This paper is a preprint of an article published in Journal of Bioeconomics, Vol. 3, Issue 2–3 (2001), p. 195–215, therefore there may be minor differences between the two versions. The published version is available at *http://www.kluweronline.com/*. Access to the article may require a contract or license with the publisher.

Making good decisions with minimal information: Simultaneous and sequential choice

THOMAS DUDEY
Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Berlin, Germany (dudey@mpib-berlin.mpg.de)

PETER M. TODD
Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Berlin, Germany (ptodd@mpib-berlin.mpg.de)

Synopsis: The adaptive pressures facing humans and other animals to make decisions quickly can be met both by increasing internal information-processing speed and by minimizing the amount of information to be used. Here we focus on the latter effect and ask how, and how well, agents can make good decisions with a minimal amount of information, using two specific tasks as examples. When a choice must be made between simultaneously-available options, minimal information in the form of binary recognition (whether or not each item is recognized) can be used in the *recognition heuristic* to choose effectively. When options are encountered sequentially one at a time, minimal information as to whether or not each option is the best encountered so far is sufficient to guide agents using a simple search-cutoff rule to high performance along several choice criteria. Both of these examples have important economic as well as biological applications, and show the power of simple fast and frugal heuristics to produce good decisions with little information.

Key words: individual decision making; ecological rationality; information search; recognition heuristic; sequential search; secretary problem; dowry problem; optimal stopping; cutoff rule; heuristic; search; simulation

JEL Classification: D81 (Criteria for Decision-Making under Risk and Uncertainty), D83 (Search, Learning, and Information), C60 (Mathematical Methods and Programming, General)

1 Introduction — Meeting the need for speed

Humans and other animals frequently must make decisions in as rapid a manner as possible. To hesitate is often to be lost, whether this means losing an opportunity for a meal or a mate to a competitor, or losing one's life or limb to a predator or otherwise hostile environment. Organisms seeking to make choices and take action as quickly as possible can either speed up the input side of the decision process, reducing the time needed to gather information on which to base those choices, or speed up the processing/output side, reducing delays in processing the information and converting it into behavior. While evolution has developed faster processing mechanisms over the eons (e.g., myelinization to increase nerve impulse travel speed), the greatest time advantage is likely to come from reducing the amount of information sought before making a decision (Todd, 2000). But then what kind of adaptive decisions can an organism make, and how useful can they be, with a minimal amount of information?

In this paper, we present answers to these questions for two common but distinct types of decisions: choosing between two simultaneously-available options, and choosing one option from a string of sequentially-encountered possibilities. Both of these types of decisions are not only important for many adaptive problems faced by a wide range of species, but are also frequently seen in more modern human domains studied by psychologists and economists. This means that the minimal-information decision mechanisms for which we can find evidence in non-human animals could also underlie aspects of human economic behavior.

We take here an evolutionarily inspired view of how decision-making mechanisms should work and what they should achieve, leading to a focus on the need to minimize (or reduce as much as possible) the amount of information that must be gathered and used in reaching an inference, the importance of simplicity and speed in information processing, and the vital role of the environment in providing structured information to enable rapid and accurate domain-specific decisions. This view stands in sharp contrast to the vision of unbounded rationality, often defined by its adherence to the laws of logic and probability theory, that forms the basis of proper economic behavior as embodied in Homo economicus. The lofty goals and strict standards of unbounded rationality require the use of complex decision-making machinery to process all available information, without regard for costs or limitations in time, processing power, or knowledge. But in recent years an increasing body of (experimental) evidence has shown that real humans often violate the principles of rational behavior. This comes as no surprise from an evolutionary perspective, which holds that evolution did not shape the mind to be rational *per se*, but rather to be well adapted to its environment.

Consider the typical economic task of choosing a particular brand of some product. A standard (rational) economic approach to solve this problem would be to find out what all the available brands are, search for all the relevant information about each brand, weigh the importance of each piece of information according to one's utility function, combine the weighted information to come up with a final judgment for each brand, and choose the highest-valued one. From an evolutionary perspective instead, we would consider the way that our minds were designed to deal with information about multiple options (say, artifacts produced in different areas). This implies considering the structure of the available information itself. For instance, in such situations information was probably incomplete (not all the available options were always known) and largely socially obtained (paying attention to the greater combined experiences of others), with the further implication that useful artifacts would probably have been heard about more often than useless ones (because people would continue to use the good ones and hence talk about them longer). Given this environment structure, we can postulate that people could make good choices by simply selecting those items they had heard about from others, rather than by personally paying the cost of investigating all the available possibilities as in the standard economic approach described above. This leads to two testable predictions: first, that people behave in this way (in particular settings), and second, that they make good choices as a consequence. Exactly this has been proposed in the form of the recognition heuristic, which we discuss in the next section.[1] In a similar manner, we can apply the investigatory principles of

[1] Similar arguments underlie the investigation of a related mechanism, conformist transmission, in which people tend to adopt the most common behavior exhibited in a population (Boyd & Richerson, 1985; Henrich & Boyd, 1998); however, this mechanism relies on more information, namely counts of behaviors, than just the minimal presence or absence of recognition explored here.

evolutionary psychology to propose and explore psychologically plausible mechanisms for making decisions in other (modern) economic contexts.

To reach decisions and make choices in a reasonable amount of time, real agents must employ limited search for information about available options (the first decision case we consider here) or for options to choose between (the second case we consider), whereas models of unbounded rationality assume that search can go on indefinitely. In realistic models, search must be limited because real decision makers have only a finite amount of time, knowledge, attention, or money to spend on a particular decision. Limited search requires a way to decide when to stop looking for information, that is, a stopping rule. One way to stop information search is to use a stopping rule that optimizes search with respect to the time, computation, money, and other resources being spent. More specifically, this 'optimization under constraints' vision of rationality holds that the decision maker should calculate the benefits and costs of searching for each further piece of information and stop search as soon as the costs outweigh the benefits (e.g., Sargent, 1993; Stigler, 1961). This approach sounds plausible at first glance. But a closer look reveals that optimization under constraints can require even more knowledge and computation than unbounded rationality, as all of the costs and benefits of all of the possible courses of searching or not searching for each further piece of information or option must be calculated (Vriend, 1996).

Instead, search for information or options can be limited in a much simpler way, following the tenets of bounded rationality. Herbert Simon proposed bounded rationality as consisting of two interlocking components: the limitations of the (human) mind, and the structure of the environments in which the mind operates. The first component implies that humans 'must use approximate methods to handle most tasks' (Simon, 1990, p. 6). These methods include recognition processes that largely obviate the need for further information search, heuristics that guide search and determine when it should end, and simple decision rules that make use of the information found. Simon's second component is also of crucial importance because it can explain when and why simple decision mechanisms perform well: if the structure of the mechanism is adapted to the structure of the information in the environment. This leads to a new conception that proper reasoning must be not only boundedly but also ecologically rational, arising from constrained decision mechanisms that are matched (that is, adapted) to the particular structure of information in the environments in which they are applied.

