Is Decision Making Rational and General?

The Optimising Model of Decision Making

Tomasz Smolen

Pedagogical University of Krakow, Poland

tsmolen@up.krakow.pl

***Abstract*- A unifying and optimising model of decision making is presented in this paper. The model is based on two hypotheses: that there is one general strategy of decision making which can recreate many specific strategies depending on some environmental parameters, and that human decision making is rational if we take into account all cognitive limitations, features of the environment and the task that an agent actually attempts to solve. Testing of the model aimed to verify specific hypotheses about the match between the model's predictions and the empirical data and the existence of some phenomena expected due to the workings of the model's hypothetical mechanism, as well as the aforementioned general hypotheses about decision making. The model's performance has been found to be consistent with subjects' results. Particularly, the model has a better match to data than other related models.**

Keywords*- *Decision making; Rationality; Mathematical modelling

I. INTRODUCTION

Similarly as in many psychological areas, early theories developed in the domain of decision making were normative (e.g., [1]). Likewise, as in other psychological fields, these theories were soon abandoned in favour of concepts that tried to explain numerous observations which did not fit normative laws and which seemed to contradict the thesis about the rationality of human cognition. Thus, the history of decision making theorising is a history of the transition from a purely normative explanation (the expected value hypothesis), through theories that allow for the inclusion of parameters to explain those aspects of human behaviour which seem to be irrational and biased (the expected utility hypothesis [2]) and theories which accept yet more parameters (the subjective expected utility hypothesis [3], the Prospect Theory [4]), and so to the present theories which emphasise biases and mechanism-based phenomena. The direction in which theories of decision making have evolved has therefore been determined by the goal of achieving the best explanation for the data. There have, however, been many drawbacks with this evolution. The resulting loss of simplicity was but one cost of these theories' capacity to explain more of the observed phenomena. According to Paulewicz [5] another disadvantage involved a sole reliance on empirical data to provide justification for such theories and the disappearance of a direct reference to more general psychological and cognitive theories. He suggests that the latter effect was not caused by the rise of the complexity itself, but rather by either neglecting or carelessly defining such basic terms as goal, rationality and behaviour.

Currently, one of the most widely accepted theories of decision making is the Adaptive Toolbox theory [6]. This concept assumes that people use many middle-range tools that exploit regularities in the environment instead of using a "single hammer" which needs a complete representation of the environment for all purposes. There are two main features of the Adaptive Toolbox theory. First, it is based on Simon's [7] idea of bounded rationality. Second, it assumes a diversity of strategies. The strategies that belong to the Adaptive Toolbox are mostly fast and frugal heuristics [8]. They are called fast and frugal because they limit information searching and do not require much computation. They are heuristic because they are ecologically rational rather than logically consistent.

The three advantages of simple heuristics (swiftness, frugality, and ecological rationality) are not incontrovertible. The swiftness of the heuristics supposedly lies in the fact that they use simple operations. For example, the "Take the Best" (TTB) heuristic makes a decision by simply comparing the cue values, while the weighted additive algorithm (WADD), the strategy most often used as a counterexample for the Adaptive Toolbox strategies, uses more complex linear transformations. However, if people are equipped with mental mechanisms that are specialised in parallel processing of some types of information [9], or if the processes of computing cue values and weights can be automatised [10], WADD can be as fast as TTB (or even faster, since TTB is recursive and WADD is simple). The second feature, frugality, allegedly consists of the fact that simple heuristics use less information or information that is easier to collect. In fact, taking the two major strategies from both classes as an example again, WADD requires information about the validity of cues (the proportion of number of cases in which a cue indicates the correct alternative to the number of all cases in which it discriminates between alternatives), whereas TTB uses simpler information about the order of validities. The strength of this argument, however, is weakened by two facts. Firstly, the order of validities is very difficult to obtain from the environment [11]. Secondly, WADD can work efficiently with validities estimated on the basis of such order [12]. Finally, the ecological rationality of simple heuristics is more of a declared rather than an established feature. Several experiments [13, 14, 15, 16] have revealed that simple heuristics are used less frequently than following from their hypothetical accuracy.

The two hypotheses that underlie the Adaptive Toolbox Theory (namely: bounded rationality of the decision making and diversity of decision making strategies) are to be questioned in the current paper. This will be attempted both by theoretical analysis and by proposing a decision making model which attempts to find the optimal solution of the choice problem by means of one general mechanism which takes into account environmental conditions.

A. Rationality of Decision Making

An important feature of the strategies that belong to the Adaptive Toolbox is the fact that they are boundedly rational. The notion comes from Simon and is characterised as follows [17, p. 198]:

> For the first consequence of the principle of bounded rationality is that the intended rationality of an actor requires him to construct a simplified model of the real situation in order to deal with it. He behaves rationally with respect to this model, and such behavior is not even approximately optimal with respect to the real world.

What does "not even approximately optimal" mean? Is a boundedly rational agent's behaviour even not close to optimality? If it is not, how can we know what the agent is actually doing? When are we justified in claiming that an observed agent, which exhibits behaviour "A", is in fact aiming at behaviour "B", but fails to do so optimally? Paulewicz [5] points out that the only situation in which we are allowed to do so is when we know that (a) the agent has a certain goal, and (b) approximetly optimal action that leads to achieving this goal is behaviour "B". Note that the "optimal" mentioned in Point (b) means objectively optimal, not optimal with respect to some simplified, internal model. Therefore, behaviour can only be understood and analysed as far as it is known what goal it is meant to achieve and what the optimal solution of the problem is [5]. Besides, it is not assumed in the argument that the agent knows or can find the optimal solution (or rational solution, as in this paper the terms "optimal" and "rational" are used interchangeably, which is reasonable in the context and commonly practised in this field of research, see [18]).

One can hardly disagree with Simon's view that agents often cannot reach a solution which is optimal in the classical, normative way due to its computational and informational costs. But there are many cases in which agents seem to act non-optimally, whereas their behaviour in fact is optimal. It can be seen as non-optimal with regard to the task as it is defined by an observer because the task that an agent is actually trying to solve is a different one [5]. The rationality used in such situations, compared to the normative one, is not the rationality that requires a simplified model (as Simon suggests). The rationality used in such cases is the same as that in the normative case if it took into account all limitations, costs, and the way in which the agent represents the task.

A good example of a phenomenon often considered to be a common non-optimal behaviour is the probability matching strategy. The simplest task in which this strategy can be observed is the multiple-alternative iterated-choice task (the multi-armed bandit problem). The task entails the series of choices of actions from a finite set. Every action generates a reward, with the goal of the task being to maximise the cumulated reward. The probability matching strategy involves choosing the action with a probability proportional to its relative reward value. It is clear that the probability matching strategy is not an optimal solution for a well specified, static multi-armed bandit problem. The optimal agent is the one which always chooses the most promising action. Of course, if the agent does not know the reward distribution, it must explore the environment in order to find the best action. Otherwise it may end up repeatedly choosing an alternative that is not the best, but which seemed to be the best on the basis of the first few noise-laden feedbacks.

