

# Sequential evidence accumulation in decision making: The individual desired level of confidence can explain the extent of information acquisition

Daniel Hausmann\*  
General Psychology (Cognition)  
University of Zurich

Damian Läge  
Applied Cognitive Psychology  
University of Zurich

## Abstract

Judgments and decisions under uncertainty are frequently linked to a prior sequential search for relevant information. In such cases, the subject has to decide when to stop the search for information. Evidence accumulation models from social and cognitive psychology assume an active and sequential information search until enough evidence has been accumulated to pass a decision threshold. In line with such theories, we conceptualize the evidence threshold as the “desired level of confidence” (DLC) of a person. This model is tested against a fixed stopping rule (one-reason decision making) and against the class of multi-attribute information integrating models. A series of experiments using an information board for horse race betting demonstrates an advantage of the proposed model by measuring the individual DLC of each subject and confirming its correctness in two separate stages. In addition to a better understanding of the stopping rule (within the narrow framework of simple heuristics), the results indicate that individual aspiration levels might be a relevant factor when modelling decision making by task analysis of statistical environments.

Keywords: evidence accumulation, sequential information search, information acquisition, threshold models, stopping rule, level of confidence, probabilistic cue, validity, one-reason decision making.

## 1 Introduction

“Some use it sometimes.”

This is the short statement by which Newell (2005) concluded the empirical research on the well-known Take-the-Best heuristic (TTB). This fast and frugal heuristic was introduced by Gigerenzer, Hoffrage, and Kleinbölting (1991) and Gigerenzer and Goldstein (1996) to provide an alternative to traditionally proposed strategies in judgment and decision making research. According to these traditional models, a rational decision requires a comprehensive integration of all available information about all alternatives. Such integration can become computationally complex as soon as information cues are probabilistic and numerous. Gigerenzer and Goldstein (1996), however, were able to show by a simulation of a binary attribute cue paradigm that choosing the best alternative (out of two) can be very simple when searching the available cues in the sequential order of descending validity (the probability that a cue will lead to the correct decision, given that it discriminates between the alternatives) and deciding on the direction

of the first discriminating cue. In the stimulus environment they used (German cities, their population size and a number of features being present in the cities or not), this simple three-step-heuristic (containing a search rule, a stopping rule and a decision rule) led to decisions of the same quality as a comprehensive computational integration of the entire set of cue information. The termination of information search after having found a first discriminating cue has therefore been called one-reason decision making (ORDM) (Gigerenzer & Goldstein, 1999).

The term ORDM was chosen in clear contrast to models that stand for the integration of all available information (usually not reflecting how the information was compiled). These models can be of varying complexity, ranging from simple tallying to complex multiple linear regression (for an overview, see Payne, Bettman & Johnson, 1993, or Doyle, 1999. Lee & Cummins, 2004, group them as “rational models”). What they have in common is that at least two pieces of information need to be present for starting a comparison and, in case of several cues indicating the same option, an integration of the available information. In contrast to the class of ORDM, we subsume this broad class under the term “more-reason decision making” (MRDM), because the search for more than one discriminating piece of information is a prerequisite

---

\* Address: Daniel Hausmann, Psychologisches Institut der Universität Zürich, Allgemeine Psychologie (Kognition), Binzmühlestrasse 14 / Box 22, CH-8050 Zürich. E-mail: d.hausmann@psychologie.uzh.ch.

for information integration.

TTB, as the most prominent representative of ORDM, reveals an impressively high number of correct predictions, compared to MRDM models (see Gigerenzer, Czerlinski & Martignon, 1999, for simulation studies). These mathematical results allow TTB to rank amongst the rational strategies to be used in probabilistic decision making, especially (but not only) when information needs to be searched for.

However, the empirical findings concerning the use of TTB — compared to MRDM models — were heterogeneous. While the conclusions of some studies were generally against the use of TTB because they found only particular use of this strategy (Bröder, 2000; Lee & Cummins, 2004; Newell & Shanks, 2003; Newell, Weston & Shanks, 2003), a survey of more recent studies shows more evidence for the use of TTB (Bröder & Schiffer, 2003; see the overview in Bröder, 2005). One unresolved question is the fact that participants repeatedly showed a huge inter- and intra-individual inconsistency in using a single strategy (Läge & Daub, 2006). So far, there have been no statistical environments in which all participants of an experiment acted according to TTB or refrained from using it. But what is even more problematic is the inconsistent use of the ORDM stopping rule. The most surprising effect was reported by Newell, Weston and Shanks (2003): Testing a sequential search paradigm using only two available cues, they found a number of participants (32% of their sample) continuing information search in some instances even when the more valid cue discriminated, hence completely ignoring the fact that the second cue would not be able to compensate the first information! As an overall judgment, the conclusion drawn by Newell still holds: “Some of the people made choices consistent with TTB some of the time” (Newell, 2005, p. 12).