Evidence is growing that humans do indeed make decisions in an ecologically rational manner, using as little information as possible and tailoring their information and option search to the structure available in the environment. Experts have been shown to base their judgments on surprisingly few pieces of information (Shanteau, 1992). It has been found that people can trade off the effort involved in making a choice against the accuracy of that choice, and choose a quick and simple decision strategy that would achieve the desired balance (Payne, Bettman, & Johnson, 1993). And simple heuristics that use only a single piece of information to make a choice between two alternatives have been discovered to rival the performance of much more complex and information-hungry methods such as multiple linear regression (Gigerenzer & Goldstein, 1996).

Thus, researchers interested in how people and other animals make many real-world decisions in a rapid manner would do well to begin by looking at candidate 'fast and frugal' decision mechanisms that could accomplish the job with a minimum of information and processing. In the rest of this paper, we do just that for two simple cases: First, in the next section we briefly review work on the recognition heuristic, which uses perhaps the least possible amount of information — whether or not an option is recognized — to choose between simultaneously available options. Then in the third section, we consider at somewhat greater length the problem of choosing one option from a sequence of possibilities seen one after another. In this decision situation, each possibility must either be selected or rejected forevermore on the spot — there is no going back to previously-seen options. This problem, commonly called the secretary problem, can be effectively solved by simple mechanisms that use only the information of whether or not a current option is the best seen so far. In both the cases we discuss here, the way information is structured in the environment allows simple mechanisms to make good decisions despite ignoring most of the available information, demonstrating how ecological rationality can work.

2 Choosing between options using recognition alone

One of the simplest kinds of choice — numerically, at least — is to select one option from two possibilities, according to some criterion on which the two can be compared. What is the minimal amount of information that an organism could use to make a reasonable choice in this case? Of course, the organism could use no information at all, and just choose randomly — certainly this would be simple, and probably fast, but it would result in a good choice (i.e., selecting the better option in terms of the criterion dimension) only half of the time on average. What is the least increase in information that could allow an improvement over the random case?

Consider the situation in which the only information available is whether or not each option has ever been encountered by the organism before. Then either this information is uncorrelated with the criterion, in which case random choice remains the only option, or else recognition has some (positive or negative) correlation with the criterion. In this latter case, the decision maker can do no better than to rely on his or her own partial ignorance, choosing recognized options over unrecognized ones (if the correlation is positive, or the reverse if it is negative). This kind of 'ignorance-based reasoning' is embodied in the *recognition heuristic* (Goldstein & Gigerenzer, 1999): When choosing between two objects (according to some criterion), if one is recognized and the other is not, then select the recognized one. For instance, if deciding at breakfast-time between a cheddar-cheese omelette and a durian-fruit compote (on the criterion of being good to eat), this heuristic would lead (most Western) people to choose the recognized egg dish over the unrecognized fruit offering.

As just indicated, the recognition heuristic will not work in every environment, but in some it will be ecologically rational — that is, able to exploit environment structure to yield good choices more often than would random choice. This will happen in all those decision environments in which exposure to different possibilities is positively correlated with their ranking along the decision criterion being used. (Again, negative correlations can also be exploited, by reversing the decision of the recognition heuristic.) To continue with our breakfast example, using the recognition heuristic to decide what to eat can be ecologically rational because those things that we do not recognize in our environment are more often than not inedible — humans have done a reasonable job of discovering and incorporating edible substances into our diet. (Those who know durian fruit may consider its edibility debatable.) Norway rats follow a similar rule, preferring to eat things they recognize through past experience with other rats (e.g., items they have smelled on the breath of others) over novel items (Galef, 1987).

Goldstein & Gigerenzer (1999) have used a different kind of example to amass experimental evidence that people also use the recognition heuristic: Because we hear about large cities more often than small cities, using recognition to decide which of two cities is larger will often yield the correct answer (in those cases where one city is recognized and the other is not). Employing the recognition heuristic can lead to the surprising less-is-more effect, in which less knowledge can lead to more accurate decisions than does a greater amount of knowledge. More specifically, an intermediate amount of (recognition) knowledge about a set of objects can yield the highest proportion of correct answers, because it allows the recognition heuristic to be used more frequently (there are more pairs where one item is recognized and the other is not) — knowing (i.e., recognizing) more than this can actually decrease the decision-making performance (Goldstein & Gigerenzer, 1999).

The recognition heuristic can be extended beyond picking one option from two, generalizing to cases in which several options are to be chosen from a larger set of possibilities. For instance, recognition can be called on when several social partners are to be chosen for some collaborative activity such as resource exchange or hunting. Can the recognition heuristic also be put to work in a modern-day equivalent of this type of choice, choosing companies for investment? When deciding which companies to invest in from among those trading in a particular stock market, the recognition heuristic would lead us to choose just those that we have heard of before. Such a choice can be profitable assuming that more-often-recognized companies will typically have some of the better-performing stocks on the market — a testable, but not obvious, assumption. For instance, people can recognize a company because of news coverage of spectacularly poor performance as well. So testing this assumption also entails answering whether recognition is more common for companies through positive or negative associations.

This assumption has indeed been tested recently as a fast and frugal approach to investing (Borges, Goldstein, Ortmann, & Gigerenzer, 1999). Several sets of people were asked what companies they recognized (but not whether positively or negatively), and investment portfolios were formed based on the most frequently recognized firms. In this (short-term) trial of a minimal-information heuristic in a complex and unforgiving real environment, the researchers found that ignorance-driven recognition-based investment choices could beat the performance of professional stock pickers and mutual funds. This study does not, by itself, prove that people use the recognition heuristic when making such choices (though common investment advice suggests this is so) — it only provides evidence that using this heuristic can be a surprisingly adaptive strategy in the stock market environment. Experimental examination of whether (and if so when) people employ the recognition heuristic in this and other domains remains an important challenge.

3 Searching through sequential options

The recognition heuristic can help agents make good choices with a minimum of information when confronted with multiple options to choose between simultaneously. But many of the important decisions we face are not structured this way — rather, we often must choose between a set of options that we do not see all at once, but one after another, sequentially. This happens for instance when we are trying to find the best price on tomatoes as we drive from store to store, or when we are looking for a parking space outside each store. More importantly, we often encounter potential jobs one by one, or meet potential mates in a sequence, or find potential dwellings spread out over time. These situations are typically characterized by low (or zero) probability of being able to recall, or return to and choose, previously-seen options once they have been passed by (e.g., houses on the market one month will probably be sold by the next — for California, divide the time-scale by 100). The problem then becomes one of deciding when to stop searching and go with the currently-available option.

How can we make a good choice in such a sequential setting? The best way, assuming no costs for time spent searching, would be to find out what all the available options are, compute which one is best according to the choice criteria, and then wait until that option comes along. Of course, such full knowledge is rarely available. Instead we might know the distribution of possible values of the options we might encounter, and we might have a limited number of possible options that we can check, in which case we can compute the optimal point to stop search so as to balance the gain from the particular chosen option against the risk of missing a better option later on (DeGroot, 1970). But what is the *minimal* amount of information that could be used for these sequential choices? This is a difficult (perhaps ill-defined) question to answer, but one possible approach is the case in which only a single bit's worth of information is registered for each option, corresponding to whether (bit = 1) or not (0) the current option is the best one seen so far (i.e., if it is a 'candidate,' in the terminology used below). Then, each sequence of possible options would be conceived of as a binary string of 1's and 0's, and the decision maker would have to decide where to stop (usually on a 1).[2] This minimal-information formulation of the sequential search problem underlies a fair amount of earlier work in this area (though sometimes not explicitly). Given this framework, we now want to ask, how can a decision maker go about finding a good option, and how well will the decision maker be able to do?