However, even though in an uncertain situation the frequency of choosing the best action should tend towards one while the agent's uncertainty about the expected values of actions declines [19]. Despite this fact many research results (e.g., [20, 21]) have shown that people tend to use the probability matching strategy in most cases. Therefore, if people and other animals attempt to solve a multiple-handed bandit problem, more specifically: if they attempt to solve a stationary, fully defined, isolated multi-armed problem, they usually do not behave rationally using a probability matching strategy. But can we be sure that the strategy observed is really chosen as a solution to this specific problem?

The probability matching strategy seems to be suboptimal, but it can be shown that it is actually optimal in an environment characterised by at least one of two features typical of the environment in which human evolution occurs. The first feature involves competition for resources; the second is environmental dynamism. It has been demonstrated that the probability matching policy is an evolutionarily stable strategy for an individual agent facing its competitors [22, 23]. It seems obvious that its alternative, the greedy strategy, would lead to an irrational situation in which all agents share rewards from the most abundant source, while the less promising one remains unexploited. The second feature that makes the strategy evolutionarily stable is environmental dynamism [24]. As is stated above, the optimal strategy in an uncertain environment is not a greedy one [25]. An unstable environment is always, by rule, uncertain, so it requires a more sophisticated policy. What is more, Daw, O'Doherty, Dayan, Seymour, and Dolan [26] have shown that a strategy that is similar to probability matching has the best match to human behaviour among all the tested strategies in a dynamic environment. Other modelling research has also shown that the probability matching strategy may be an emergent feature of an evolving, foraging system [27]. The probability matching rule deserves wider review not only because it is a good example of an ostensibly suboptimal strategy which, after closer investigation, turns out to be, in fact, optimal, but also because it is used as a part of the proposed model which will be further described below.

The question of seeking an optimal solution can only be discussed on the grounds that people have the ability to explore strategies. Moreover, if people did not try different strategies in order to maximise results, or if they did so without any regard to optimality (e.g., relying on common belief, maladaptive routine, or random choice), the model proposed in the current paper could not be accurate. But several studies show that people are able to adapt their strategies to the environment. Bröder and Shifter [28] have shown that people choose decisional strategies appropriately. Over 65% of their choices fit an optimal strategy in a static environment, although they mostly fail to detect a change in the environment that occurs while the task is being performed. On the other hand, Newell and Lee [29] have demonstrated that throughout the whole task people can vary the amount of information taken from the environment according to changes. Furthermore, the results provided by Payne, Bettman and Johnson [30] show that people are highly adaptive when responding to changes in the environment which occurred during the experiment.

B. Generality of Decision Making

The Adaptive Toolbox theory states that there are many qualitatively different strategies of decision making. The alternative hypothesis presented herein claims that there is one general strategy which is used in all classes of decisional situations. This hypothesis may seem improbable because it is evident that people behave differently in different environments (e.g., [31, 32]). However, this fact does not falsify the hypothesis about general decision strategy. The strategy might be flexible; it could recreate different specific strategies by adjusting some parameter. Newell [33] has noticed that all major decision strategies from the Adaptive Toolbox can be ordered with respect to the amount of information required for making a decision. This has led him to the hypothesis that there is a mechanism which makes a decision on the basis of the amount of information available in the environment and that it works like a corresponding Adaptive Toolbox strategy which requires the same amount of information.

However, this interesting proposition is purely theoretical. Newell has not proposed any particular mechanism which fits the criteria that he himself had stipulated. Nevertheless, a few suggestions for a mechanism that unifies compensatory and noncompensatory strategies have been described by others. A significant example is the study conducted by Lee and Cummins [34], who tested the Evidence Accumulation Model (EAM) of decision making. The model works by accumulating the cue-based evidence indicating that a given alternative is the best one. When the difference between the accumulated evidence strength of the stronger and the weaker alternative surpasses a specified value, the respective alternative is chosen. The idea of the evidence accumulation is not quite new; it originally appeared in the Random Walk Model [35, 36], and was later applied in many decision making models (especially in Diffusion Model [37]), but an interesting feature of Lee and Cummins' model is that it can reproduce the performance of both compensatory and noncompensatory strategies depending on the value of the requested evidence strength.

Also within the Adaptive Toolbox Theory, a mechanism that adapts agents' behaviour to environmental constraints has been proposed. A model based on Strategy Selection Learning theory (SSL [38]) learns to use different strategies from a hypothetical repertoire. The model uses the reinforcement learning paradigm [25] to learn which strategy leads to the highest dividend in different environments or classes of decision situation. The basic action which is evaluated after getting feedback is the choice of a strategy. Also the basic state of the environment is defined with general terms—it is the whole decisional situation. At every step the model picks a predefined strategy from its repertoire and applies it. After receiving feedback it updates the metastrategy (the strategy for choosing the elementary strategy) according to the reinforcement learning rules. Although the SSL model does not apply one general strategy but rather uses a metastrategy to manage many simple strategies, it is nevertheless possible that in some conditions it will emergently exhibit a continuous adaptation of the informational intake.

Although the two previously mentioned strategies accommodate information intake, they differ considerably. While the EAM does indeed recreate different classes of strategy, SSL just uses different predefined strategies according to a policy determined by a metastrategy. On the other hand, SSL tends to have optimal performance as it is based on a general optimising paradigm of reinforcement learning, while EAM does not aim to achieve optimality directly.

Taking into consideration these objections to the widespread notions of decision making, it seems reasonable to propound a thesis according to which the assumptions behind the contemporary view on this field (involving, namely the bounded rationality and variety of strategies), are a little too far-reaching. Verification of the model proposed will also be a verification of its two fundamental claims: (a) human decision making is close to optimal, and (b) there is one general and flexible strategy which recreates many specific strategies. Phenomena which seem to deny the optimality of decision making have usually been caused by having a task or an environment poorly defined or by not having properly understood the actual goal that an agent attempted to achieve. These phenomena have also emerged due to neglecting the fact that agent's performance depended on the environment in which the agents operated or on agents' features and preferences. The model described below aims to take into consideration the previously mentioned premises.

Moreover, the verification of the model is the verification of an entire class of models because the Optimising Model integrates many models, as it performs differently depending on the features of the environment. It includes both models belonging to adaptive toolbox as well as models from outside of it, as long as they can be located on Newell's continuum of information requirement.

II. THE MODEL

The optimising model relies on the assumption that a decision maker's main goal is to maximise a reward. The reward is equal to the difference between received payment and incurred costs, so the decision maker must achieve two subgoals: maximising the probability of a correct choice and minimising the costs. These two subgoals conflict because the more information from the environment one receives, the higher the probability of choosing the right alternative becomes, but always at the expense of a rise in costs incurred by an agent.

