satisficing principle is nothing but a heuristic, there is mathematical optimization at work when aspiration levels are used for these kind of problems. After the introduction of the principal argument – that there is a pertinent rationality in this kind of bounded rationality – we shall deal with the question whether and how an optimal aspiration level can be determined.² In other words, in the following it will be shown that *sequential decision-making* using aspiration levels will *always* lead to advantages over mere guessing, even when no information about the underlying probability functions is available.

2 The problem of rationally determining an adequate aspiration level (in Simons words an “acceptable pay-off”) is also discussed in Simon’s (1955) appendix, but with the assumption that the probability distribution of offers is known. The argument that is presented in this paper will presume much weaker assumptions and offer interesting insights about the adequate use of aspiration levels.

The rationality of using aspiration levels

Thomas Bruss (2000) presents a mathematical idea, beginning with an illustration in which an individual wants to sell a house. The example is similar to the example given in Simon (1955), and it, like Simon's piece, is intended to refer to an abstract problem-structure that we encounter in everyday life. The problem structure underlying this class of problems can be grasped in an abstract game:

We are often embarrassed to snap at the chance without knowing whether the things which follow are better. For the sake of clearness, let's assume the scenario in an abstract game. You ask your son and your daughter to pick a number (really any number you want, big or small, negative, with a decimal point, anything goes), each one in secrecy, to write it on a sheet of paper and to put it on the table covered. There is no consultation between your son and your daughter. You turn around the sheet of your son and have the choice to accept the number or to refuse it. If you refuse, you automatically get the number from your daughter. Afterwards both numbers have to be compared with each other. If you have chosen the bigger number you win, otherwise you lose. The difference between the numbers is irrelevant, what counts is to win the game. If both numbers are equal, which is very unlikely, the game will be played once again. In order to exclude from consideration that you can infer the chosen numbers due to the fact that you know your children, replace them in your thoughts by strangers. It could also be one person filling out two sheets.

Now this really seems to be a game of pure chance, with a 50% probability of winning. But now the surprise: Here is a strategy which may increase your probability to win.

(translated from Bruss 2000, p. 106).

In order to show how the probability to win can be increased, let us call the number which is drawn at first X and the second number Y . Z gives our aspiration level.³ Now the strategy works as follows: If the first number surpasses Z (that is $X > Z$), then choose X ; otherwise choose Y . Thus three different possible events can be distinguished:

- (A) Both X and Y are lower than or equal to Z ,
- (B) Z lies between X and Y (inclusive),
- (C) Both X and Y are greater than Z .

In Case (A) you choose Y ; corresponding to that, in case (C) you choose X . In both cases your probability to win is 50%, because your decision results from a random choice. But in case (B) you win with certainty! You automatically choose the higher of the two numbers, independent of whether it is X or Y . Thus the total probability of winning is $w = (a+c)/2 + b$, where a , b and c give the probabilities of the events (A), (B) and (C). Since one of these occurs with certainty it holds that $w = (a+b+c)/2 + b/2 = 1/2 + b/2$ ⁴. Your probability of winning is augmented by $b/2$, compared to a pure guess, and b is greater than zero. It is absolutely possible that your Z value lies between X and Y . (translated from Bruss 2000, p. 106).

3 Please note that this aspiration level is similar to the one in Simon's example about house-selling, where the success or failure of a decision is measured by the binary function $V(s)$. You either win or lose with your decision. The selling-price itself is irrelevant. Arguments concerning why such a conceptualization makes sense can already be found in Simon's (1955) paper itself.

4 This is because $w = (a+c)/2 + b = a/2 + c/2 + b/2 + b/2 = (a+b+c)/2 + b/2$

Thus it can be maintained that it is always advantageous to use aspiration levels in sequential decision situations, even if no further assumptions can be made with respect to the random process which leads to the numbers X and Y . In the worst case we hold up the probability of a pure guess, but the use of aspiration levels can also increase the probability to win. This provides an essential argument for the use of aspiration levels in heuristics like Herbert Simon's satisficing principle!

Now Bruss pursues the issue as to how this strategy might be optimized. He asserts that the probability b of the event (B) should be preferably maximized. Therefore Z should be chosen such that

the probability that Z lies between X and Y is maximized. But there is just no knowledge about the values of these two numbers. Therefore generally no definite recommendations can be given. But in concrete cases there are often good clues.
(translated from Bruss 2000, p. 106).

Is there an optimization within bounded rationality? About optimal aspiration levels in 2 choice-tasks

Only partial consent can be given to Bruss' argument. If, in contrast to Bruss, we don't conceptualize X and Y as concrete values, but as two different realizations of one random variable beyond, some more considerations with respect to an optimal value of Z can be added. In the extreme case, the probability distribution of the random variable may just consist in a distribution of the two numbers X and Y : but the probability distribution can also, for instance, be a normal distribution with a determined value of the mean and the variance. It is presumed that the realized values of X and Y are independent from the succession. This assumption is plausible for both the example given by Bruss and the example given by Simon, although in the latter case, in the midterm run – when a sequence consists of more than two selling offers – people may make use of the price development information. In other words: In this case people may make use of what they have learned before they make their decision.