As we will see, the answers all depend on what our definition of 'good' is. We start with perhaps the most widely studied version of this sequential search problem, in which 'good' is defined as having a high probability of finding and choosing the single best applicant in the available set of options. This version is known in probability theory as the *secretary (or dowry) problem*. In the secretary problem it is the searcher's aim to find the (one) best of N applicants for a job as a secretary, with the assumption that they can all be arranged on a single dimension of overall quality. The applicants are presented to the searcher sequentially in a random order and the searcher has no knowledge about the distribution of the applicants' quality values. With each new applicant, the searcher learns the quality (or current rank in the strictest version) of this applicant, and then must choose between stopping the search and thus hiring the current applicant or continuing the search to look for a better applicant. If the searcher continues she cannot go back and choose an earlier applicant — that is, there is no ability to 'recall' past applicants in this search.[3]

To maximize the chance of selecting the best applicant in the secretary problem setting, the searcher should sample the first 37% of the applicants and then select the first candidate thereafter who is better than all previous applicants. (See Ferguson, 1989, for a review of the literature on the secretary problem and this optimal solution.) One simple way to do this is to set an aspiration level equal to the quality of the best applicant seen so far, and then after 37% of the applicants have been seen, fix this aspiration level and use it to stop the search with the next applicant seen who exceeds that aspiration — an approach akin to Simon's (1990) notion of *satisficing*.

Thus, given that the searcher knows the optimal 37% sample size — which assumes that the searcher first knows the total number of applicants that could be encountered — following this *37%*

[2] Note that a fair amount of processing may have to go into determining the value of the *best-so-far?* cue for each option, comparing its criterion value to that of the previous best-so-far candidate (which must therefore be stored somehow); but the same is true of the recognition heuristic, which makes use of a simple binary *recognized?* cue that must itself be computed with further (potentially complex) machinery such as a face recognition mechanism. We are interested here in the simpler higher-order cognitive decision mechanisms that can be built upon the outputs of possibly more complex lower-level mechanisms including perception and categorization — see Todd (1999).

[3] These assumptions restrict the set of search situations that the secretary problem can reasonably model; others, e.g., Hey (1982), and Moon & Martin (1990), have explored decision rules that can be used in those situations where full recall is possible or further assumptions are relaxed, and heuristics for the case of mutual or two-sided mate choice are considered by still others, e.g., Todd & Miller (1999), and Bergstrom & Real (2000).

rule is simple. The searcher must only keep in memory the optimal sample size (37%) and the quality of the best applicant seen so far. With each new applicant seen, a simple pair of comparisons (about the applicant's quality, to yield the best-so-far? cue, and its position number) is sufficient to decide whether to stop or continue search. But a simple rule does not necessarily imply that the actual search is short. On the contrary, the 37% rule leads to long searches: The first 37% of applicants are *always* searched through, and 37% of the time the search will go to the very end of the applicant pool, because the best option was already observed during the sample phase. On average the searcher examines 74% of the applicants before making a final choice. And even more striking, the simple rule does not guarantee great success: The 37% rule fails to find the best applicant in the whole set 63% of the time (i.e., it only succeeds 37% of the time). But no other rule can do better — at least not on this criterion of success. (Later we will consider other, possibly more reasonable, criteria for measuring search performance.)

While applying the 37% rule is simple, its derivation is not. First, it is not immediately obvious that the general approach of sampling a certain number of applicants, setting an aspiration level based on that sample, and then picking the next encountered applicant who exceeds the aspiration level thereafter is indeed the optimal procedure. Second, the fact that for large $N \to \infty$ the optimal sample size (for setting the aspiration level) approaches $1/e \approx 0.368$ is also not intuitive. (But see Mosteller, 1965, for a clear explanation of why the 37% rule is the optimal strategy.) Obviously it is unrealistic to assume that humans faced with such a search problem actually derive this optimal rule and then follow it. At this point economists often argue that while humans do not actually carry out such complicated calculations they still act *as if* they were doing them. But this still leaves us with the question of how people actually do make these kinds of search decisions (such as in the secretary problem) and how well they do given a particular definition of success (such as the optimal approach described here). We now turn to a review of past research on people's performance on the secretary problem, and mechanisms that have been proposed to account for the observed patterns of behavior.

3.1 Experiments on the secretary problem

Several experimental investigations of the secretary problem have indicated that the 37% rule is not the best explanation of observed behavior — that is, people do not seem to be acting optimally in this setting. Kahan, Rapoport, & Jones (1967) conducted some of the earliest experiments in this area. In their studies, participants had to search for the largest of 200 different numbers that were presented one by one. They had no knowledge of the underlying distributions, and no ability to recall (return to) previously-seen numbers. Overall Kahan et al. found that up to half of all participants searched less than directed by the 37% rule. The researchers suggested that people could instead be using a subjective cut-off rule, which fitted the observed behavior rather well. This rule says to stop search on any number that is far enough (in terms of standard deviations) above the mean of the previous offers — but note that this requires knowing much more than just the binary best-so-far? cue (i.e., all the actual values), and thus goes beyond the strict definition of the secretary problem.

Corbin, Olson, & Abbondanza (1975) ran a study that showed that such a cutoff model does not adequately explain some observed search behavior. In particular, they wanted to show that the exact order in which applicants (or options) are seen can exert an effect on the stopping behavior of searchers — even though standard notions of rationality (and the 37% rule) indicate that order of encountered options should not matter for deciding when to stop. In their setup, participants had to find the highest of five numbers presented sequentially on cards, again with no recall possible, in each of 120 trials. Subjects were told that the numbers were drawn randomly, though this was not true. In half of the trials the second card was higher than the first, thus making it a ‘candidate’ (defined as a number that is the highest seen so far, and thus a possible stopping point when searching for the highest value). These 60 trials were prepared such that this candidate was low (around 50), medium (around 500) or high (around 4400) in 20 trials for each level (‘magnitude’ differentiation). Moreover, in each of these three groups the first card was either extremely low or extremely high relative to the second (‘history’ differentiation). In 40 of the 120 trials the third card was a ‘candidate’. These 40 trials were also arranged in groups concerning the ‘pattern’ (relation of first to second card), ‘magnitude’ (of third card) and ‘history’ (size of first two cards relative to the third). The remaining 20 trials used rank orders that were not under special examination but were left in the study to ensure that participants would not develop a bias for stopping at a certain position. (Again, the inclusion of magnitude and history information departs from the standard secretary problem, but relevant conclusions can still be drawn from the effects of candidate pattern.)