There have been many studies trying to find the relation between the costs and expected reward value. The earliest ones based the optimal criterion for terminating an information search on the difference between odds that a given alternative is correct. The results of research are ambiguous. Edwards and Slovic [39] have showed that people get as much information from an environment as is optimal, taking into account its cost and expected gain. On the other hand Fried [40] have found out that although people use an optimal amount of information when they have to decide in advance how much information they are going to use, they get less than an optimal amount of information when they can stop collecting information at any time. Other studies [41, 42] have shown that people do not use an optimal amount of information. A criterion of optimality which was derived from Bayesian theory was tested by Rapoport [43]. The subjects' performance was only partially consistent with model predictions.

A. The model's Subgoals

1) Maximising the probability of correct decision

As stated by Edwards [44, p. 188] "information should be sought only if the expected cost of obtaining it is less than the expected gain from it". This thesis can hardly be considered to be false (at least as a normative statement), but how can "expected gain" be assessed? The solution is quite simple if the reward value is known, as the expected gain is then proportional to the probability of receiving the reward. The more information is received from the environment, the more likely it is that the decision made on the basis of this information will be correct. Thus in the model the relation between the probability (*P*) of correct decision and the amount of the analysed information (*n*) is supposed to increase with negative acceleration (fig. 1), and is determined by the formula 1,

$$P(n) = 1\text{-}(1\text{-}U(n))^n \tag{1}$$

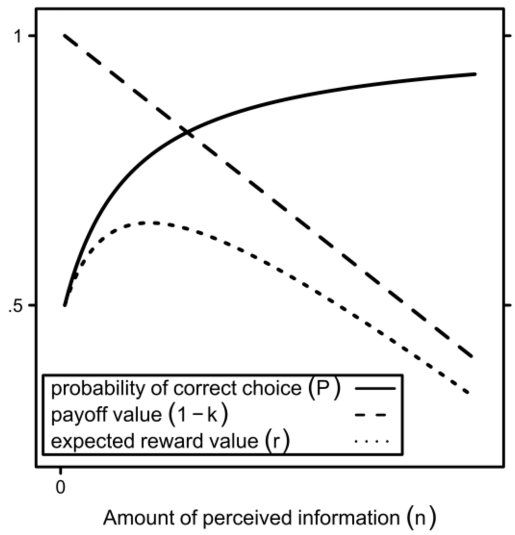

Fig. 1 The expected reward value as a product of the probability of the correct choice and the payoff value

Function *P* uses Function *U*, which returns the probability (in a discrete case) or the density of the probability (in a continuous case) of choosing the better alternative on the basis of the information in Point *n*. Note that the Function *U* can either be fitted to observed data or set on the basis of the properties of the environment. The latter case is justified when it can be supposed that subjects know the properties of the environment (e.g., when they are provided to them explicitly or when they can be learned during training). If the Function *U* is determined by the properties of the environment, the model has no free parameters.

2) Minimising the costs:

An environment in which information is free is both unchallenging and unreal. It is unchallenging because the optimal strategy in such an environment is simply to gather all possible information as such a policy maximises the probability of making correct decisions. It is unreal because in the natural environment information is hardly ever free. Gathering information usually consumes material resources and energy or carries risk, but even if it does not, it still requires time and cognitive resources to process the data. There are many possible functions that describe how costs increase as the amount of gathered information rises (e.g., linear, proportional to cue value or random). Since the costs are cumulative, all of them are ascending and the expected value of the reward (which is equal to the difference between the starting value of the payoff and the value of the cost) declines over *n*, regardless of an interpretation of *n* (e.g. time, information portion, or processing progress).

As both subgoals described above depend inversely on the amount of information obtained from the environment, this amount must be optimised with regard to an overall reward. Function *r* (see formula 2) returns the expected reward on the basis of the probability of the correct choice (*P*), gain of the correct decision (*g*), the function of the information cost (*k*) and the cost of a decision (*K*):

$$r(n) = P(n)g - k(n) - K \tag{2}$$

The optimal amount of received information is $\text{argmax}_n r(n)$ (where $\text{argmax}_x f(x)$ is the value of *x* for which function *f* returns the highest value).

As long as both functions *P* and *k* are continuous and integrable, function *r* is also continuous and integrable. The feature of continuity is helpful as *k* is not necessarily a discrete cue charge, but it can reflect any costs (e.g., time costs). Integrability is important due to the possibility of applying the probability matching strategy.

Since optimal strategy depends on the environment in which the decision is taken, people could be expected to take the optimal amount of information in every trial only if they explicitly computed the expected reward on the basis of the amount of information taken. It is far more probable that people learn the profitability of actions on the basis of feedback. Thus, because the greedy strategy has previously been said to be non-optimal when learning about an uncertain environment, it seems plausible that the probability matching rule will be applied to seeking information.

To reflect the probability matching rule in the optimising model, Function *R* has been built into it. The function returns the probability that a decision will be made after obtaining the amount of information that is in the range [*l*, *h*] and it is described by the formula 3:

$$R(l,h) = \frac{\int_l^h r(n)dn}{\int_0^I r(n)dn}, \tag{3}$$

where *I* is the maximum amount of information available. The function reflects the hypotheses that the probability of getting an amount of information in given interval is proportional to the expected reward for the action taken on the basis of the information. Point estimation for the expected reward is given by the function *r*, so the relative value is a proportion of the integral of *r* over the given interval to the integral of *r* over all the function domain.

B. The model performance

Due to its abstract formulation, equation 1 can reflect various probability functions by using various *U* functions. Let us consider the two most interesting instances of function *U*.

1) The multiple cue choice task:

Symbol *n* refers to a cue revealed in the *n*-th step. Then Function *U* specifies the probability of choosing the better alternative on the basis of cue *n*, which takes into consideration its validity and its discrimination rate (frequency of cases in which the cue value is different for different alternatives). Property that meets this criterion is “success” [45]. In that case, the Function *U* and (as a consequence) the Function *P* are discrete and not differentiable.

2) The item recognition task:

In the task (as described by Ratcliff [37]) *n* refers to time, and the probability of choosing the better alternative on the basis of information perceived in a given time interval rises linearly with the length of the interval. Then Function *U* is equal $(d/dn)cn$, where *c* is Ratcliff’s drift parameter, and it takes the following form: $U(n) = c$.

Depending on the Function *U*, the Function *P* can be more or less concave. The more concave it is, the earlier Function *r* achieves its maximum, and therefore (as $-k(n)$ decreases) the bigger the maximum value is. Accordingly it is optimal to use the order of cues that maximises the concavity of *P*, which is, in the case of the multiple cue choice task, the order based on success. Therefore the success determines the order of cues which allows the most efficient information search, but the model itself does not include a mechanism for finding proper cue search order. Empirical findings [45, 46] confirm the expectation that cue search order is based on success.