### 1.0.1 Modelling the information acquisition process under uncertainty

The order in which probabilistic cues are to be searched depends on the individual weight of the cue parameters (at least cue validity and discrimination rate, i.e. the proportion of occasions on which a cue value is different for two objects in a two alternative comparison task; see Martignon & Hoffrage, 1999; Newell, Rakow, Weston & Shanks, 2004; Dieckmann & Todd, 2005; Läge, Hausmann, Christen & Daub, 2005; Todd & Dieckmann, 2005). Whenever the discrimination rate is not relevant, only two search strategies are reasonable: a cue search in descending order of validities (as specified in TTB) or a random selection strategy of cues which has been called the Minimalist search rule (Martignon & Hoffrage, 1999). The decision maker’s knowledge about cue validi-

ties determines which of these two search rules is adaptive: TTB search rule is normatively required when cue validities are known, whereas Minimalist search rule is the only possible search strategy when cue validities are unknown.

In both cases, the crucial question is about the stopping rule: at which point do people stop their search for information? When focussing on the decision rules (ORDM versus information integration in MRDM), the answer should be related to the general strategy of decision making that an individual prefers (one-reason stopping after the first discriminating cue versus more-reason stopping depending on the specific integration model). Focussing on the information acquisition process, however, an alternative answer would be that people will (ideally) terminate search when they have subjectively collected enough information, which means they have enough evidence to determine which of the options is the right or the best one.

This assumption can be based on experimental findings that show a very clear relation between stopping behaviour and estimated cue validity in a learning experiment conducted by Hausmann, Christen, and Läge (2005): over a learning period of 50 trials, people observed the ratio of correct/incorrect answers from four experts who gave advice in a quiz show. In a second stage, they became participants of that quiz show and were able to seek advice from one or more of these experts. As to be expected, people started their search with the expert they had perceived to give the most valid answers. The stopping rule, however, very much depended on the perceived validity of this best adviser: the higher the perceived validity was, the higher the tendency towards ORDM.

### 1.0.2 Models with an explicit level of confidence as a stopping rule

Approaches that place the stopping rule in the foreground and make it dependent on the quality of the searched cue information have been called evidence accumulation, or evidence accrual, or sequential-sampling process models, for they consist of an evidence threshold, the reaching or exceeding of which results in the termination of the information search (for a list of specific models according to this class of approaches, see Lee & Cummins, 2004, p. 346; Rouder, 2001, p. 335; or Leth-Steensen & Marley, 2000, p. 65). The direction of the decision (the preferred option) has already crystallized and the effective choice (the decision rule) therefore remains trivial.

The “satisficing principle” of Herbert Simon (1955; 1956) was an early threshold model in decision making theories, intended as a direct rejection of the a priori normative view provided by the utility maximization principle of Subjective Expected Utility theory. Rather than maximizing the overall utility of all available options, Si-

mon postulated that people search for options only until the first one met or exceeded an aspiration level. Whereas Simon had focused on the search for options and the evaluation of their overall utilities, one can assume a “sufficiency” principle for the search for probability cue information when options are given: how much cue information do people need to make a good or correct decision? People perhaps want to find out not only what the best decision is, but they also want to reach a certain level of confidence before making their final decision.

Lee and Cummins (2004) described such a sequential-sampling process as a random walk, acquiring probabilistic cue information step-by-step either until an evidence threshold is reached or the remaining information cannot outperform the current evidence gathered for the best option. By that way they construct a unifying model for ORDM and MRDM, integrating both as special cases given that the individual evidence threshold is reached immediately (ORDM) or not (MRDM) during the information acquisition process.

An evidence (or confidence) threshold can be viewed in two ways: either as a relative threshold (value difference) or as an absolute threshold (level of confidence). On the one hand, some of the approaches of evidence accumulation models dealing with a sufficiency-principle expect termination of information search as soon as sequentially accumulated value differences of cues have reached or exceeded a certain threshold. The searching process “is stopped and an alternative is chosen when a person has accumulated enough evidence to be convinced that one alternative is better than the other” (Aschenbrenner, Albert & Schmalhofer, 1984, p. 154). The criterion dependent choice (CDC) model for binary choices from Aschenbrenner et al. (1984) has until now predicted choices in various domains better than, for example, an additive model. On the other hand, some evidence accumulation models indicates stopping as soon as the desired confidence level of an option has reached, or exceeded, a certain threshold. A sufficiency principle, for example, is formulated in the heuristic-systematic model by Chaiken, Liberman, and Eagly: “People will exert whatever level of effort is required to attain a sufficient degree of confidence...” (1989, p. 221). Processing efforts are assumed to be a function of the discrepancy that exists between actual and desired levels of confidence (Eagly & Chaiken, 1993).<sup>1</sup>