For trials in which the second card was a candidate, there was a clear tendency for its magnitude to increase the probability that search would be stopped at this point. These candidates were also significantly more often chosen after a low history (low first value seen) than after a high history. Similar strong effects were observed for magnitude, history and pattern when the third card was a candidate. Pattern was especially important: Whenever the second value was lower than the first and the third was higher than both (making a down-up pair of transitions, or a 1–0–1 pattern of best-so-far? cue values), participants almost always stopped on the third value. In comparison, when there was a monotonic increase from the first to the third card (a 1–1–1 pattern), the third card was picked only rarely, even though the 37% rule would direct searchers to stop at that point. Corbin et al. related their findings to the idea that people use general heuristics for solving this problem — for instance, the magnitude of the current observation (in absolute terms, or relative to the entire experiment, or relative to the trial’s previous observations) and the trend of the previous observations have a strong impact on the decision of the participants. What is not made clear is what those heuristics might be, and how particular patterns would exert influence on participants’ decisions to stop or continue a search.

The results from both Kahan et al. (1967) and Corbin et al. (1975) show that people do not typically follow the 37% rule. As we noted above, however, the original secretary problem assumes no knowledge of the distribution. While participants in the experiments just described did not know anything about the distribution when they started their search, they did see actual number values throughout their search and thus could learn at least something about the underlying distribution they were searching from, thus violating one of the secretary problem’s assumptions. This problem was avoided by Seale & Rapoport (1997), who were the first to experimentally investigate the original form of the secretary problem. In their secretary-hiring setting, only the *relative rank* of those applicants interviewed so far was revealed, rather than some real value.

Seale & Rapoport regarded the optimal 37% rule solution of the secretary problem as only a benchmark and put more emphasis on finding simple heuristics that would be a better explanation of the actual behavior of the participants in their studies. They proposed three such heuristics, namely a *cutoff rule*, a *candidate count rule* and a *successive non-candidate rule*. The *cutoff rule* is a generalization of the optimal solution, where searchers simply pass by a certain number $r - 1$ (not necessarily the optimal number) of applicants and then hire the next encountered top-ranked applicant (so the 37% rule is a cutoff rule with the cutoff set at 37% of the possible alternatives). As defined before, each applicant that is top-ranked at the moment of being interviewed (i.e., when the option is assessed) is termed a *candidate*. The *candidate count rule* then simply implies choosing the j^{th} candidate seen. The *successive non-candidate rule* on the other hand chooses the first candidate that is interviewed after observing at least k consecutive non-candidates — that is, it stops searching after the gap between successive candidates has grown sufficiently large. All of these heuristics demand only minimal cognitive requirements (mainly counting and computing the best-so-far? cue), and thus are simpler than Kahan et al.'s subjective cut-off rule that must calculate means and standard deviations; but they can account for only some of the order effects found by Corbin et al. (1975).

Seale & Rapoport (1997) first showed by means of Monte Carlo simulations that their proposed heuristics can be very effective at finding the best applicant, in comparison with the benchmark 37% rule.[4] With the best parameter choice, the cutoff rule is by definition identical to the optimal solution (i.e., when the cutoff is set to 37; all trials here recalculated with $N = 100$ applicants) and is thus successful in 37% of the search trials. But even with small to medium deviations of the cutoff parameter, e.g. cutoffs between 26 and 50, this rule still robustly finds the best candidate in more than 35% of the trials. The successive non-candidate rule also achieves a very high success rate of more than 35% by setting the threshold for gap-size between candidates to the best value of 19. In contrast, the candidate count rule is successful in only 21% of the trials with the best parameter choice (stop on the 5^{th} candidate). (These results can be seen in both Figure 1 and Table 1, described later.)

Next, Seale & Rapoport compared the predictions of the three search heuristics with the actual behavior of their participants when searching through sequences of 80 values (presented as relative ranks). The cutoff rule was most successful in this regard, being most consistent with observed search behavior (accounting for between 41% and 74% of the decisions made) for 21 out of 25 participants. The successive non-candidate rule gave the best fit for 8 of the 25 (equalling the performance of the cutoff rule for some participants), while the candidate count rule best accounted for only one participant's decisions.

Seale & Rapoport (2000) extended this work to the case where the number of applicants N is ex-ante unknown — instead, the participants only knew that the number of applicants was randomly chosen between 1 and some upper bound U. The optimal solution in this case is to follow a cutoff rule akin to the optimal solution of the standard case, looking at 13.5% of U and choosing the first candidate thereafter. The chance of finding the best applicant overall is then slightly above 27%. In this setting, most results are very similar to those for the original secretary problem. The strategies again perform very close to the optimal success rate (equalled by the cutoff rule), with the candidate count rule reaching 25% and the successive non-candidate rule 27%. Again the cutoff rule is best at explaining the observed behavior of the participants in the experiment.

One of the key results in both of these studies (similar to those of Kahan et al., 1967) is the observation of early stopping: A majority of the participants stopped earlier than prescribed by the optimal solution. As a consequence, the human searchers had success rates of 30–32% (compared to 37%) in the 1997-study and 21–25% (compared to 27%) in the 2000-study. Seale & Rapoport (1997) offer one possible explanation for this early stopping, arguing that participants would stop before the optimal point if they were including endogenous search costs in their stopping decisions. According to Seale & Rapoport's calculations, the modal observed behavior would be the optimal search length if participants calculated in a search cost at every step of 1% of the value of selecting the best applicant (or 3.5% in the modified unknown-N version). Note that this does not necessarily imply

[4] Seale & Rapoport report performance "efficiency" and "effectiveness" in terms of the ratio of a strategy's success rate to the success rate of the optimal rule, but here we just report success rates.

that people are embodying a fully rational approach by actually calculating with search costs — they could still be using a simple rule, mindless of costs, that leads to shorter searches and thereby produces behavior in line with the optimal search length incorporating costs. Another explanation for early stopping could be provided by the order effects found by Corbin et al., a possibility we are currently exploring.

The fact that people search too little in these experiments may also indicate that they are not understanding the problem in the same way that experimenters intend it, or that they are behaving in ways more appropriate for search situations that they typically encounter. So to better understand how living agents can perform sequential search with minimal information, we should think again about the search tasks that they would naturally face — how many of these adhere to the form of the secretary problem? Probably not too many. As mentioned earlier, there may be some possibility of recalling past alternatives, or some costs to search, or some uncertainty about the number of available alternatives, or some knowledge about the available distribution (as could be learned in the Kahan et al., 1967, and Corbin et al., 1975, studies). Such modifications to the search situation have been studied both analytically to find optimal solutions (e.g., Corbin, 1980, for search with varying possibilities of recall, Seale & Rapoport, 2000, for the case with an uncertain number of possible alternatives) and experimentally to see how people behave in these settings (e.g., Schotter & Braunstein, 1981). The first three factors have predictable effects — greater costs and greater uncertainty lead to shorter searches, and the ability to recall leads people to go stop when costs grow too large and go back to the best option previously seen. More interesting is the finding that more information has little effect on individuals' search behavior (Schotter & Braunstein, 1981), providing support for our investigation of minimal-information search strategies. What has been less explored is the effect of a fifth factor: The payoff function for the search itself. As we noted earlier, the definition of what is 'good' search performance can have an impact on what search strategies are appropriate. Thus, we should now ask what payoff functions, or searcher goals, are likely to be used? And how well can different minimal-information search strategies meet those goals?