One of the main features of the model is its generality, which is defined here as a capacity for recreating many decisional strategies which are considered to be elements of the Adaptive Toolbox [6] and some strategies outside of it. Taking Newell’s [33] assumption about the continuous quantitative difference between strategies, it is expected that the model will be able to recreate strategies which use different amounts of information. As the difference is continuous, the model’s performance should not recreate a set of particular strategies but rather any point on the continuum from the most to the least informationally demanding (e.g., from WADD to “take one” [14], or even the extreme case of random choice). Since the model has no free parameters, all variation within performance must derive from environmental features.

The environmental feature which has been widely proved to influence the amount of information taken for a decisional process is the information cost. Available results concern both the direct costs expressed in reward units [13, 46] and time costs [30, 47] as well as memory retrieval or computational costs [48, 31]. The model can reflect the impact of costs of any type on the amount of information collected from the environment as long as the relation between the costs and the information can be expressed as a function. In particular, the model can account for well-specified cognitive costs of information processing. It is probable that people treat the costs of processing similarly to the objective external costs: Newell and Lee [29] have shown that people tend to minimise the amount of information processed during decision making. Indeed, the optimising model uses various amount of information depending on the information costs. The higher the costs are, the lower both the optimal amount of information and the expected amount of information taken by an agent who uses the probability matching strategy are (see table I).

TABLE I INFLUENCE OF THE COSTS ON THE NUMBER OF CUES REVEALED BY THE MODEL

Cue cost	**Number of cues revealed**		
	1st quartile	**Optimal number**	**3rd quartile**
.1	1.13	2.25	2.39
.2	1.02	1.54	2.26
.3	.84	1.14	1.92
.4	.49	.87	1.01

Note: The quartiles refer to informational intakes according to the probability matching rule.

Another feature of the environment which is supposed to affect the amount of information processed by an agent is the relation between the validities of cues. The decisional environments can be divided into two classes: compensatory and noncompensatory. A noncompensatory environment is defined as one in which, in every subset of cues, the most valid cue has a higher validity than the sum of validities of the remaining cues [49]. In compensatory environments the validities of cues do not fit the above description. The feature of being compensatory or noncompensatory can be generalised as being continuous. Let us assume that the more the most valid cue surpasses the validities of the remaining cues in every subset, the more noncompensatory the environment is.

According to rational analyses, compensatory strategies like WADD will perform better in compensatory environments, whereas noncompensatory strategies, like TTB, will similarly perform better in noncompensatory environments [50]. The actual performance of the examined subjects confirmed these predictions. It has been shown [13, 51] that in noncompensatory environments people use compensatory strategies less frequently than in compensatory environments. The model recreates this phenomenon as well; the more noncompensatory the environment is, the less cues it reveals (see table II).

TABLE II INFLUENCE OF THE TYPE OF ENVIRONMENT ON THE NUMBER OF CUES REVEALED BY THE MODEL

Validity imparity	Number of cues revealed		
	1st quartile	Optimal number	3rd quartile
-1	1.02	1.75	2.33
-.7	1.07	2.04	2.35
-.4	1.11	2.18	2.37

Note: The quartiles refer to information takes according to the probability matching rule. The bigger the validity imparity, the more compensatory the environment (imparity is an exponent of validity regression line).

C. The model's Predictions

There are three rules by which every decisional strategy can be characterised. The searching rule determines which algorithm is being used to get information from the environment, particularly in what order the cues are revealed. The stopping rule determines how much information the strategy takes from the environment. Finally, the decision rule determines which decision is made on the basis of the information taken. Most studies, due to practical reasons, examine just one of the strategy's features. In the present paper the searching rule is the feature which is at the centre of attention. There are two reasons for this. Firstly, the model's predictions concerning the amount of information taken are quite different from other model's predictions (as was stated previously, the cue order used by the Optimising Model is consistent with cues' success). Secondly, according to both the searching rule and the decision rule, the model gives deterministic predictions, which are more difficult to test quantitatively than indeterministic predictions.

The process of optimising the amount of information taken from the environment in the model described herein involves a trade-off between two opposite goals (namely maximising the probability of making the correct decision and minimising the costs). One of the ways to determine whether there actually are two subprocesses which aim to fulfil these two goals is to try to affect the subprocesses selectively. The hypotheses assuming the separation of the subprocesses are based on the Prospect Theory [4], which predicts what conditions can influence the perceived scale of gain and loss. The Prospect Theory can therefore be used to manipulate the representation of the values of costs and rewards.

The Prospect Theory states that subjective perception of loss or gain is not a linear function of objective changes. The actual function has four features. (a) It is defined in the domain of gains and losses instead of absolute values. (b) In the case of two identical objective changes the perception of loss is bigger than the perception of gain. (c) The function is convex for losses and concave for gains. (d) The function increases with positive acceleration for losses and negative for gains.

The complex progress of the function of the subjective value change gives rise to the expectation that in different conditions the motivation to avoid loss and the motivation to gain will have different strengths. One possible influence on the disproportion between these motivational strengths relies on the fact that in the case of two identical objective changes the subjective perception of loss is bigger than the subjective perception of gain. The expected value of the subjective change (which is an effect of an objective change of a given size) is negative when the direction of the change is unknown (e.g., when no feedback on rewards is given). Therefore, the expected strategy chosen by people whose goal is to maximise the accumulated reward would be a conservative one (i.e., by using more information). Such a strategy would allow for assurance from expected loss. By contrast, when people can monitor their progress and react to the negative effect of excessively risky behaviour, there is no need for assurance in the form of a conservative strategy.

Another way to affect the trade-off between the two subgoals of the presented model is related to the fact that the function of the subjective value of the changes is convex for losses and concave for gains, thus explaining the reflection effect, which in turn implies that risk aversion in the positive domain is accompanied by risk seeking in the negative domain [4]. The consequence of this variation in the curvature of the function is the following. The expected value of the change (when the direction of the change is unknown) is positive for a person who perceives their position as negative in comparison to some reference point and, similarly, the expected value is negative for a person whose position is perceived as relatively positive. So, if there is a predecisional trade-off described above that works by adjusting the amount of information to be processed, the output of the trade-off should be affected by (a) the possibility of monitoring the rewards and by (b) the subjective relative position of the decision maker in the dimension of losses and gains.

The best criterion for choosing one of the models described above is the one that includes their match to the observed data and their complexity. The latter can be evaluated only in comparison to other models. The models that were chosen to be compared against the Optimising Model were TTB and the Evidence Accumulation Model. Firstly, because they are the most widely discussed simple models of decision making, and secondly, because they predict among other phenomena the amount of information taken from the environment, which is the main feature of the model's strategy examined in this research.