<sup>1</sup>To distinguish, the paper deals with three quantitative concepts: validity, evidence, and confidence. Validity describes the predictive quality of a cue, whilst evidence and confidence are related to the alternatives. Evidence (as the probabilistic quality of an alternative) is a rather cognitive term, whereas confidence describes the emotional state of a person (etymologically referring to social relations). “Confidence ratings” are a common measure in decision making to research the degree of certainty a person has that a chosen alternative is correct. This can be modelled either as discriminating difference between alternatives (for

### 1.0.3 Research question

According to an evidence accumulation model, the decision to stop or to continue information acquisition depends on the degree of certainty for the preferred option, and can therefore be very inconsistent in terms of the number of cues which need to be acquired: one discriminating cue can be sufficient when it provides a person with a high degree of evidence. However, when the validity is not high enough, the person prefers to continue information search. Different from ORDM (always stopping after the first discriminating cue) and MRDM (needing at least two discrimination cues), the stopping rule would be a certain degree of confidence (a threshold which can vary from individual to individual and may also depend on the accessibility of cues).

Hence, it is the aim of this paper to propose a validity threshold as a stopping rule and to test in two experiments whether such a threshold is better able to predict the individual behaviour than a strict one-reason or more-reason stopping rule. We assume that a threshold, like a desired level of confidence (DLC), is the effective stopping rule that people follow when searching for probability cues. As soon as the actual accumulated evidence (= current confidence) has reached or exceeded the DLC, people will stop further searching and decide on the direction of the preferred option.

We present two experiments to measure and verify a DLC in a multi-attribute probabilistic decision task with several options. To be able to focus on the stopping rule, the order of probabilistic cues to look at becomes a defined aspect of the statistical environments of these experiments: search rules will be fixed either as descending validity (TTB environment) or in random order of validity (minimalist environment). By fixing the order of cues, search rules become obsolete in that the only decision to be made during the information acquisition process is when to stop. For the experiments we will therefore talk of “information acquisition” whenever there is no choice of which cue to look at next and of “search” theoretically in a more general way or when there is free choice in choosing cue information.

example as the difference between the probabilities of the best two options to be the correct one) or as the probability of the best option. The latter model is chosen to be used for this paper because it directly refers to the degree of evidence a person has accumulated during the information acquisition process. (In case of more than two alternatives, this seems to be important because a difference model could lead to a high value, but on a very low level of evidence.) Hence, confidence and evidence are used in a similar manner because they are both expressed by the same quantitative measure.

## 2 Experiments

### 2.1 Experiment 1: Measuring and verifying a confidence threshold<sup>2</sup>

#### 2.1.1 Empirical evidence for the desired level of confidence (DLC).

The fundamental idea in the empirical testing of a judgment confidence threshold consisted in measuring the stopping behaviour of people when confronted with a first discriminating cue. A consistent ORDM-stopping rule is given when people always terminate the search for information after uncovering the first piece of discriminating information, independent of the cue validity. On the other hand, with an MRDM-stopping rule, one will never be satisfied with a single individual discriminating cue (independent of the cue-validities). People will always search for at least two discriminating cues in order to integrate these into their final decision. A consequent application of a judgment confidence threshold with a certain desired level of confidence (DLC) indeed looks very different: the first discriminating cue is examined according to its validity. If the validity of this cue corresponds to, or even lies above, the DLC, the search is immediately stopped and one will decide for that option based on the direction in which the cue points (the observed behaviour corresponds to an ORDM-stopping rule strategy). If, on the other hand, the validity of the first discriminating cue is lower than the DLC, the search will be continued (the observed behaviour tends towards an MRDM-stopping rule strategy). In order to examine a confidence threshold, it is sufficient to measure the stopping behaviour of a person dependent on the validity of the first discriminating cue by checking whether this information is accepted for making an immediate decision, or whether the search is continued (by looking at the next cue).

In order to keep the method of measurement as simple as possible, the discrimination rates of the cues in the experiments were set at the maximum of 1.00: every uncovered cue provided useful — discriminating — information. The maximum discrimination rate was achieved by using the “personified cues” proposed by Läge, Hausmann, Christen, and Daub (2005). These cues are direct predictions given by persons with a certain amount of expertise, indicating which of the options they would chose. Whilst the concept of attribute cues (like in the Gigerenzer & Goldstein, 1996, set of cities and their attributes) leads to a limited discrimination rate, the personified cues always discriminate as long as the cue predicts exactly one option as being the best one. Both cue classes can be used in the same way for information board designs be-

<sup>2</sup>Data from this experiment are presented in the proceedings book of the German Cognitive Science Conference (Hausmann & Läge, 2005).

cause they lead to identical structures of binary cue values. (Since the experimental design as described below deals with four options per trial, the attribution cue concept would lead to many non-discriminating cues making the experiment long-lasting and does not support efficient research on the stopping rule.)