3.2 Searching with broader goals

To repeat, how and how well a searcher can find good options using the minimal best-so-far? cue depends on the searcher's definition of a good option. In the secretary problem, the single best option was the only good one. But this strictest-possible payoff function would be found in few natural situations. In many species, most animals find some mate, some food source, and some place to live (and indeed most companies find some secretary), and thus receive some payoff, even if not the highest possible payoff. Thus in these cases, a payoff proportional to the quality of the alternative chosen (e.g., the typing speed of the hired secretary) is more appropriate than the all-or-none payoff of the standard secretary problem. For other species, available alternatives (e.g., mates or habitats) may be limited so that not everyone can succeed in making a viable choice — for instance, only a quarter of the possible habitats in which one could settle may provide enough resources to raise offspring. In such cases, the search payoff function could fully reward only choices made in the top 25% of all available alternatives and give zero payoff to all other choices.

Todd & Miller (1999) considered these questions by taking the original secretary problem (in the equivalent *dowry problem* form for mate choice) as a starting point and broadening the goal of the searcher to biologically more realistic settings. They used Monte Carlo simulations to examine how the cutoff rule (which they called *Take the Next Best*) performs in the usual secretary problem search setting, but with a variety of different payoff measures. These included payoff proportional to the mean value of chosen alternatives (across multiple searches), payoff only for finding an applicant in the top 10% or in the top 25% of the population, and payoff only for finding an applicant that is *not* in the bottom 25% of the population (i.e., avoiding the worst possible alternatives). For each of these measures, Todd & Miller assessed the length of search, in terms of cutoff point, needed to maximize the chance of meeting the particular goal (or maximizing the mean payoff value for the first goal).

In the first set of simulations the number of applicants was fixed at $N = 100$. As stated earlier, the goal of finding the single best applicant requires first checking 37% of the applicant population before setting the aspiration level, and it succeeds only 37% of the time. But when the goal is changed to maximizing the average value of the chosen applicant, much less search is needed to achieve the best possible performance: here Todd & Miller found that an aspiration-setting cutoff at 9% of the applicants led to the maximum average value of chosen applicants (equal to 92 out of 100 given applicant values evenly distributed from 1 to 100). For the other new goals, again little search was needed and the rate of success was much higher (see Table 1 entries for the cutoff rule for a summary of these results, which vary slightly from those originally reported by Todd & Miller owing to random fluctuations).

To test how these search times are affected by a change in the number of available applicants, Todd & Miller ran a second set of simulations with $N = 1000$. Whereas for the original secretary problem goal the optimal solution is still the same (cutoff = 37%), the best cutoff values for the other goals do not grow in a similarly linear (percent-based) manner. Instead, with this larger N only 1–3% of the population need be looked at to set an aspiration level for finding 'good' applicants (see Table 1). Based on these findings, Todd & Miller proposed an even simpler (cutoff-style) heuristic, applicable for a wide range of N, which they call 'Try a dozen': Take the first applicant that is better than the best of the first 12 applicants. This heuristic neither needs to assess the expected total number of possible alternatives N, nor compute a percentage of that number, to perform reasonably well on the variety of goals Todd & Miller considered. (This is in contrast to the optimal 13.5% cutoff rule for the standard secretary problem with an unknown number of applicants between 1 and N, which was carefully explored by Seale & Rapoport, 2000.)

How well do Seale & Rapoport's other two rules perform on these different goals? We extended their simulations to find out, using 100 applicants. We computed the average length of search required for different parameter values of all three heuristics (e.g., how long would one need to search on average with a candidate-count parameter of 5?), and plotted this against success of each heuristic on each goal. (In contrast, Seale & Rapoport looked at the number of correct selections versus parameter values; but because the three heuristics use different parameters, we chose to look at total search time instead to allow direct comparison between heuristics.) First, in Figure 1, we show success on the standard secretary problem goal of selecting the single best applicant. As Seale & Rapoport (1997) found, the cutoff and successive non-candidate rules are very close in performance — similar search lengths yield similar levels of success — while the candidate count rule lags behind. For the goal of picking an applicant in the top 10% (Figure 2), the striking result is that the already small difference in performance between the cutoff rule and the successive non-candidate rule decreased even further, with the cutoff rule succeeding in 81% of the trials while the successive non-candidate rule reached 80% success, after a mean search length of 43. Finally, when maximizing the mean value of the applicant selected (Figure 3), the pattern looks about the same, but with maximum success being reached with less search.

The performance comparison for all three rules on the five goals, for both $N = 100$ and $N = 1000$,

Table 1: Results of an extended Monte Carlo simulation, following Seale & Rapoport (1997) and Todd & Miller (1999), comparing the performance of three search heuristics on five goals (selecting the single best applicant, an applicant in the top 10% of the quality ranks, in the top 25%, avoiding the bottom 25%, or maximizing the mean quality value selected). For each rule and goal, entries are shown for the best parameter value found (i.e., yielding the highest success); the mean success of the rule on that goal given the best parameter value, in percent of time the goal was achieved or percent of quality range returned for the Max Mean goal; and mean total search length given the best parameter. All results are computed over 100,000 rounds.

	Goal				
	Top 1	Top 10%	Top 25%	Not Bottom 25%	Max Mean
			$N = 100$ Applicants		
Cutoff Rule					
Best Parameter Value	37	15	9	4	10
Mean Success in %	37	82	92	99	91
Mean Search Length	73	42	29	14	31
Successive Non-Candidate Rule					
Best Parameter Value	20	8	4	1	5
Mean Success in %	35	80	91	99	90
Mean Search Length	71	43	28	13	32
Candidate Count Rule					
Best Parameter Value	5	4	4	3	3
Mean Success in %	21	64	81	98	84
Mean Search Length	59	38	38	19	19
			$N = 1000$ Applicants		
Cutoff Rule					
Best Parameter Value	354[a]	33	15	5	33
Mean Success in %	37	97	99	100	97
Mean Search Length	722	143	74	27	143
Successive Non-Candidate Rule					
Best Parameter Value	194	19	9	2	18
Mean Success in %	35	96	98	100	97
Mean Search Length	695	150	85	29	144
Candidate Count Rule					
Best Parameter Value	7	5	4	3	4
Mean Success in %	16	83	92	99	92
Mean Search Length	526	210	100	36	100

[a]The cutoff rule with parameter 354 led to 36,719 correct choices in 100,000 simulation rounds, while parameter value 371, closer to the theoretical optimum, produced 36,718 correct choices — indicating the flat maximum in this problem — with a mean search length of 739.