"Take The Best" is a lexicographic, recursive and deterministic model of decision making. It works by comparing cue values for two alternatives in order of the cue's validity. If the most valid cue discriminates, the alternative which is indicated by the cue is chosen. If the cue does not discriminate, the TTB runs through the set of remaining cues. When the set is empty, random alternative is chosen. Thus, the frequency of revealing n cues by TTB is equal to d_1 for the first cue, $(\Pi_{k=1}^{n-1}\ 1\text{-}d_k)d_n$ for the middle cues $\Pi_{k=1}^{n-1}\ 1\text{-}d_k$ for the last cue, where d_k is the discrimination rate of the k-th cue.

The Evidence Accumulation Model is generally similar to TTB in the way that it reveals cues in order of their validity and makes a decision if cue-based information indicates one of the alternatives with sufficient credibility. However, computing the number of cues revealed is a bit more complicated in this case as more information must be taken into consideration. Apart from the cues' discrimination rate, their validities and threshold parameter must also be included in the analysis. However, the frequency with which a cue is revealed is constant in some intervals of threshold parameter values. The interval limits depend on cue validities. For example, if the threshold parameter is low enough, the EAM behaves exactly like TTB because any discriminating cue makes the evidence strength pass the threshold, thereby causing a decision to be made that is consistent with the cue's indication. On the other hand, if the threshold parameter is higher than some critical value, EAM applies a compensatory strategy which makes a decision on the basis of all cues.

These two models can, however, be a little awkward, as they are both deterministic. It means they cannot be evaluated by the use of the measures that take into account information about the probability of the model given the observed data (e.g., *BIC*, *MDL*, or *K* – bayes factor [52]). The models also make predictions that render them too easily falsifiable on the basis of some observations: both TTB and EAM predict that people would reveal at least one cue. Since the likelihood of at least one success given the model that predicts zero successes is zero, these models would be falsified if a decision was made after revealing no cues. Rejecting these models on the basis of the aforementioned fact, were it ever observed, would be nonetheless a little premature. Such observations, providing their number is small, can stem from an error rather than a systematic trend. Besides, both models can be easily extended with some error parameters to eliminate the problems discussed. After adding the parameter s that describes the probability of stopping the information search, which was typically continued in the original versions, both models become probabilistic and capable of predicting a non-zero likelihood for all numbers of cues revealed. The basic models are then specific cases of the extended models with $s = 0$.

Three hypotheses based on inference presented above in this section were tested. (a) In the loss condition, subjects will collect less information that in the gain condition because of asymmetry of the subjective perception of loss and gain (1st experiment). (b) In the feedback condition, subjects will collect less information that in the no-feedback condition because of the non-linearity of the subjective perception of loss and gain (2nd experiment). (c) There will be a match between the amount of information collected by the subjects and the amount predicted by the model in both experiments. Particularly: the match will be better than for other models considered.

III. METHOD

A. Participants

Two experiments were performed. All 48 subjects (26 women) participated in both. The mean age of the subjects was 20.52 (range: 19-24). The subjects were college students from Krakow, Poland. All were randomly drawn from a pool of students who had agreed to participate in experiments for course credits. Participants were paid for taking part in the experiment. The amount of payment depended on the level of task performance, and on average it was 25.21 PLN (≈ €6). In each experiment the participants were randomly divided into two groups: control or experimental and, independently, they were assigned to one of the two conditions: involving low or high cue costs.

B. Apparatus

The subjects were tested with a computer application consisting of a forced-choice multiple-cue task. The task consisted of a series of decisional situations. In each of them, subjects had to choose one of two given alternatives. The choices could be based on the values of cues. In every decisional situation three binary cues were available. All were covered, and revealing any of them involved some specified costs. Revealing a cue would result in displaying the values of the cue for both alternatives. Choosing the correct alternative was rewarded with a specified number of points. Every decision, irrespective of its correctness, cost another number of points. The costs of revealing a cue could be either high (.1 of reward value) or low (.05 of reward value) depending on the condition. After each decision subjects were given feedback about its correctness, they also had their points balance displayed. Cues validities were: .9, .8, and .7, whereas discrimination rates were: .33, .4, and .47. The cue display order was randomised.

C. Models implementation

The Optimising Model was implemented in R whereas TTB and EAM were implemented in Lisp. Although in the multiple-cue forced-choice task subjects used discrete pieces of information, applying the probability matching strategy requires the use of a continuous version of function r. So instead of discrete values of cues' success (.632, .620, and .594), which provide values of r only in three points ($U(1)$, $U(2)$, and $U(3)$) the linear function of the success was used ($U(n) = -.019n + .6533$). The line was regression line of the three above-mentioned points. The model's prediction was generated using Function R (formula 3) with environment based values of costs and payoffs. The probability of revealing n cues was defined as $R(n\text{-}.5, n\text{+}.5)$. The threshold parameter in EAM was optimised analytically, whereas the s parameter in extended versions of TTB and EAM was optimised using the Nelder-Mead method. Both parameters were optimised in order to minimise root mean square deviation (RMSD).

D. Procedure

First, the subjects were asked to imagine that they were engineers looking for a deposit of some mineral. They were told that they should choose from two possible searching locations "A" and "B". They were also told that searching itself costs a given number of points and that they gain a specified number of points for making the correct choice. The subjects were informed that they could use a few tests to examine the locations and that these tests had a described price, validity and discrimination proportion. The costs of all the cues were equal and visible for subjects, but the validity and discrimination rate were hidden and different for all the cues. As the cues' properties were not provided to subjects explicitly, they had to try to discover all features of the cues that they needed during the training phase (either directly or as some features of the strategy that they found optimal). All the cues' features remained unchanged between training and test.

The task consisted of two parts: training and actual task. Performance in the training phase was not rewarded because it could prevent subject from exploring different strategies. Subjects began the training phase with 0 points and their goal was to collect as many points as possible within 75 trials. Points gained during the training phase were not included in the overall sum. After the training phase subjects went on to the actual task, which consisted of 55 trials, their goal was again to collect as many points as possible starting with 0 points. The reward depended on the number of points collected in the latter phase.

In the loss condition (experimental condition of the first experiment) subjects began with negative number of points (-450 in the training phase and -375 in the actual task) and their goal was to reach zero points (in no more than 150 trials). In the no feedback condition (experimental condition of the second experiment) the subjects received no feedback and they did not have the point balance displayed.

All subjects participated in both experiments described. In each of the experiments they were randomly divided into two equinumerous groups: control and experimental, and independently assigned to high or low cost condition. Thus, each subject performed one condition on one cost level in each of the two experiments (assignment to the manipulation and the cost were independent for each experiment). The cue order was randomised and did not reduplicate between experiments for any subject.