Acquisition costs were introduced in order to avoid participants being able to uncover all the cues. The acquisition costs per cue in relation to the potential profit increase were fixed at a ratio of 1:10, because earlier investigations have shown that this ratio evokes a high proportion of ORDM (Hausmann, Christen & Läge, 2006).

#### 2.1.2 Overcoming the measuring problem

The evidence of a judgment confidence threshold was examined in two separate stages in the same experimental session. The trials in Stage 1 were used for individually fitting the “desired level of confidence” parameter. The obtained value was then tested in Stage 2 by independent trials using the same set of data. (Participants were not informed about this fitting/testing separation nor did they know that three stopping rules were examined.)

The easiest way to measure a “desired level of confidence” is to check whether a person is satisfied with a piece of information of certain validity. This satisfaction can be documented by the stop of information acquisition, whilst dissatisfaction would lead to a continued search for information. Hence, we used this stopping behaviour as a predictor of an individual DLC by providing people with cues of differing validities and observing which cues led to their satisfaction and which did not. The DLCs were solely fitted from the first cues observed in each trial (preventing us from assuming certain models of information integration). Whenever people stopped further information acquisition after the first cue, the DLC was assumed to be lower than its validity. Whenever an individual continued information acquisition, the DLC was assumed to be higher. Testing the range of possible validities, the individual “changing point” could be detected.

Should an individual be inconsistent in his or her stopping behaviour (accepting a lower validity for stopping than the highest “rejected” validity — and vice versa), the optimal threshold needs to be calculated; this was done by minimizing the sum of squared errors.<sup>3</sup>

<sup>3</sup>The following trials of a subject were counted as errors a) where the validity of the first uncovered cue above an assumed threshold was not accepted to terminate information acquisition, and b) any validity below this threshold that was accepted to stop acquisition after the first cue. The magnitude of each error was then calculated as the difference between threshold level and cue validity. The sum of the squared errors determines the degree of appropriateness of this assumed level as an estimate of the DLC of that individual. That individual threshold was finally chosen for which the sum of these squared differences was

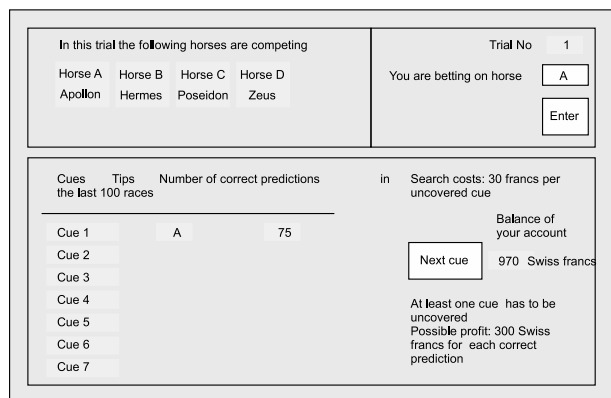


Figure 1: A typical judgment trial on the information board with four options (horses A to D). In this example, a cue with a validity of .75 (indicated as number of correct predictions in the last 100 races) was first uncovered. The participant stopped looking for further cues and decided for horse A (in which direction the cue pointed). In general, search costs (CHF 30 per uncovered cue) were subtracted independent of a correct prediction (CHF 300).

After having completed Stage 1, the best fit was automatically calculated for each participant. This value was taken as hypothetical DLC to be tested by independent data conducted in Stage 2 of the experiment: it contained 20 trials which were individually compiled from a list of trials starting with all different validities. Half of the trials covered the range between .03 and .21 above the individual DLC, the other half were below that value (same range between .03 and .21). Assuming the DLC measured in Stage 1 to be correct, the participant should stop information acquisition in exactly those 10 trials which were above this value immediately after having seen the first cue, whilst he or she should continue information acquisition in exactly the other 10 trials. (An ORDM model, in contrast, would assume that the participant stops in all 20 trials after the first cue, whilst all MRDM models predict a continuing information acquisition in all 20 trials.) These clear predictions can be tested for each individual by a simple binomial test for the 20 trials.