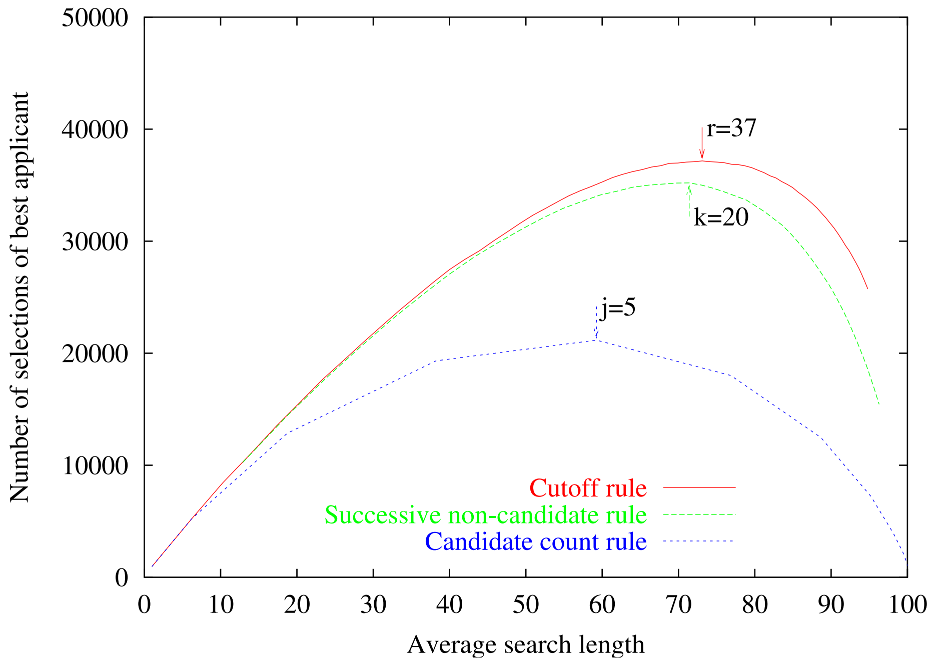

Figure 1: Performance of three search rules (see text) for the goal of selecting the single best applicant out of $N = 100$ applicants (simulation results computed over 100,000 rounds). For each value of the parameters r, k, and j, a point determined by the success of the corresponding rule with that parameter value and the mean search length using that rule with that parameter value is plotted, with all points for each rule connected by a line. The specific parameter value that yields the greatest number of correct selections is highlighted for each rule (see also Table 1).

is shown in Table 1. Here we can see, as Todd & Miller (1999) found, that a ten-fold increase in the number of available applicants results in only a doubling to tripling of parameter values and search length needed by the cutoff and successive non-candidate count rules to maximize performance on the new measures (slightly more for the mean value criterion). The candidate count rule still performs least well at $N = 1000$, though the gap has narrowed — but more surprisingly, it does this with little or no change in parameter values (e.g., choosing the fourth candidate works best for finding an applicant in the top 25% for both $N = 100$ and $N = 1000$, though the mean search length increases because the maximum search length is so much longer in the second case). Still, an agent with no knowledge of N (at least in the range 100 to 1000) would do better picking a low parameter value (e.g., 12) for one of the other two rules, which outperform the candidate count rule over a wide parameter range on all of the criteria.

Another real-world feature that we should consider in searching for ecologically rational search rules is the fact that sequential choice environments are often not static and stable. Rather, the distribution from which available options are drawn may change over time. For example, as (or if) one gets older and wealthier, the set of available and attainable options to consider when buying a car or a house (or attracting a spouse) may increase in value as well. Conversely, if the pool of available options shrinks over time, like unattached dance partners at a party or unsold lots left in a new housing development, the distribution of quality may also fall over time. How well will different search rules fare in such changing environments?

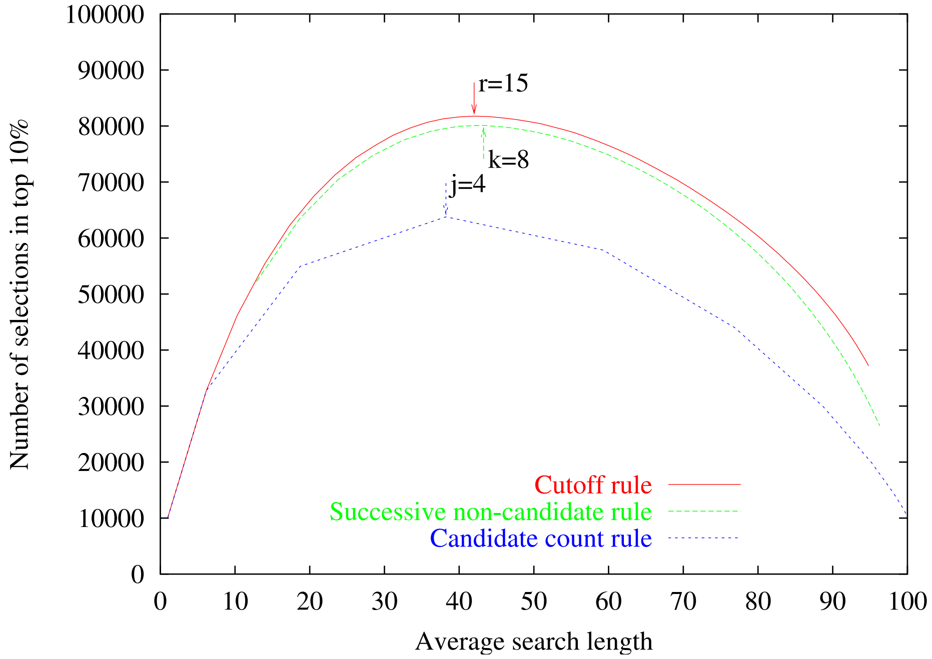

Figure 2: Performance of three search rules for the goal of selecting an applicant in the top 10% of the quality ranks of all $N = 100$ applicants (simulation results computed over 100,000 rounds and displayed as in Figure 1 but with different y-axis range).

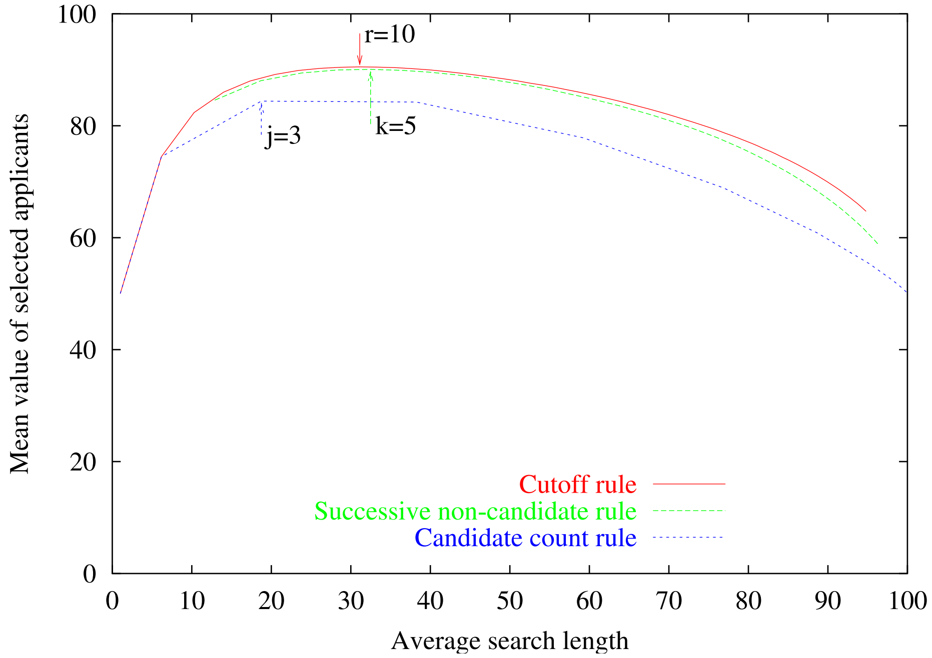

Figure 3: Performance of three search rules for the goal of maximizing the mean value of the selected applicant (simulation results computed with $N = 100$ applicants over 100,000 rounds and displayed as in Figure 1 but with y-axis showing mean value selected as percent of total value range).