IV. RESULTS

All the analyses described below were performed with R language [53]. Since the model predicts a particular distribution of the number of cues revealed and because distribution carries more information than a mean number of revealed cues, a proportion test was used in all analyses. The analyses do not include the training phase results. In both experiments a statistically significant influence was found concerning the cost of revealing a cue on the number of cues revealed. In the first experiment the proportion for 0, 1, 2, and 3 cues was respectively: .02, .27, .32, and .39 for the .05 cost condition and .03, .31, .38, and .28 for the .1 cost condition ($\chi^2 = 110.19$, $p < .0001$, $df = 3$, $N = 1262$). In the second experiment the proportion was .06, .25, .38, and .32 for the low-cost condition and .03, .32, .33, and .32 for the high-cost condition ($\chi^2 = 44.19$, $p < .0001$, $df = 3$, $N = 1310$). In the first experiment the proportion of guessing (choosing an alternative with no cues revealed) was almost the same in both cost conditions (proportion test based 95% confidence intervals for .024: [.016, .034], and for .029: [.019, .038]). The difference between conditions was supported by the fact that in the high-cost condition participants revealed all cues less frequently (.28 in comparison to .39, 95% CI for .39: [.363, .417], and for .279: [.255, .304], they more often made decisions after revealing one or two cues. In the second experiment there was no difference between the proportions of revealing three cues, and a very small difference (although statistically significant) between proportions of guessing (95% CI for .056: [.044, .07], and for .027: [.019, .037]). Most of the loss of the revealed cues number in the high cost condition was caused by making a decision on the basis of one instead of two cues in that group.

There were significant differences between the proportions of cues revealed in the control and the experimental conditions in both experiments. During the first experiment, in the loss condition subjects more often ended exploration after revealing less cues than the optimal value, and in the gain condition they more often revealed more cues than the optimal value ($\chi^2 = 57.96$, $p < .0001$, $df = 1$, $N = 1242$). Similarly, in the second experiment, in the feedback condition subjects more often ended exploration after revealing less cues than the optimal value, and in the no-feedback condition they more often revealed more cues than the optimal value ($\chi^2 = 50.2$, $p < .0001$, $df = 1$, $N = 1312$, see Table III). The optimal number of cues revealed was 1.75 for the low cue cost condition, and 1.1 for the high cue cost condition.

The task was performed in three conditions (one control identical for both experiments and one experimental for each experiment). Since there are no theoretical reasons for considering any of the conditions to be basic or to be the primary version of the task, and the manipulation was supposed to change the number of revealed cues, the model's general predictions were compared with mean performance in all conditions. The model performance proves to match subjects' performance well (RMSD = .02, $r^2 = .98$), see the figure 2. In particular the model recreates the change in frequency of revealing one and three cues between two cost conditions as well as the relative constancy of frequency of revealing zero and two cues. In both model's predictions and subjects' behaviour, one cue was revealed more often in the high cost condition than in the low cost one. By contrast three cues were revealed more often in the high cost condition than in the low cost one.

TABLE III PROPORTION OF NUMBERS OF CUES REVEALED IN DIFFERENT CONDITIONS

Exper iment	Condition	Proportion of the number of cues revealed	
		0 or 1	**2 or 3**
I	Gain, cost = .05	.25	.75
	Gain, cost = .1	.29	.7

	Gain, indiscriminately	.27	.73
	Loss, cost = .05	.34	.66
	Loss, cost = .1	.4	.61
	Loss, indiscriminately	.37	.63
II	Feedback, cost = .05	.39	.61
	Feedback, cost = .1	.36	.64
	Feedback, indiscriminately	.37	.63
	No feedback, cost = .05	.22	.78
	No feedback, cost = .1	.34	.66
	No feedback, indiscriminately	.28	.72

Note: All differences are statistically significant (p < .001, with Benjamini and Yekutieli's correction).

The Optimising Model's match was compared to the match of two other models: TTB and EAM in both their basic and extended versions. For EAM, the threshold parameter was optimised for each cost condition. The EAM achieved the best match for the threshold in the range (0, 1.386], for both cost conditions, and its predictions of the proportions are the same as TTB's: 0, .33, .268, and .402 for 0, 1, 2, and 3 cues revealed. EAM's match is equal within the whole interval because as long as the threshold is not larger than ln $v^2/(1\text{-}v^2)$ (where v^2 is validity of second most valid cue), any discriminating cue exceeds the threshold or is the last available one–in both cases the choice is made after picking the cue. TTB's and EAM's RMSD and r^2 (RMSD = .07, r^2 = .81) turned out to be worse than the Optimising Model's one.

The extended versions of TTB and EAM achieved the best match for s = .048 (and for the EAM's threshold parameter remaining in the same range as for its basic version). Their match to data was slightly better than for their basic versions (RMSD = .05, r^2 = .83). Since the simplest model (The Optimising Model) clearly achieved a better match than the more complex ones, there is no further need to compare the measures which take into account their complexity, as it would not provide any additional information.

The figure 2 shows the TTB's and EAM's match to data (in the base versions of both models). Although both models predict the frequency of guessing (and frequency of revealing one cue in high cost condition), they fail to match the frequencies of revealing two or three cues. The models predict that three cues will be revealed more often than two, while subjects make a decision on the basis of two cues as often as on the basis of three ones (in low cost condition), or even more often (high cost condition).

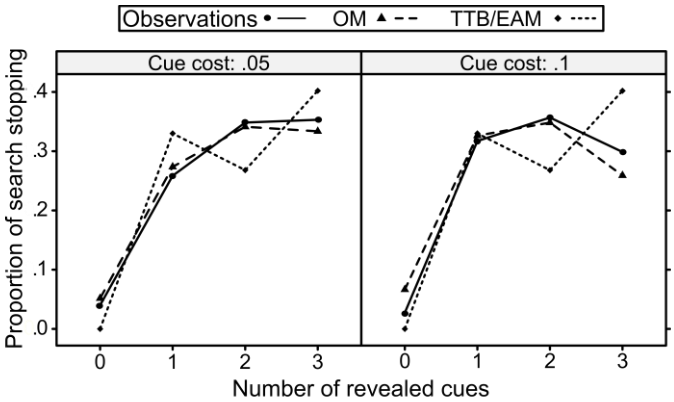

Fig. 2 Match between the models' predictions and subjects' performance

Contrary to the specific hypotheses concerning the mathematical model, it is difficult to confirm any strong hypotheses about the generality and optimality of human decision making by one series of experiments alone. However, the results show that people behave in the same way as a model that is designed to achieve an optimal solution to a decisional problem and is so general that it can recreate many strategies which are believed to be qualitatively different. Therefore, it can at least be said that the general hypotheses have been partially corroborated, and that the results are a good starting point to follow up studies which further explore these questions.

The rationality that underlies the proposed model has a specific meaning. It is neither the classical economic rationality which does not allow for cognitive biases and limitations, nor Simon's [7] bounded rationality which assumes that people are not able to efficiently progress towards an optimal solution. The rationality assumed in the present paper is defined in relation to an optimal solution of a given task and reflects the biases which are known (e.g., the probability matching rule) or expected (e.g., subgoals asymmetry depending on the gain-loss condition). These biases are neither random nor derived from limitations of human cognitive abilities. Although they may appear non-adaptive in certain environments, they are, in fact, adaptive in more general and common conditions.