### 2.1.3 The experimental paradigm: the horse race

In order to measure stopping behaviour, we chose an attractive information board (see Figure 1) on which participants could uncover cue information actively and sequentially. As a cover story, a fictitious horse race was introduced allowing people to bet on the horse they thought had the best chances of winning (based on the cues provided). In each trial, four new horses were presented randomly (assembled out of a pool of a total of 125 names minimal.

of Greek, Roman, Egyptian, Celtic and Germanic gods). There was no prior knowledge of the winning probabilities (base rate = .25). The external cues served as an information basis to increase the probability of predicting the winning horse in the current trial. These cues were introduced to the participants in terms of personified insider-knowledge cues. Every uncovered cue pointed to one of the four horses (A, B, C or D) with a certain level of probability. The degree of expertise (cue validity) was displayed as the number of correct predictions the personified cue had made in the last 100 races (possible values were randomly distributed between 0 and 100). Thus, the number of correct predictions corresponds to an observable measure in terms of frequency of correct and incorrect tips of a cue. Each trial contained a new set of personified insider-knowledge cues so that participants could not calculate, conclude or transfer any validity from previous trials.

A maximum of seven cues were available in each trial, and these could only be uncovered sequentially in the order provided (top to bottom) (see Figure 1). At least one cue had to be uncovered in order to guarantee an evaluation with the first value. The number of cues the participants wanted to uncover was left to their discretion. The information acquisition was followed by the decision concerning the expected winning horse (A, B, C or D). As soon as a person had bet on a certain horse he or she received feedback as to whether this decision turned out to be correct or not. The only other feedback consisted of displaying the updated account: A correct prediction resulted in winning 300 Swiss francs (CHF), which was added to the participant's gambling account, minus the spent acquisition costs (CHF 30 per uncovered cue). A wrong prediction resulted in a commensurate loss, because the invested acquisition costs were deducted from the gambling account. Participants were urged to maximize their gambling account by making as many correct predictions as possible and collecting (fictitious) money. The motivation of the participants was additionally increased by the fact that the 10 participants with the highest total score entered a real lottery in which three lots of CHF 100 were drawn.

Twenty-two participants were given a total of 152 forced-choice decision trials, each consisting of four options in the form of different horses. The 152 trials included one practice trial, 51 trials for estimating the hypothetical level (Stage 1), while 20 tested the calculated level of confidence (Stage 2); (the remaining 80 trials in Stage 2 served to investigate another research question, namely varying the number of options; for results see Hausmann & Läge, 2005). All participants were faced with the same trials in Stage 1 (but in different random orders), whilst the selection of trials for Stage 2 was adaptive to the individual DLC. In Stage 2, the 20 testing trials

were randomly distributed within the set of 100 trials.

### 2.1.4 Results

The individually-calculated hypothetical DLCs in Stage 1 were situated between .52 and .88 ( $M = .73$ ,  $SD = .09$ ). The minimum squared sum of errors was between .00 and .11 ( $M = .03$ ,  $SD = .03$ ), meaning that some individuals strictly followed a certain DLC (error .00), while others showed more variability in their stopping behaviour.

The individually-calculated DLC was examined in Stage 2 with 20 comparable trials. The expectation of the error distributions in accordance with three decision models (DLC, ORDM and MRDM) and one random model (expecting a uniform distribution of correct predictions and errors) is presented in Figure 2. Out of the 440 trials (20 trials for 22 participants), the DLC model predicted 400 trials correctly (91%). The superiority to the other models is evident: ORDM correctly predicted only 214 trials (49%), MRDM 226 trials (51%). Therefore, no more than 9% errors clearly speak in favour of a preferred and consistent use of a confidence threshold model (DLC) at a group level.

On an individual level, each single participant can be classified as a user of his or her confidence threshold (binomial test,  $\alpha = .05$ , which means that at least 15 out of the 20 trials had to follow the DLC prediction). The same tests for the alternative models (ORDM and MRDM) failed for all participants: only in the case of one participant could it not be ruled out that she followed an inconsistent MRDM-stopping rule (15 hits out of the 20 trials for both, MRDM prediction and DLC prediction). In conclusion, people behave significantly according to the measured DLC.

## 2.2 Experiment 2: Stopping rules in different environments

The empirical testing of a desired level of confidence (DLC) in the above experiment deals with environments in which the information cues occur in an unforeseeable random order (unstructured cue order according to their validity). As explained in the introduction, the tested environment resembles the search rule of the Minimalist heuristic. Hence, the next step is to examine whether the existence of the DLC can be generalized to environments in which cues appear in a descending order of their validities (structured cue order), just as specified for TTB. If people were to know that the first available cue information is also the most valid one, would they maintain a level of confidence, or would they switch to an exclusive ORDM-stopping rule strategy? ORDM seems plausible because a second piece of information can never overrule the most valid cue. Therefore, people would have

to be prepared to continue the information search if inconsistent information should arise. Hence, TTB — as a representative of ORDM — should be a very reasonable strategy in such a scenario. In terms of the DLC model, however, one would expect that in some instances (when the validity of the first cue exceeds the individual DLC) people would stop their information search after the first cue, while continuing in all those cases where the validity of the first cue is lower than their DLC. The reason for continuing information search would not primarily be to overrule the first cue, since more cues may point in a different direction, but to gain more confidence before making the decision. Considering such a motivation for continued information search, even stopping after the second cue (if supporting the first cue) would be considered a “win” situation by the participant and not as a loss of money. Therefore, obtaining more evidence would be a driving force in information search. The fact that we confront the same person with both environments (structured versus unstructured cues) allows a direct testing of the stopping behaviour in a TTB-heuristic “friendly” environment.