Conceptualizing X and Y as two realizations of a random variable means that it is assumed that the values X and Y stem “from the same population” in the mathematical sense of the word. As a consequence of this conceptualization, it is possible to give a precise measure for Z , which optimizes the probability of winning in the sequential decision task earlier presented. For the sake of clarification: it is merely necessary to assume that X and Y are realizations of the same random variable; no further assumptions are needed about the concrete definition of the underlying distributions such as Binomial distribution, normal distribution, or anything else.⁵ Furthermore, no information about the form of the distribution – such as skewness or variance – need be given.

5 Again, this is in contrast to Herbert Simon's (1955) reflections about optimal levels, which he offers in the appendix of his paper. There he makes parametric derivations. The conclusions given here are non-parametric.

In this case, the median is the optimal value for Z needed in order to win the abstract game as often as possible. The median maximizes the probability of winning, because the probability that event (B) will occur (from the events (A), (B) and (C) given before) is maximized. Using the median as the aspiration level, there is a 50% chance that Z will lie between X and Y . This is because the probability $p(X \text{ is below } Z)$ is equal to the probability $p(X \text{ is above } Z)$, which is equal to 50% if the value of Z is the median. The same is true for Y with respect to $p(Y \text{ is below } Z) = p(Y \text{ is above } Z) = 1/2$, because Y is a realization of the same random variable. The consequence of taking the median as a value for Z is that, in 25% of all cases, both X and Y are below the median. The probability of winning is $1/2$ in this case (see above, case (A) in Bruss' example). In 25% of all cases both X and Y are above the median; here the probability of winning is also $1/2$ (case (C) in Bruss' example). In 50% of the cases Z lies between X and Y ; in this case you win with certainty (case (B) in Bruss' example). The probability of winning results from the sum of the products, where the products are composed of the probability of the occurrence of a case, multiplied by the probability of winning in this case:

$$p_{\text{win}} = (p_{\text{case A}})(p_{\text{win in case A}}) + (p_{\text{case B}})(p_{\text{win in case B}}) + (p_{\text{case C}})(p_{\text{win in case C}}) = (1/4)(1/2) + (1/2)(1) + (1/4)(1/2) = 3/4.$$

Any other value for Z would decrease the probability that Z will lie between X and Y , and therefore it would also decrease the probability of winning. Thus choosing the median as a value for Z , or, in other words, choosing an aspiration level which corresponds to the median, raises the probability of winning that is associated with a pure guess ($p = 1/2$) to a maximal possible winning probability of $p=3/4$.

To summarize: The question about optimal aspiration levels can be answered. In case of a choice between two alternatives in sequence, it is the median, as long as the goal is to win the game and the extent or magnitude of the gain is ignored. This conclusion can be drawn without any detailed knowledge of the probability distribution, which leads to the values X and Y . Compared to a pure guess, where no strategy is used, the probability of winning increases from $1/2$ to $3/4$. However, does this knowledge have any practical use? In other words, if the intention is to use this knowledge in the sequential game presented before, there is no chance to make use of it as long as there is no estimation of the median available. The problem of estimating the median, or more generally, of estimating parameters of a population, is a typical statistical problems.

This leads to an interesting question with respect to the learning mechanisms and the adaptations of heuristics in such choice tasks between two alternatives in sequence. When people have experiences, in a statistical sense they take samples. Such samples can be used to make estimations. In the case of the satisficing principle and Simon's example, learning the prices of former transactions should lead to adaptations in aspiration levels. In cases where only two choices are given in a sequence, such learning mechanisms⁶ would be optimal with respect to the problem if the adaptation of the aspiration level were such that it corresponded to the median. It's a moot ques-

6 Learning can take place by both our own experiences and by observing others' decisions.

tion whether people do indeed adapt their choices in this task by using an estimation of the median. Some general doubts arise in light of the knowledge that people learn differently under implicit and explicit learning conditions (cf. Sedlmeier and Betsch 2002). Therefore, it is possible to maximize the occurrence of such adaptations of aspiration levels to the median only under one of the two conditions.

This threshold is adequate and optimal when there is a sequential choice between two alternatives and the “game” is only about winning or loosing, ignoring the absolute values. In the context of ecological rationality, good arguments can be found that such games occur in the struggle to survive. Barrett and Fiddick (2000) argue that survival does not so much require that a certain expected value be kept as that a certain level, which guarantees survival with an utmost probability, be sustained. For instance, organisms without the ability to build stocks have to choose between different foraging options in order to get enough energy to survive the next night. Falling below the energy level means complete loss or death

However, not all problems have the underlying formal properties of the abstract game and Herbert Simon’s example. For instance, in some cases information on the probability distribution, which leads to the values of X and Y , is available, and the absolute value of gains or losses matters. In such cases the median is only optimal if the underlying distribution is symmetrical. However, if the absolute values don’t matter, the median is optimal even in asymmetric cases. This point is clear if we think about Bruss’ example. Let us, for instance, say that $Y = -1$. In this case it is irrelevant whether the value of X is 10 or 100. X is above Y in both cases; both 10 and 100 win; and 100 does not represent a better result than 10. However, in the other cases it is necessary to calculate the expectation (i.e. the expected value) of the distribution. Then it is necessary to calculate the sum of the probabilities (or densities) of each event multiplied by the respective values of the events. In this case, asymmetry leads to different optimal aspiration levels than the median: the mean becomes adequate in this case.