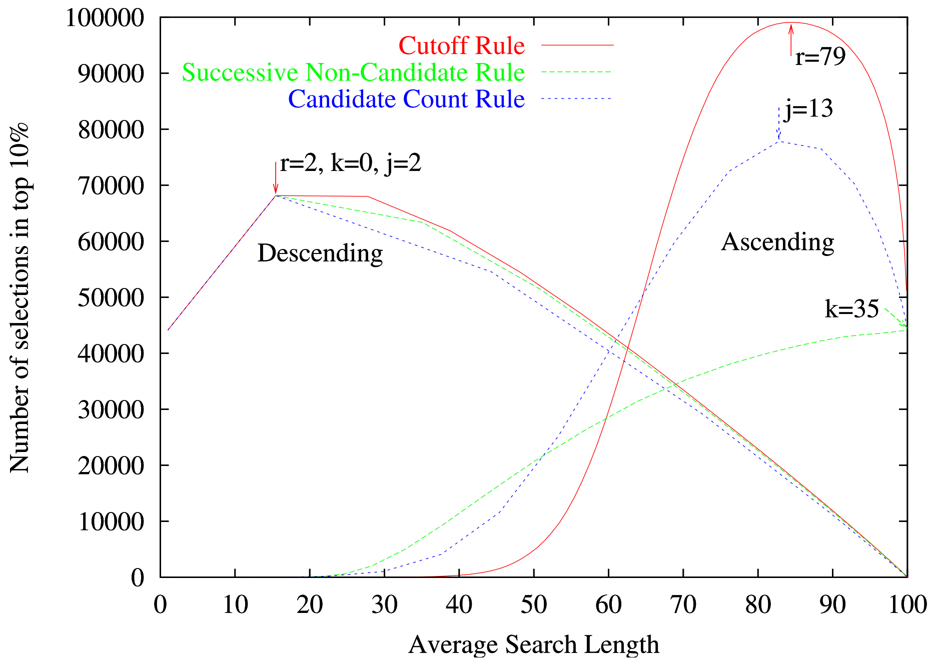

Figure 4: Performance of three search rules (see text) for the goal of selecting an applicant in the top 10% of the quality ranks of all $N = 100$ applicants, but with the distribution from which the applicants are drawn changing over time, either descending by 100% (peaks on left) or ascending by 100% (peaks on right) over the course of the 100 applicants (simulation results computed over 100,000 rounds and displayed as in Figure 1 but with different y-axis range).

To find out, we have simulated the three rules discussed so far in a range of settings where the distribution of available values shifts over time. Here we report results for two environments: one in which the distribution of values rose by 100% over the course of the 100 items that could be seen (i.e., item values were drawn from the uniform distribution (0.01, 1.01] for the first item, (0.02, 1.02] for the second item, (0.03, 1.03] for the third, and so on up to (1.0, 2.0] for the 100th), and one in which the distribution of values fell by 100% over the 100 items (i.e., from (-0.01, 0.99] for the first to (-1.0, 0.0] for the 100th).

In Figure 4 we show the results of the three search rules, for the goal of selecting an applicant in the top 10% of the quality ranks of all $N = 100$ applicants, in both the ascending and descending environments. In the ascending environments, all three rules not surprisingly must search much longer than in the static environment case (Figure 2) because the best applicants will tend to come at the end of the overall-rising sequence. What is more surprising is that, while the cutoff rule still does the best (achieving 99% success with $r = 79$), the candidate count rule now greatly outperforms the successive non-candidate rule. The reason for this is that the environment structure does not match the "expectations" of the latter rule, causing its performance to suffer: In a rising sequence, new candidates are encountered often, and the gap between them does not grow sufficiently to allow appropriate triggering of the non-candidate rule. The opposite problem plagues all three rules equally in the descending environments: New candidates are rarely encountered as the sequence falls overall, so the rules must stop search as quickly as possible (e.g., with $r = 2$ for the cutoff rule, which yields a 68% success rate). Rules that could stop search on applicants that are not candidates would clearly fare better in these environments.

3.3 Searching for further answers

In general, in stable environments where possible options are encountered independently, one after another, with no knowledge of their distribution and no recall available, searchers can be ecologically rational by simply using the best-so-far? cue with the cutoff (or successive non-candidate count) rule and a low parameter value. This approach (e.g., 'Try a dozen') is a simple way to perform well, according to multiple criteria, in sequential search. However, its empirical match to observed behavior remains to be determined. In a preliminary study, we have observed that people change their search behavior appreciably when directed to search with different goals. We had 29 participants search 9 sequences of 100 numbers (positive integers drawn uniformly from different ranges) with three different goals. Each person thus saw three sequences for each goal, with the first for each goal being a practice run. The three goals (and their corresponding payoff conditions) were: select the highest of the 100 values (and receive payoff only if successful), select a number in the top 10% of the values in the current sequence (again receiving payoff only if successful), or select a high number (receiving payoff proportional to how high the selected number was in the current sequence).

How did participants search in these three different conditions? For the first goal (select highest), the mean search length was 41 and the success rate was 18%; for the second goal (select in top 10%), the mean search length shortened to 28 while the success rate grew to 95%; and for the third goal (select a high value), the search length shrank further to 24 and the selected value was on average at the 94$^{\text{th}}$ percentile of the possible range (with 79% of the selections landing in the top 10% of the range). Thus people seemed to adapt their strategy to the particular search goal they were presented with, rather than using a single simple rule as suggested by the 'Try a dozen' approach. Further experiments of the sort done by Seale & Rapoport are needed to determine just what strategies were used in each case.

Not only can it be reasonable to adapt one's search strategy to the particular situation at hand, but it also can be advantageous to adapt one's strategy over the course of a particular single search. For instance, if the final search horizon is approaching and a suitable option has not yet been found (possibly leading to 'end-panic'), then lowering one's aspiration level can be sensible to increase the chances of making a choice before it is too late. Such an adaptive strategy has been proposed for animals searching for mates during a fixed breeding season — as the end of the breeding season looms, aspiration levels should drop, because any mating opportunity is better than none (Johnstone, 1997). For a modified secretary problem setting with the specific goal of finding one of the three highest values in a sequence of (for instance) 100 numbers, Quine & Law (1996) derived the optimal strategy, which also relies on falling aspiration levels: First follow the cutoff rule presented earlier with $r = 33$ until either search is stopped at a candidate or the 58$^{\text{th}}$ applicant is seen; after this point, stop on any best-so-far or second-best-so-far applicant until the 78$^{\text{th}}$ applicant is seen; from that point on, select any applicant qualifying in the best three so far. In additional simulations, we have found that similar *adaptive cutoff rules* can increase search performance considerably on the search goals (aside from selecting the single highest) we considered here.

To make further progress studying the psychological mechanisms that people use in sequential search tasks, we must continue to combine three main approaches. First, we need to identify a set of plausible mechanisms that could be used in different settings, something that can be done with simulation and mathematical analysis. We are currently exploring a range of environments (e.g., with different and time-varying distributions of values) and variations in agents' knowledge of them (e.g., whether they observe just the best-so-far? cue, or ranks, or actual values) to find what heuristics can be ecologically rational in those settings. Second, we need to perform laboratory experiments in which we look for evidence of the use of particular proposed mechanisms. And third, we must gather empirical observations of decision making in real-world sequential search situations (such as, when driving, choosing which rest area to pull over and stop at, or when stocking a warehouse with goods whose price varies on a daily basis, choosing when to buy new supplies — see Bolle, 1979), to gain more evidence about how people actually behave in these tasks.