### 2.2.1 Participants

Thirty-six participants (a majority of [mainly undergraduate] students from the Philosophy Faculty of the University of Zurich and several employed individuals) took part in the experiment. Twenty of the participants were female and 16 male. The average age was 29.2 years (range 18–44,  $SD = 5.1$ ).

### 2.2.2 Materials and procedure

The experiment was programmed with Microsoft Visual Basic 6 and run on IBM-compatible laptops. The program interface and the horse race scenario were taken from Experiment 1. This time, the participants were given a total of 60 forced-choice decision trials, each of which again consisted of four options. Mean validity of the best cue in each trial was at .75 ( $SD = .12$ ), the least best cue had a validity of .57. Once more, a maximum of seven cues was available in each trial and they could only be uncovered sequentially in the order provided, while at least one cue had to be uncovered so as to guarantee an evaluation with the first value. Provided with the cue validities of the uncovered cue(s), the participants were required to bet on the horse with the best chances of winning. Acquisition costs per cue and the possible winning amount remained identical to Experiment 1 (only the displayed currency on the information board was changed from Swiss francs to Euro).

Models and their predictions			
A pure ORDM-Stopping model		A pure MRDM-Stopping model	
	<i>ORDM</i>	<i>MRDM</i>	
$V_1 > DLC$	10	0	$V_1 > DLC$
$V_1 < DLC$	10	0	$V_1 < DLC$
Stopping by chance (random model)		Confidence threshold (DLC)	
	<i>ORDM</i>	<i>MRDM</i>	
$V_1 > DLC$	5	5	$V_1 > DLC$
$V_1 < DLC$	5	5	$V_1 < DLC$
Results			
	<i>ORDM</i>	<i>MRDM</i>	
$V_1 > DLC$	9	1 <sup>2</sup>	
$V_1 < DLC$	1 <sup>1</sup>	9	

Figure 2. The four examined models and their predictions. A pure ORDM-model would predict stopping after the first uncovered cue (= ORDM-stopping behaviour) for all testing trials per person ( $N = 20$ ) independent of the degree of the validity of the first uncovered cue. A pure MRDM-model would predict a further search (no stopping after the first cue) for all trials (= MRDM-stopping behaviour). A threshold model with an individual desired level of confidence (DLC) would predict a stopping of information search as a function of the degree of validities: showing an ORDM-stopping behaviour if the validity of the first uncovered cues reaches or exceeds the DLC and correspondingly a MRDM-stopping if the validity is lower than the DLC. In comparison, a random model would result in an unspecific stopping behaviour.

Results (on the third line) support the threshold model with only 4% ( $M = 0.77$ ,  $SD = 1.11$ ) errors of Type 1 (1<sup>1</sup> first cue validity wrongly accepted) and 5% ( $M = 1.05$ ,  $SD = 1.46$ ) errors of Type 2 (2<sup>2</sup> first cue validity wrongly ignored), with a total of 22 participants.

### 2.2.3 Design

We varied one two-level within-subjects factor, namely the experimental condition of unstructured versus structured cue orders. Each condition consisted of 30 trials. (The first trial served as a practice trial and was excluded from analysis.) After the first 30 trials, a new window appeared on the screen with the information that either “from now on the cues will be sorted in descending order of insider expertise”<sup>4</sup> (structured cue order) or, “from now on the cues will no longer be sorted”<sup>5</sup> (unstructured cue order). Participants were randomly assigned to start with one of these orders.

<sup>4</sup>Followed by a more detailed explanation: “From now on the first cue you uncover is automatically the one with the highest number of forecasts (in the last 100 races). Within a trial every further uncovered cue has therefore a smaller validity than the previous uncovered cues.”

<sup>5</sup>Followed by a more detailed explanation: “From now on the first cue you uncover is no longer automatically the one with the highest number of forecasts (in the last 100 races). Within a trial every further uncovered cue can have an arbitrary validity.”

### 2.2.4 Results

The results of the calculations for the individual data are summarized in the Appendix. For each participant, the number of shown ORDM and MRDM trials is given (amounting to 29 in each of the two conditions). On average, MRDM was in the majority, even in the TTB-friendly condition ( $M = 63.4\%$ ,  $SD = 26.9\%$ , range 14–100%). However, there is a significant increase of the decisions based on one cue only (ORDM) from the Minimalist condition ( $M = 8.1$ ,  $SD = 5.1$ ) to the TTB condition ( $M = 10.6$ ,  $SD = 7.8$ ;  $t[35] = 2.63$ ,  $p = .007$ , one-tailed paired). This tendency is to be expected because people had to conclude from the instruction that they already had the best cue information at hand and that the immediately following piece of information would not be able to overrule it.