This is also the result of Simon’s (1955) analysis in the appendix of his paper. The vital issue of the present paper is to show that aspiration levels generally imply strategic advantages in sequential decision situations under uncertainty, even in situations where uncertainty is very high or the available information is very low! It is also important to see that it is much easier to get good estimations about the location of the median than to arrive at good estimations or realistic assumptions about the shape of the whole probability distribution underlying the uncertain situation. Employing the median makes it possible to easily adapt the aspiration levels to learning and experience.⁷

7 But take care! This argument might be misunderstood if it is assumed that the median ought to be a better estimator of the central tendency of a distribution than the mean. Of course, this is not the case. It is well known in statistics that the variance of the sampling distribution of the mean is smaller than the respective variance of the median. Therefore, it makes sense to take the mean in cases or “games”, where such information is available.

Generalization on decision-making in n choices-tasks

So far we have discussed sequential decision-making with respect to two options. It now seems appropriate to ask about generalizations when more than two options are available. Such a problem can be grasped in the following way:

Suppose a boy is to have a date with his choice of one of n unseen and unknown girls, and suppose he wishes to choose the prettiest. The girls are presented for him to see one at a time in a random order, and he must choose or reject a girl when she appears. Once he chooses, he sees the rest, and he is disappointed if his date is not the prettiest. How can he maximize his problem of choosing the prettiest in the lot. (Gilbert and Mosteller 1966, p. 35).

In different articles this problem is presented as the “secretary problem”, the “dowry problem” or the “beauty contest”. The problem can be grasped more precisely in the following isomorphism:

An urn contains n tags, identical except that each has a different number printed on it. The tags are to be successively drawn at random without replacement (the $n!$ permutations are equally likely). Knowing the number of tags n , the player may choose just one of the tags, his object being to choose the one with the largest number. The player’s behavior is restricted because after each tag is drawn he must either choose it, thus ending the game, or permanently reject it. The problem is to find the strategy that maximizes the probability of winning and to evaluate that probability. To emphasize that the player has no access to prior information, we could specify that the player is told after each tag is drawn only the rank of its numbers drawn so far. (Gilbert and Mosteller 1966, p. 38).

The optimal strategy in this case consists in rejecting s tags and accepting the first tag after draw s where the rank of the number drawn so far is 1. The asymptotic ($n \rightarrow \infty$) optimal solution gives $s = n/e$ and $p(\text{win}) = 1/e$, where $e = 2.718$ (cf. Gilbert and Mosteller 1966, p. 36 or Bruss 2004).

Our problem could also be generalized slightly differently, by assuming that people get the number printed on the tags instead of the present rank of their number they have drawn. In this case, learning about the underlying distribution of the numbers is possible. In the extreme case, the underlying distribution is already known exactly *a priori*. This case (besides many others) is also discussed in Gilbert and Mosteller (1966).

One by one, a sample of n measurements is drawn from a population with continuous cumulative distribution F . The continuity assures that ties have probability zero. After each draw, the player, who knows F and n , is informed of its value, whereupon he must decide whether or not to choose that draw. He is to maximize the probability of choosing the draw with the largest measurement in the sample. (Gilbert and Mosteller 1966, p. 51).

In case of $n=2$, the analogy to the problem we discussed before is obvious: the optimal value for the selection is the value that separates the distribution in 0.50 and 0.50 (fifty-fifty split) and yields a probability to win of 0.75. In case of $n = 3$, Gilbert and Mosteller’s analysis yields the value that separates the distribution in 0.69 and 0.31 as the first threshold for the first draw; the second threshold is the same as before, yielding a probability to win of $p = 0.68$. The asymptotic

analysis with $n \rightarrow \infty$ yields a probability to win of $p = 0.58$. (cf. tables 7 and 8 in Gilbert and Mosteller 1966).

So what is new in the argument presented in the text at hand? It transcends the existing analyses on sequential decision-making, because the earlier analyses have not yet grasped the general advantage of thresholds. Gilbert and Mosteller separate their analyses into a non-parametric and a parametric part by assuming no knowledge of the underlying distribution at all or complete knowledge of the underlying distribution. In their non-parametric analysis, they have overlooked the fact that the probabilities calculated there are worst-case probabilities, resulting from making no use of thresholds at all in the secretary problem. This is why they find a probability of $p = \frac{1}{2}$ for the secretary problem with $n = 2$. However, this problem is isomorphic to the Brussian problem presented above, where we could see that the use of a threshold may raise this probability, and the use of the median (in the ideal case) would lead to a probability to win of $p = \frac{3}{4}$ (this result corresponds to Gilbert's and Mosteller's analysis when there is complete information of the underlying distribution and $n = 2$; cf. table 8 there). Only if the threshold used is completely out of the range of the underlying distribution is the probability of a right guess $p = \frac{1}{2}$.