The generality which is defined here as the capability of the model to recreate behaviour of many qualitatively different strategies by manipulation in some continuous parameter is not complete. The model is supposed to imitate most decisional strategies from the Adaptive Toolbox [6] and some strategies from outside of it (e.g., WADD), but it still remains within Newell's [33] continuum of strategies ordered with regard to information requirements. Therefore, if the presented results reveal anything valid about rationality and generality, they do not tell us that human decision making is absolutely rational and general, rather they show how general it is and what kind of rationality it involves. The assumptions about optimality and generality are not applied in the model independently. The model emulates many strategies with the use of one mechanism because it tries to behave optimally in different environments, and as long as the environments are substantially different, the strategies which are optimal in them are also different.

V. CONCLUSION

There were two groups of questions that the experiments had to answer. The first group considered the issues of rationality and the generality of human problem solving. Do people behave optimally while making a decision? Do they use one general strategy? Most importantly, can they use a strategy that is both general and optimal? The second group of questions considered the proposed model itself. Does the model describe human decision making behaviour accurately? Can the two opposing subprocesses which make up the suggested strategy be noticed in people's behaviour?

Next to the importance of the general findings, the presented results might be interesting in the context of this particular model. As stated earlier “this particular model” is actually an entire class of models which can be recreated by the Optimising Models depending on the environmental features.

The match between experimental data and the model’s prediction is good, especially considering the fact that the model has no free parameters, so its output values spring from a simple application of the postulated mechanisms to the features of the environment, not from a manipulation of parameters aimed at obtaining the expected results. The model also performs well in comparison to other models of decision making. The two models that it was compared with, TTB and EAM, are based on different theories. “Take The Best”, most commonly examined strategy from the Adaptive Toolbox, is being opposed to the “rational”, highly demanding strategies like WADD. The Evidence Accumulation Model is supposed to integrate two kinds of models: the “fast and frugal” as well as the “rational”. Both models perform worse than the Optimising Model despite their higher complexity (in their extended versions).

Moreover, since the Evidence Accumulation Model achieves the best match for a low threshold parameter, it performs identically to TTB. Therefore, the threshold parameter does not improve the fitness of the model and as a consequence the TTB, being simpler, should be considered the better of the two models in the domain of predicting the amount of information obtained in the tested environments.

The mere fact that the suggested model’s predictions match empirical data does not give reason to find this particular model true, as there could be other models that make the same predictions. Therefore, further hypotheses were presented that focused on the postulated mechanisms of the model. It was assumed that if there are two opposing goals of the decision process (maximising the probability of making the correct choice and minimising the costs), they can be influenced separately under some conditions. The results confirm these hypotheses: along with some controlled changes in the features of a decisional environment (namely: providing or withholding feedback and affecting the subjects’ sense of loss or gain) people focused accordingly on one of the goals mentioned above, which in turn had an influence on the number of cues revealed.

It requires further research to uncover the specific impact that different factors have on the disproportion between validity of the subgoals and to provide a detailed description of the mechanism in question. But, before addressing the problem of the exact mechanism used in decision generating, it is necessary to answer two preliminary questions about what, on the one hand, decision making itself is, and, on the other hand, how on a general level more specific questions can be raised. The results presented allow the supposition that, firstly, the decision making consists of optimising behaviour in order to maximise the reward value, and secondly, that hypotheses concerning decision making can be formed on the most general level, and for various decisional behaviours seem to emerge from one process of maximising the difference between gains and costs.

Although the Optimising Model has been confirmed by empirical results, both by achieving a good match to data and by recreating qualitative phenomena, more research must be done to fully confirm the general hypotheses concerning the optimality and rationality of decision making. To unmistakably show its generality, the model must be tested in many different environments, which differ not only with regard to costs but also to cost increment function or to the relation between validity and discrimination rate. Also the hypothesis concerning the optimality of human behaviour awaits broader verification as well as theoretical analysis. Finally, further research could reveal the particular structure of the decisional mechanism, for both of the discovered subprocesses are likely to consist of multiple components which affect their working in the same way as the subprocesses themselves affect the general decision making process.

REFERENCES

[1] B. Pascal. *Pensees*. Forgotten Books, Charleston, 2008.

[2] J. von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton University Press, Princeton, 1953.

[3] L. J. Savage. *The foundations of statistics*. Dover Publications, New York, 1972.

[4] D. Kahneman and A. Tversky. “Prospect theory: An analysis of decision under risk”. *Econometrica*, vol. 47, pp. 263-292, 1979.

[5] B. Paulewicz. *Interaction and adaptation - a proposition in behavioural science* [*Interakcja i adaptacja - propozycja metateoretyczna w badaniach nad zachowaniem*]. PhD thesis, Krakow: Jagiellonian University, unpublished.

[6] G. Gigerenzer and R. Selten. *Bounded Rationality. The Adaptive Toolbox*. MIT Press, Cambridge, MA, 2001.

[7] H. A. Simon. *Models of Bounded Rationality. Vol. 3*. MIT Press, Cambridge, MA, 1997.

[8] G. Gigerenzer, P. M. Todd, and the ABC Research Group. *Simple heuristics that make us smart*. Oxford University Press, New York, 1999.

[9] A. Glöckner and T. Betsch. “Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making”. *Judgment and Decision Making*, vol. 3, pp. 215-228, 2008.

[10] W. Schneider and R. M. Shiffrin. “Controlled and automatic human information processing: I. detection, search, and attention”. *Psychological Review*, vol. 84, pp. 1-66, 1977.

[11] P. M. Todd and A. Dieckmann. “Heuristics for ordering cue search in decision making”. In L. K. Saul, Y. Weiss, and L. Bottou, editors, *Advances in neural information processing systems 17*, pp. 1393-1400. MIT Press, Cambridge, MA, 2005.

[12] T. Smolen and S. Wichary. “Bottom-up model of strategy selection”. In B. C. Love, K. McRae, and V. M. Sloutsky, editors, *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, pp. 1517-1521. Cognitive Science Society, Austin, TX, 2008.

[13] A. Bröder. “Assessing the empirical validity of the «take-the-best» heuristic as a model of human probabilistic inference”.

Journal of Experimental Psychology: Learning, Memory, and Cognition, vol. 26, pp. 1332-1346, 2000.

[14] B. R. Newell, N. J. Weston, and D. R. Shanks. “Empirical tests of a fast-and-frugal heuristic: Not everyone «takes-the-best»”. *Organizational Behavior and Human Decision Processes*, vol. 91, pp. 82-96, 2003.

[15] B. R. Newell and D. Fernandez. “On the binary quality of recognition and the inconsequentiality of further knowledge: Two critical tests of the recognition heuristic”. *Journal of Behavioral Decision Making*, vol. 19, pp. 333-346, 2006.