In addition to ORDM and MRDM stopping (which could be counted directly), the individual DLC was calculated for each condition (analogous to Experiment 1).

On average, DLCs remained stable at  $M = .81$  in both conditions ( $t[35] = .75, p = .228$ , one-tailed paired): the higher percentage of ORDM in the TTB condition did not affect these calculated values.<sup>6</sup>

Conformity with the different stopping rules (DLC versus ORDM and MRDM) was individually tested by separate binomial tests ( $\alpha = .05$ , requiring at least 21 correct predictions in the 29 trials). One participant showed a clear MRDM strategy in the Minimalist condition and maintained this behaviour in the TTB condition (no. 8). Surprisingly, four more participants joined this strategy in the TTB condition showing a DLC of 1.00. From the remaining 31 subjects being classified as DLC-orientated in their stopping rule during the Minimalist condition, 26 could be identified as DLC-orientated in the TTB condition as well, whereas the remaining five participants showed no significant stopping strategy in the TTB condition.

There was an ambiguous classification of four subjects of the TTB condition (nos. 9, 22, 18 and 12) with equal or more than 21 trials with an ORDM-stopping behaviour (out of 29): They could, alternatively to DLC, also be significantly classified as one-reason decision makers (applying the same binomial test criterion as for the DLC-strategy). Two of them (nos. 9 and 22), however, were better described by the DLC measured in the TTB condition.

In conclusion, ORDM as a major strategy in the TTB-friendly environment with structured cue validities is far from being a common option for the participants in this experiment. There are 13 participants decreasing their DLC by  $>.02$  (compared with the Minimalist condition), but 11 participants use the same DLC in both conditions, and 12 participants even increase it in the TTB-condition by  $>.02$ . In both condition people are therefore not willing to accept any cue validity as the single reason to make a final decision.

### 2.2.5 Performance

Performance of individuals should be seen in the light of “normative” behaviour, which in this particular experiment is the best general constant rule (or combination of general rules) to stop information search. Hence, the critical question is whether it is appropriate to continue information acquisition after the first (and mandatory) cue

is uncovered. Additional information acquisition is only of purpose when the initial cue information is overruled by the following cues. Should this be the case, the resulting expected value must exceed the initial expected value after the first uncovered cue. An a posteriori analysis of all trials in this experiment shows that this is never the case in the TTB condition. In the Minimalist condition three trials could be identified in which continued information acquisition would be the optimal behaviour (in one of these instances, three cues would have to be uncovered, in two trials even four cues). Since these trials (starting with cue validities of .70, .63, and .35) are not the ones providing the worst initial cues, it is unpredictable for the subjects that they should continue information acquisition in exactly these trials. Since these cases cannot be subsumed under a general rule, a consequent ORDM would be the normative behaviour for the subjects regarding the state of knowledge they have when conducting the experiment. This strategy would lead to a performance of € 5370 in the Minimalist condition and € 5430 in the TTB condition. It can only be outperformed if one subject accidentally violates the ORDM stopping rule in the unpredictable three trials mentioned above: In the Minimalist condition, there was such a chance to beat the ORDM performance with a constant threshold of .88 because this led to 100% correct answers. For the TTB condition, there was no such chance by whatever constant threshold. However, unpredictable inconsistent behaviour could by chance outperform the optimal strategy (and did in one case).<sup>7</sup>

Average performance in Minimalist condition was at € 5395 (SD = 435) which means that participants turned out to perform as well as the rational model (€ 5370). Individual performance and threshold were negatively correlated ( $p = -.56$ ): expanded information acquisition did not pay. Average performance in the TTB condition was at € 4438 (SD = 667) and therefore much lower than the rational ORDM stopping rule would have been (€ 5430). Only one participant was able to outperform the rational model (using a threshold of .66 and breaking his rule by pure luck in the right moments). Again, individual performance and threshold were negatively correlated ( $p = -.61$ ) which indicates that the more a participant behaved according to the ORDM stopping rule the more money he or she could earn.

<sup>6</sup>Participants consequently following the TTB stopping rule in the environment of descending cue validities would be expected to show a DLC of .57 (since this was the minimum “best cue” in the 29 analysed trials). Not a single participant showed this behaviour: everybody violated the TTB stopping rule at least from time to time. Even the individual (no. 18) showing the lowest DLC — therefore being closest to the ORDM stopping rule — reported after the experiment that he had established an explicit DLC of approx. .62 as his stopping rule during the experiment.