Generally, Gilbert's and Mosteller's analyses always deduce optimal thresholds by determining percentiles of the distribution that have to be surpassed in order for a choice to be accepted. Thus, determining thresholds is inherent to the task of calculating optimal choices, regardless of what the choice task looks like in detail. Depending on the details, different optimal thresholds can be calculated.

This is exactly where critique about such optimal choices comes up. Peter Todd (1997) convincingly argues that different optimization standards may lead to very different results. Thus, for example, it makes a big difference whether the task is to find the best candidate in a series of n choices, or either the best or the second-best candidate. Even worse, Gilbert's and Mosteller's optimal solution would pick over 11% of choices of the lowest quartile of the underlying distribution. This is unacceptable from the point of view of risk-averse standards. However, this argument does not concern the issue of this paper, because – in contrast to Gilbert's and Mosteller's analyses – this paper does not presume any methodological optimization standards. This paper rather shows that there is a general – you might even say a universal – advantage in the use of aspiration levels, even when there is no previous knowledge of the underlying distribution, although the determination of optimal aspiration levels depends on the optimization standards. With the argument of the paper at hand, the question can be posed how organisms adapt their aspiration levels in dependence of learning processes. The knowledge of such adaptations would make it possible to draw conclusions about the underlying optimization standards. By eliciting these standards, interesting conclusions about the tasks in our evolutionary past can be drawn; to do so, however, the assumption must be made plausible that these mechanisms have been developed in order to solve problems of evolutionary fitness.

Such arguments, for example, are made in the context of heuristics, when Gigerenzer et al. (1999) introduce the notion of evolutionary rationality, or in Cosmides and Tooby (2002), when they argue about the mechanisms of cheater detection.

Following the argument of the text at hand, it seems that aspiration levels can be found, at least in heuristics that are useful for decision-making in sequential choice tasks. Gigerenzer's heuristics can be used in such tasks, as Peter Todd (2002) demonstrates in his analysis, where he makes use of a take The Next Best heuristic. This heuristic is a slight variation of the Take The Best heuristic, presented in Gigerenzer et al (1999). In the following it will be argued that indeed Gigerenzer's heuristics make implicit use of aspiration levels. Thus, these heuristics can ensure an advantage greater than mere guesses, even in cases of sequential decision-making with complete uncertainty. This advantage might also have been relevant in evolution.

Some implications for the research on heuristics

So far we have learned that – given an optimization standard – the use of aspiration levels may optimize choices in sequential decision tasks. This is true, although aspiration levels are used within approaches of bounded rationality! The bounded rationality approaches consist in simplifying complex problems. However, the heuristics used with respect to such simplified problems may optimize towards this standard. And, in contrast to the original problem, optimizing becomes simple.

In the following we consider an interesting, still unanswered question. In cognitive science, besides Simon's satisficing principle, many different heuristics are in use. Despite the fact that researchers in cognitive science did not yet become aware of these mathematical insights, nevertheless there might be an implicit use of such advantages, e.g. by making use of thresholds that are implemented in the architecture of heuristics and by using adaptation mechanisms that come close to the mathematical insights. This question is interesting both for models already used, like the ACT-R cognitive architecture model (cf. Anderson 1993) and for empirical experiments about what and how subjects learn. In the latter case, evolution might have taken place such that human learning of thresholds takes place in a way that corresponds to the mathematical insights

In the last section it will be shown that many well-known heuristics make implicit use of aspiration levels. The question is: Besides Simon's satisficing principle, which heuristics have aspiration levels or thresholds?

The Recognition Heuristic

Gigerenzer & Todd (1999) take up Herbert Simon's conception of a bounded rationality, connected with methodological consequences, and his critique of perfect rationality. Like Simon, they presume some limitations in human information processing and the contextual influence of the environment. As they note:

Simon's vision of bounded rationality has two interlocking components: the limitations of the human mind, and the structure of the environment in which the mind operates.
(Gigerenzer & Todd 1999, p. 12).

In connection with these arguments, they introduce the term "ecological rationality", bringing to bear the particular role of the environment.

We use the term "ecological rationality" to bring environmental structure back into bounded rationality. A heuristic is ecologically rational to the degree that it is adapted to the structure of an environment... Thus, simple heuristics and environmental structure can both work hand in hand to provide a realistic alternative to the ideal of optimization, whether unbounded or constrained.
(Gigerenzer & Todd 1999, p. 13).