[16] D. Hausmann and D. Läge. “Sequential evidence accumulation in decision making: The individual desired level of confidence can explain the extent of information acquisition”. *Judgment and Decision Making*, vol. 3, pp. 229-243, 2008.

[17] H. A. Simon. *Models of man*. Wiley, New York, 1957.

[18] C. Cherniak. Rational “Agency”. In Robert A. Wilson and Frank C. Keil, editors, *The MIT Encyclopedia of the Cognitive Sciences*, pp. 698- 699. MIT Press, Cambridge, MA, 1999.

[19] N. Vulkan. “An economist’s perspective on probability matching”. *Journal of economic surveys*, vol. 14, pp. 101-118, 2000.

[20] L. G. Humphreys. “Acquisition and extinction of verbal expectations in a situation analogous to conditioning”. *Journal of Experimental Psychology*, vol. 25, pp. 294-301, 1939.

[21] M. E. Bitterman. “Phyletic differences in learning”. *American Psychologist*, 20, pp. 396-410, 1965.

[22] D. G. C. Harper. “Competitive foraging in mallards: `ideal free’ ducks”. *Animal Behaviour*, vol. 30, pp. 575-584, 1982.

[23] F. Thuijsman, B. Peleg, M. Amitai, and A. Shmida. “Automata, matching and foraging behavior of bees”. *Journal of Theoretical Biology*, vol. 175, pp. 305- 316, 1995.

[24] W. Gaissmaier, L. J. Schooler, and J. Rieskamp. “Simple predictions fueled by capacity limitations: When are they successful?” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 32, pp. 966-982, 2006.

[25] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998.

[26] N. D. Daw, J. P. O’Doherty, P. Dayan, B. Seymour, and R. J. Dolan. “Cortical substrates for exploratory decisions in humans”. *Nature*, vol. 441(15), pp. 876-879, 2006.

[27] Y. Niv, D. Joel, I. Meilijson, and E. Ruppin. “Evolution of reinforcement learning in uncertain environments: A simple explanation for complex foraging behaviors”. *Adaptive Behavior*, vol. 10, pp. 5-24, 2002.

[28] A. Bröder and S. Schiffer. “Adaptive flexibility and maladaptive routines in selecting fast and frugal decision strategies”. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 32, pp. 904- 918, 2006.

[29] B. R. Newell and M. D. Lee. “Learning to adapt evidence thresholds in decision making”. In N. A. Taatgen & H. van Rijn, editor, *Proceedings of the 31th Annual Conference of the Cognitive Science Society*, pp. 473- 478. Cognitive Science Society, Austin, TX, 2009.

[30] J. W. Payne, J. R. Bettman, and E. J. Johnson. “Adaptive strategy selection in decision making”. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 14, pp. 534-552, 1988.

[31] A. Bröder and W. Gaissmaier. “Sequential processing of cues in memory-based multiattribute decisions”. *Psychonomic Bulletin & Review*, vol. 14, pp. 895-900, 2007.

[32] B. R. Newell, P. Collins, and M. D. Lee. “Adjusting the spanner: testing an evidence accumulation model of decision making”. In D. McNamara and G. Trafton, editors, *Proceedings of the 29th Annual Conference of the Cognitive Science Society*, pp. 533-538. Erlbaum, Mahwah, NJ, 2007.

[33] B. R. Newell. “Re-visions of rationality?” *Trends in Cognitive Sciences*, vol. 9, pp. 11-15, 2005.

[34] M. D. Lee and T. D. R. Cummins. “Evidence accumulation in decision making: Unifying the «take the best» and the «rational» models”. *Psychonomic Bulletin and Review*, vol. 11, pp. 343-352, 2004.

[35] S. Link and R. Heath. “A sequential theory of psychological discrimination”. *Psychometrika*, 40, pp. 77-105, 1975.

[36] M. Stone. “Models for choice-reaction time”. *Psychometrika*, vol. 25, pp. 251- 260, 1960.

[37] R. Ratcliff. “A theory of memory retrieval”. *Psychological Review*, vol. 85, pp. 59-108, 1978.

[38] J. Rieskamp and P. E. Otto. “SSL: A theory of how people learn to select strategies”. *Journal of Experimental Psychology: General*, vol. 135, pp. 207-236, 2006.

[39] W. Edwards and P. Slovic. “Seeking information to reduce the risk of decisions”. *American Journal of Psychology*, vol. 78, pp. 188-197, 1965.

[40] L. S. Fried and C. R. Peterson. “Information seeking: optional versus fixed stopping”. *Journal of Experimental Psychology*, vol. 80, pp. 525-529, 1969.

[41] G. F. Pitz. “The influence of prior probabilities on information seeking and decision making”. *Organizational Behavior and Human Performance*, vol. 4, pp. 213-226, 1969.

[42] G. F. Pitz, Helen Reinhold, and E. Scott Geller. “Strategies of information seeking in deferred decision making”. *Organizational Behavior and Human Performance*, vol. 4, pp. 1-19, 1969.

[43] A. Rapoport. “Sequential decision-making in a computer-controlled task”. *Journal of Mathematical Psychology*, vol. 1, pp. 351-374, 1964.

[44] W. Edwards. “Optimal strategies for seeking information”. *Journal of Mathematical Psychology*, vol. 2, pp. 312-329, 1965.

[45] B. R. Newell, Tim Rakow, Nicola J. Weston, and David R. Shanks. “Search strategies in decision making: The success of «success»”. *Journal of Behavioral Decision Making*, vol. 17, pp. 117-137, 2004.

[46] T. Rakow, B. R. Newell, K. Fayers, and M. Hersby. “Evaluating three criteria for establishing cue-search hierarchies in inferential judgment”. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 31, pp. 1088-1104, 2005.

[47] A. Glöckner and T. Betsch. “Multiple-reason decision making based on automatic processing”. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 34, pp. 1055-1075, 2008.

[48] A. Bröder and S. Schiffer. “Stimulus format and working memory in fast and frugal strategy selection”. *Journal of Behavioral Decision Making*, vol. 19, pp. 361-380, 2006.

[49] L. Martignon and U. Hoffrage. “Fast, frugal, and fit: simple heuristics for paired comparison”. *Theory and Decision*, vol. 52, pp. 29-71, 2002.

[50] B. R. Newell and D. R. Shanks. “Take the best or look at the rest? Factors influencing «one-Reason» decision making”. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 29, pp. 53-65, 2003.

[51] A. Bröder. “Decision making with the «adaptive toolbox»: Influence of environmental structure, intelligence, and

working memory load”. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 29, pp. 611- 625, 2003.

[52] R. E. Kass and A. E. Raftery. “Bayes factors”. *Journal of the American Statistical Association*, 90(430), pp. 773-795, 1995.

[53] R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2009.