4 Conclusions

Good decisions can be made with little information. Here we have shown how simple binary values, whether or not something is recognized or is the best option seen so far, can be sufficient when paired with appropriate heuristics operating in appropriately-structured environments to yield high performance on choice and search tasks. Other examples of this general trend abound; for instance, when more information than just recognition is available about two available options (and hence the recognition heuristic cannot be applied), simply tallying the number of pro and con attributes for each choice (Dawes, 1979) or even basing choice on a single distinguishing attribute ('one-reason decision making' — Gigerenzer & Goldstein, 1996) can rival appropriately-weighted processing of all information.

In fact, in contrast to the usual image, handed down from traditional views of rationality, that more information will always be better, recent research has begun turning this notion on its head by showing ways in which less information can actually lead to *greater* decision-making performance (Hertwig & Todd, in press). We saw this in the less-is-more effect with the recognition heuristic; another example is simple heuristics leading to more robust generalization performance on new data sets (which may also hold for the sequential search tasks we have examined here). Thus, people and other animals may not only suffer little from using ecologically rational strategies to make decisions with minimal information, they may even gain an edge.

References cited

Bergstrom, Carl T. & Real, Leslie A. (2000). Towards a theory of mutual mate choice: Lessons from two-sided matching. *Evolutionary Ecology Research*, 2(4):493–508. 6

Bolle, Friedel (1979). *Das Problem des optimalen Stoppens: Modelle und Experimente*. Ph.D. thesis, University of Hamburg, Germany. Verlag Peter D. Lang, Frankfurt am Main. 17

Borges, Bernhard, Goldstein, Daniel G., Ortmann, Andreas, & Gigerenzer, Gerd (1999). Can ignorance beat the stock market? In Gerd Gigerenzer, Peter M. Todd, & the ABC Research Group, *Simple Heuristics That Make Us Smart*, chapter 3. Oxford University Press, New York. 5

Boyd, Robert & Richerson, Peter J. (1985). *Culture and the Evolutionary Process*. University of Chicago Press, Chicago. 2

Corbin, Ruth M. (1980). The secretary problem as a model of choice. *Journal of Mathematical Psychology*, 21:1–29. 10

Corbin, Ruth M., Olson, Chester L., & Abbondanza, Mona (1975). Context effects in optional stopping decisions. *Organizational Behavior and Human Performance*, 14:207–216. 7, 8, 9, 10

Dawes, Robyn M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, 34:571–582. 17

DeGroot, Morris H. (1970). *Optimal Statistical Decisions*, chapter 13. McGraw–Hill, New York. 6

Ferguson, Thomas S. (1989). Who solved the secretary problem? *Statistical Science*, 4(3):282–296. 6

Galef, Bennett G., Jr. (1987). Social influences on the identification of toxic foods by norway rats. *Animal Learning and Behavior*, 15:327–332. 4

Gigerenzer, Gerd & Goldstein, Daniel G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103:650–669. 3, 17

Goldstein, Daniel G. & Gigerenzer, Gerd (1999). The recognition heuristic: How ignorance makes us smart. In Gerd Gigerenzer, Peter M. Todd, & the ABC Research Group, *Simple Heuristics That Make Us Smart*, chapter 2. Oxford University Press, New York. 4, 5

Henrich, Joe & Boyd, Robert (1998). The evolution of conformist transmission and between-group differences. *Evolution and Human Behavior*, 19(4):215–242. 2

Hertwig, Ralph & Todd, Peter M. (in press). More is not always better: The benefits of cognitive limits. In Laura Macchi & David Hardman, editors, *The Psychology of Reasoning and Decision Making: A Handbook*. John Wiley & Sons, Chichester, UK. 17

Hey, John D. (1982). Search for rules for search. *Journal of Economic Behavior & Organization*, 3:65–81. 6

Johnstone, Rufus A. (1997). The tactics of mutual mate choice and competitive search. *Behavioral Ecology and Sociobiology*, 40(1):51–59. 16

Kahan, James A., Rapoport, Amnon, & Jones, Lyle V. (1967). Decision making in a sequential search task. *Perception & Psychophysics*, 2:374–376. 7, 8, 9, 10

Moon, Philip & Martin, Andrew (1990). Better heuristics for economic search—experimental and simulation evidence. *Journal of Behavioral Decision Making*, 3:175–193. 6

Mosteller, Frederick (1965). *Fifty Challenging Problems in Probability with Solutions*. Dover Publications, Inc., 31 East 2nd Street, Mineola, N.Y. 11501. Reprint published by Dover in 1987. Originally published in 1965 by Addison-Wesley, Reading, MA. 7

Payne, John W., Bettman, James R., & Johnson, Eric J. (1993). *The Adaptive Decision Maker*. Cambridge University Press, New York. 3

Quine, Malcolm P. & Law, J. S. (1996). Exact results for a secretary problem. *Journal of Applied Probability*, 33(3):630–639. 16

Sargent, Thomas J. (1993). *Bounded Rationality in Macroeconomics*. Oxford University Press, Oxford. 3

Schotter, Andrew & Braunstein, Yale M. (1981). Economic search: An experimental study. *Economic Inquiry*, 19:1–25. 10

Seale, Darryl A. & Rapoport, Amnon (1997). Sequential decision making with relative ranks: An experimental investigation of the 'secretary problem'. *Organizational Behavior and Human Decision Processes*, 69(3):221–236. 8, 9, 11, 12, 16

Seale, Darryl A. & Rapoport, Amnon (2000). Optimal stopping behavior with relative ranks: The secretary problem with unknown population size. *Journal of Behavioral Decision Making*, 13:391–411. 9, 10, 11

Shanteau, James (1992). How much information does an expert use? Is it relevant? *Acta Psychologica*, 81:75–86. 3

Simon, Herbert A. (1990). Invariants of human behavior. *Annual Review of Psychology*, 41:1–19. 3, 6

Stigler, George J. (1961). The economics of information. *Journal of Political Economy*, 69:213–225. 3

Todd, Peter M. (1999). Simple inference heuristics versus complex decision machines. *Minds and Machines*, 9(4):461–477. 6

Todd, Peter M. (2000). Fast and frugal heuristics for environmentally bounded minds. In Gerd Gigerenzer & Reinhard Selten, editors, *Bounded Rationality: The Adaptive Toolbox*. MIT Press, Cambridge, MA. 2

Todd, Peter M. & Miller, Geoffrey F. (1999). From pride and prejudice to persuasion: Realistic heuristics for mate search. In Gerd Gigerenzer, Peter M. Todd, & the ABC Research Group, *Simple Heuristics That Make Us Smart*, chapter 13. Oxford University Press, New York. 6, 10, 11, 12, 13

Vriend, Nicolaas J. (1996). Rational behavior and economic theory. *Journal of Economic Behavior & Organization*, 29:263–285. 3