In contrast to Herbert Simon, they also take decision situations into account in which different alternatives for action are simultaneously available. Their "fast frugal heuristics" may also be useful in these situations:

Fast and frugal heuristics employ a minimum of time, knowledge, and computation to make adaptive choices in real environments. They can be used to solve problems of sequential search through objects or opinions, as in satisficing. They can also be used to make choices between simultaneously available objects, where the search of information (in the form of cues, features, consequences, etc.) about the possible options must be limited, rather than the search for the options themselves.
(Gigerenzer & Todd 1999, p. 14).

The recognition heuristic is the simplest heuristic presented by Goldstein & Gigerenzer (1999). This heuristic works by the simple principle: "If one of two objects is recognized and the other is not, then infer that the recognized object has a higher value." (Goldstein und Gigerenzer 1999, p. 41). This heuristic uses a threshold⁸:

8 In practice the thresholds in this heuristic may not be adaptable. Nevertheless, especially from an evolutionary point of view, it makes sense to ask where an optimal threshold should lie.

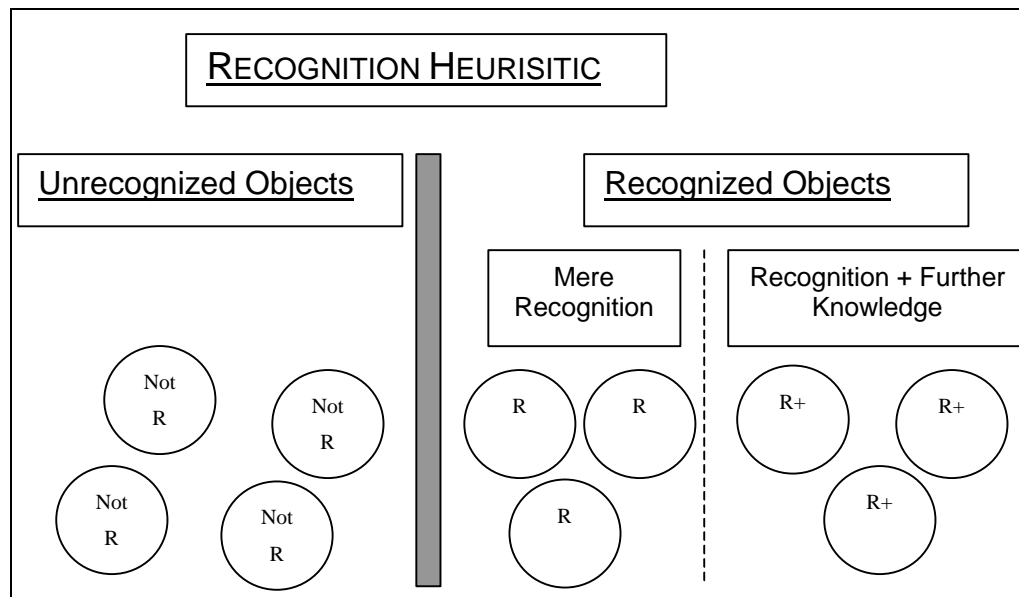


Figure 1: The Recognition Heuristic

The recognition heuristic makes sense if the recognition of an object is correlated with the adequacy of the chosen recognized object. Increasing this correlation increases the usefulness of this heuristic for questions like “Which U.S. city has more inhabitants: San Diego or San Antonio?” In one study, 100% of German students chose the correct answer, but only 62% of the students from the University of Chicago. On the other hand, in answering whether Dortmund or Munich has more inhabitants, Americans were much more apt than Germans to choose the right answer (cf. Goldstein und Gigerenzer 1999). This is because, with respect to German cities, Americans can use the recognition heuristic (taken for granted, they know of Munich but not of Dortmund); and vice versa, with respect to American cities. There Germans can use it (taken for granted they know San Diego but not San Antonio):

Both studies illustrate the ecological rationality of the recognition heuristic. The recognition heuristic is ecologically rational in the sense that it exploits the structure of information in natural environments: Lack of recognition in these environments is systematic and not random. Ignorance is beneficial if it is correlated with what one wishes to infer. The heuristic is not a general purpose strategy because this correlation holds in some situations, but not in all. (Goldstein und Gigerenzer 1999, p. 44).

In the following we will see that the use of the recognition heuristic comprises an indirect choice of aspiration levels. This choice is indirect in the sense that the median is a value like the number of inhabitants, which separates the set of all cities in two subsets of equal size, one subset containing all the cities below the median (that is, the small cities), and the other subset containing all the cities above it (that is, the big cities). However, the optimal aspiration level in the recognition heuristic is not determined by the median itself (that is, a size of a city), but rather by the optimal size of the recognized objects n . If recognition of a city correlates perfectly with the size of a city (in other words: if only the biggest cities are recognized), then the optimal value for the size of the set of recognized objects is $n=N/2$, which corresponds to the separation of a sample in

two sub-samples by the median split. In order to demonstrate this, we will go into some formal details of Goldstein and Gigerenzer's analysis.

In connection with their presentation of the recognition heuristic, the authors give a formula which calculates the proportion of correct answers when the recognition heuristic is applied in a two-alternative choice task. They presume a reference class consisting of N objects. A single test (question) consists of pairs of randomly chosen objects. There are three different ways the pairs of objects can be pieced together: both objects can be recognized; both objects can fail to be recognized; and only one of the two objects can be recognized. The N objects are compounded from n recognized objects and $N-n$ unrecognized objects. In all the following different pairs of objects are possible: $(N^2-N)/2 = N(N-1)/2$.⁹ From this total number, there are $n(N-n)$ pairs with one recognized and one unrecognized object, $n(n-1)/2$ pairs where both objects are recognized, and $(N-n)(N-n-1)/2$ pairs where neither object is recognized. In calculating the proportion of possible pair matches in each of the three cases given above, the respective calculated number of possibilities has to be divided by the total number of pair matches. For instance, the proportion of matches with two recognized objects is $n(n-1)/(N(N-1)) = (n/N)((n-1)/(N-1))$.

The correlation between the recognition of an object and the actual criterion is grasped by recognition validity α and knowledge validity β . α provides the probability of getting a correct answer when one object is recognized and the other is not; β provides the probability of getting a correct answer when both objects are recognized. In those cases in which both objects are unknown, a guess must be made; and the probability of getting a correct answer is $1/2$. Thus the proportion of correct inferences is (cf. Goldstein and Gigerenzer 1999, p. 44):

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

In the extreme example where all objects are unknown, this formula yields a 50% proportion of right answers, independent from α and β . Another question arises in extreme cases: namely, what is the ideal size of n for maximizing the proportion of right answers if the knowledge validity and the recognition validity are optimal – that is, if they have a value of 1? If the knowledge validity $\beta = 1$, then ideally all objects should be known with $n = N$. In this case the proportion of right answers $f(n)$ is equal to 1. Now we can ask about the ideal size of n if the knowledge validity $\beta = 1/2$ and the recognition validity of α is optimal, that is $\alpha = 1$. Please remember that the knowledge validity β is defined as the conditional probability of choosing the right object if both objects are known, for example, when the task is to choose whether Rome or Madrid has more inhabitants. Setting $\beta = 1/2$ implies a pure guess, with a probability of $1/2$ of choosing the right answer if both cities in the example are known. Setting $\alpha = 1$ means that the recognition heuris-

9 These are complete pair-by-pair comparisons. This is why this is equivalent to $\binom{N}{2}$.

tic always yields the right answer if one object is known and the other object is unknown.¹⁰ With respect to our example, this means that all known cities are also the cities with more inhabitants than all unknown cities. That is, $\alpha = 1$ means that if I compare an arbitrary known city with an arbitrary unknown city, the known city will always have more inhabitants than the unknown city. Now the question is: What is the optimal size of the known objects n , making it possible to maximize the proportion of right answers in this case? It can be calculated that the optimal size of n in this case corresponds to the reference value for the determination of the median¹¹ with $n = N/2$, since the median gives the value of a variable that separates the sample into two equal parts with $n_1 = n_2 = N/2$ when all individuals are arranged in order. This provides an interesting link to the theoretical consideration discussed above about the theoretic derivations from Goldstein and Gigerenzer (1999). From our considerations, we can conclude that the condition for the recognition heuristics is ideal if the sample is divided up into two equal parts so that one part contains the low values of the sample, and the other part, the high values. If this ideal condition holds, the proportion of right answers reached by using the recognition heuristic raises from 50% (mere guessing) to nearly 75% if the pairs of cities that have to be compared with each other are randomly drawn.

To sum up, it can be demonstrated that the use of the recognition heuristic comprises an indirect choice of aspiration levels. This choice is indirect in the sense that the median is a value like the number of inhabitants, which separates the set of all cities in two subsets of equal size, one subset containing all the cities below the median (that is the small cities) and the other subset containing all the cities above it (that is the big cities). However, the optimal aspiration level in the case of the recognition heuristic is not determined by the median itself (that is, a size of a city), but rather by the optimal size of the recognized objects n . But if $\alpha = 1$, the optimal value for the size of the set of recognized objects is $n = N/2$, which corresponds to the separation of the sample in two sub-samples by the median split. This is why it makes sense to say that the use of the recognition heuristic, when assuming an optimal recognition validity ($\alpha = 1$) and a nonexistent knowledge validity ($\beta = 1/2$), is optimal if and only if the aspiration level corresponds to the median. Because the result of the optimization is not the median itself, but the consequence of the application of the median in doing a median split of the total sample, the median as an optimal threshold has been determined “indirectly”.

Let us now consider the heuristic Take The Best, which is also presented in Gigerenzer, Todd & ABC-Group (1999), and where the use of aspiration levels is also an essential functional feature.

10 Therefore, it makes sense to assume that $\beta \in [1/2, 1]$ and $a \in [0, 1]$. If $a < 1/2$, then recognizing an object should lead to a decision against that object, whereas in the case of $a > 1/2$, recognizing one should lead to a decision in favor of it.

11 The reason for a marginal deviation from $N/2$ is that in total pair matching, the matches between an object with itself fall out so that, given N objects, instead of N^2 , only $N(N-1)$ pair-matches take place.

Take the Best

Let us assume that we know absolutely nothing about the cities, but recognize all of them. In this case, recognition is worthless and the knowledge validity about the size of the cities is $\frac{1}{2}$. Nevertheless, a somehow more complicated heuristics may be useful. Such a heuristic is Take The Best. What we will see in the following is that, in this heuristic, aspiration levels play an essential functional role as well.

In principle, this heuristic is based on the recognition heuristic, and like the recognition heuristic it can be used to make a choice between two objects. Making the assumption given before, i.e. that both objects are recognized, the Take The Best heuristic can be seen as a refinement of the recognition heuristic. This is especially valuable if the knowledge validity β is near $\frac{1}{2}$, which means that we have no knowledge about the size of the recognized cities. In this case, the question “Which of these cities has more inhabitants: Rome or Madrid?” can not be answered by using the recognition heuristic.

However, if we have some knowledge about the cities besides their size, and if some of this knowledge is correlated with the size of the cities, the Take The Best heuristic can be applied. Take The Best gives a fixed sequence of cues to be checked. These cues are correlated with the size of the cities. In Take The Best, the sequence of cues is ordered with respect to their validity, so that the cue with the highest correlation to city size is checked first; the cue with the second-highest correlation is checked next and so on. In contrast to this optimal check-order, the minimalist heuristic checks that the cues are in a random sequence. The principal method behind all these heuristics consists in checking all these cues until one is found that is known for both cities and that is only fulfilled in one of the two objects or that has a higher cue value in one of the two cities:

For instance, one of the heuristics that we propose, called *Take The Best*, decides between two available options by searching through cues in order of their validity, stopping when the first cue is found that distinguishes the options, and selecting the option indicated by the higher cue value. (Todd, Fiddick und Krauss 2000, p. 375).

Thereby two significant properties tag each cue: its ecological validity and its discrimination rate.

The ecological validity of a cue is the relative frequency with which the cue correctly predicts the criterion, defined with respect to the reference class... The discrimination rate of a cue is the relative frequency with which a cue discriminates between pairs of objects from the reference class. (Gigerenzer and Goldstein 1999, p. 84f.).

The ecological validity can be calculated by the following formula:

$v_i = (\text{number of correct predictions})/(\text{number of predictions, where the cue is used})$ (cf. Gigerenzer & Goldstein 1999, p. 80). The authors don't compare the applications of the recognition heuristic and the Take The Best heuristic from a formal point of view. Nevertheless, the latter is an alternative to the former when both objects are recognized. If the recognition heuristic is

applied in this case and the choice of an object is based on the knowledge about the objects, the parameter β gives the probability of right decisions or hits. It is clear that Take The Best outperforms the recognition heuristic if a cue can be found which has a higher ecological validity v_i than the knowledge validity β of the recognition heuristic.¹²

In our discussion it is important to keep in mind that Take The Best can be grasped as a generalization of the recognition heuristic. To recognize an object in the sense of the recognition heuristic can be understood as a cue (“recognized”, “unrecognized”) in the sense of the Take The Best heuristic. In those cases in which both objects are recognized, the recognition cue can be supplemented by further cues. Thus the formal treatment of each cue in the Take The Best heuristic is an analogue to the formal treatment of the recognition heuristic. Therefore in the case of the Take The Best heuristic we can also speak of an indirect use of aspiration levels: their utility is crucially dependent on the correlation between the cues and the criterion.

Summary

There is a strategic advantage in the use of aspiration levels for many heuristics and even cognitive models that have been developed in the context of research on bounded rationality. Thus far, this strategic advantage in sequential decisions under uncertainty has been ignored. Integrating these theoretic considerations could help to refine the corresponding adaptation mechanisms. For instance, an aspiration level could be arbitrarily set at the beginning and then, with new information, be modified towards the estimated median. Given the additional information, decision-making under uncertainty could be enriched by a complete expectation calculus, which is only necessary if the underlying probability distribution is skewed. Consequently, the considerations offered here open up new approaches for dealing with decisions under uncertainty on the basis of previous experiences.

Besides being of use for the models in the context of bounded rationality, these considerations of the advantages of aspiration levels may also be productive in other domains where aspiration levels are used, for example, in the context of models on attitudes. A concrete example is the Heuristic Systematic Model (HSM) of persuasion (cf. Böhner 2001). In accord with this model it is postulated that people strive for sufficient consistency in their attitudes. Thereby “sufficient” is operationalized by two constructs: the sufficiency threshold (ST), which gives the desired consistency, and the actual confidence (AC). As long as the AC value lies below the ST value, the respective person is motivated to look for more information. This model of attitudes has a clear link to Simon’s (1955) satisficing principle. It goes beyond Simon (1955) in maintaining that the sufficiency threshold depends on situational realities (for instance, whether the situation is im-

12 Besides the ecological validity, the value of cues depends on their discrimination rate. That is the proportion of pair-matching between the objects where the cue is applicable to the total number of possible pair matches between two objects from the reference class. Because this consideration has no relevance for the introduction of the argument presented here, a formal presentation of the discrimination rate has been omitted. Such a formal presentation can be found in (Gigerenzer und Goldstein 1999, p. 85).

portant or not) as well as on variables of the individuals involved (like the subjective responsibility felt by the person – cf. Bohner 2001, p. 261). Nevertheless, it uses aspiration levels, which can be modified in reference to external and internal changes. The considerations here show how the modifications of sufficiency thresholds could be optimized with reference to external changes.

To summarize the paper: Even within the framework of models of bounded rationality, the question of optimal adaptation makes sense. However, it does so only if it is simple to perform. Otherwise, it is inefficient to give up the assumptions of rational. The use of aspiration levels in sequential decisions under uncertainty seems to allow such simple optimization procedures by just trying to take the median as the aspiration level. This simple rule leads to (strategic)¹³ advantages in decision-making processes.

13 Strategic advantages come into play from an evolutionary point of view.

References

- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, New Jersey: Lawrence Erlbaum.
- Barret, C. H., & Fiddick, L. (2000). Evolution and Risky Decisions. *Trends in Cognitive Science*, 4(7), 251-252.
- Bohner, G. (2001). Attitudes. In M. Hewstone & W. Stroebe (Eds.), *Introduction to Social Psychology* (pp. 239-282). Malden, Mass.: Blackwell.
- Brockhaus. (1996). *Brockhaus - die Enzyklopädie* (20 ed.): Brockhaus-Lexika.
- Bruss, F. T. (2000). Mathematische Unterhaltungen: Der Ungewissheit ein Schnippchen schlagen. *Spektrum der Wissenschaft*, 6, 106.
- Bruss, F. T. (2004). Mathematische Unterhaltungen: Strategien der besten Wahl. *Spektrum der Wissenschaft*, 5, 102-104.
- Buchanan, B., & Shortliffe, E. (Eds.). (1984). *Rule based Expert Systems*. Reading, MA: Addison-Wesley.
- Cosmides, L., & Tooby, J. (2002). Knowing thyself: The evolutionary psychology of moral reasoning and moral sentiments. In E. R. Freeman & P. H. Werhane (Eds.), *Business, Science, and Ethics* (Vol. 4, pp. 91-127). Charlottesville, Virginia: Philosophy Documentation Center.
- Gigerenzer, G., & Goldstein, D. G. (1999). Betting on One Good Reason – The Take The Best Heuristic. In G. Gigerenzer & P. M. Todd & A. R. Group (Eds.), *Simple Heuristics That Make Us Smart*. Oxford: Oxford University Press.
- Gigerenzer, G., & Todd, P. M. (1999). Fast and Frugal Heuristics – The Adaptive Toolbox. In G. Gigerenzer & P. M. Todd & A. R. Group (Eds.), *Simple Heuristics That Make Us Smart* (pp. 3-34). Oxford: Oxford University Press.
- Gigerenzer, G., Todd, P. M. and the ABC Group (Eds.). (1999). *Simple Heuristics That Make Us Smart*. Oxford: Oxford University Press.
- Goldstein, D. G., & Gigerenzer, G. (1999). The Recognition Heuristic – How Ignorance Makes Us Smart. In G. Gigerenzer & P. M. Todd & A. R. Group (Eds.), *Simple Heuristics That Make Us Smart* (pp. 37-72). Oxford: Oxford University Press.
- Jamikovsk, P. (1999, Dezember 9). *Proseminar Künstliche Intelligenz – Heuristische Suche*, [URL]. Available: <http://www.uni-karlsruhe.de/~ueh5/> [2001, June 7].
- Johnsonbaugh, R. (2001). *Discrete Mathematics*. Upper Saddle River, NJ: Prentice Hall.

- Johnsonbaugh, R. (2001). *Discrete Mathematics*. Available:
<http://www.depaul.edu/~rjohnson/dm5th/> [2001, August 15].
- Newell, A., Shaw, J. C., & Simon, H. A. (1963). Chess-playing programs and the problem of complexity. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thoughts* (pp. 39-70). Berkely: McGraw Hill.
- Rubinstein, A. (1998). *Modeling Bounded Rationality*. Cambridge, MA: MIT Press.
- Sedlmeier, P. and Betsch, T. (Eds.). (2002). *Frequency Processing and Cognition*. Oxford: Oxford University Press.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics* (69), 99-118.
- Simon, H. A. (1956). Rational Choice and the Structure of Environments. *Psychological Review* (63), 129-138.
- Todd, P. M., Fiddick, L., & Krauss, S. (2000). Ecological Rationality and its Contents. *Thinking and Reasoning*, 6(4), 375-384.
- Todd, P. M. (1997). Searching the next best mate. In R. Conte, R. Hegselmann P. Terna (Eds.), *Simulating social phenomena*. (pp. 419-436). Berlin: Springer. Available:
www-abc.mpib-berlin.mpg.de/users/ptodd/publications/todd97/todd97.pdf