<sup>7</sup>Information integration models perform less well because additional acquisition becomes too costly and does not substantially enhance the quality of the decisions. In the best case (stopping immediately after finding two cues indicating the same decision), the performance is at € 4380 in the TTB condition and € 3570 in the Minimalist condition.



### 3 Discussion

One of the advantages of the sequential evidence accumulation approach is that researchers will have a more precise view of the individual decision making process. Classification into the three building blocks (search, stopping and decision rule) (Gigerenzer, Todd & the ABC Research Group, 1999) has already proved to be a useful instrument in describing individual decision behaviour in multi-attribute settings with probability cues. Although the decision rule has been predominant for a long time in the research history of decision making, the empirical support for an evidence accumulation approach reveals that the stopping rule is crucial indeed: it expresses the need for people to know at which point they have accumulated sufficient information on the pending decision problem. If they reach this point by accumulating evidence, they already “know” how they have to decide because in most of the cases it is obvious. As in the best case, with only one discriminating cue, the decision rule becomes trivial: “decide in the direction the cue points to” (see Lee & Cummins, 2004).

#### 3.1 Empirical indications for a DLC

In Experiment 1, we tested the desired level of confidence (DLC) with a very simple and direct method on the basis of the first uncovered cue. The horse race experiment showed very clear empirical data confirming the hypothesis of individual confidence thresholds (with a mean DLC of .73 in Experiment 1). The “inconsistent” behaviour Newell and other colleagues noted earlier can now be seen as a systematic search and decision pattern in nearly all of the tested participants: if the evidence from the first uncovered cue (validity) reached or exceeded the DLC, participants followed an ORDM, and if the value of the current cue validity was lower than the DLC, participants correspondingly searched for further information (violation of the stopping rule of ORDM) in order to increase their confidence.

At present, we still know very little about how people deal with conflicting information or, generally, what they do by continuing to search after discriminating cues have been found. Our experimental design was consequently confined to the question as to whether information search is stopped or continued after having the first cue information in hand, so that only assumptions about possible information integration can be made. At present, the maintenance of the DLC and the assumptions of an integration of validities are highly hypothetical and require further systematic investigation with a different or expanded experimental design.

In Experiment 2, we furthermore found a certain robustness, contrary to important changing environmental

factors such as cue structure (Minimalist- versus TTB-related cue order). People apply a confidence threshold not only in environments with cues appearing in a random order, but they also do so if they explicitly know that cues are well structured and cue validities appear in a descending order. This is far from normative behaviour because a second discriminating cue with a lower validity has no extra benefit in terms of a cost-benefit analysis. In the best case, new information can only confirm the direction of the first cue and therefore increase confidence. Precisely this desire for more confidence could be the key to understand the ongoing search in special cases (if current confidence < desired confidence). Therefore, being confident could have a value per se.

The desired level of confidence would be only one (nevertheless crucial) component in an integral process model of evidence accumulation. Therefore, influencing factors should be evaluated and tested empirically. A central environmental factor is the importance of the decision consequences for the individual (see also Lanzetta & Driscoll, 1968; Böckenholt, Albert, Aschenbrenner & Schmalhofer, 1991): Hausmann and Läge (2005), for example, were able to show that participants adjusted their DLC with changing amounts of possible winnings. Other environments should also be included in further research; for instance, Browne and Pitts (2004) suggested different types of problems (for example, choice problems versus design problems) in which people could have used different stopping rules (convergence towards a solution versus sufficiency of information).

As a surprising result, Hausmann and Läge (2005) showed an independence of the DLC from the number of options. It seems as if most people attach greater importance to confidence (“most probably I am correct”) than to normative behaviour (“I am significantly better than the base rate”) when making their decisions.

#### 3.2 Theoretical indications for a DLC: evidence accumulation and threshold models

In general, sequential sampling models assume an accumulation of information until there is sufficient evidence to favour one option. More or less independently from each other, different authors have similarly termed such an internal confidence threshold: “level of aspiration” (Lanzetta & Kanareff, 1962), “desire to produce an accurate response” (Holland & Kleinmuntz, 1994), “desired level of judgmental confidence” (Eagly & Chaiken, 1993), “desired level of confidence” (Hausmann & Läge, 2005), or “evidence accrual” (Lee & Cummins, 2004).

### 3.2.1 Heuristic and systematic information processing and the Principle of Sufficiency within and beyond the persuasion context

The “heuristic-systematic model” (HSM) from Chaiken, Liberman, and Eagly (1989) distinguishes between two different modes of information processing: systematic information processing on the one hand needs a considerable cognitive effort in comprehending, evaluating and integrating the message’s arguments in forming a final judgment, whereas on the other hand heuristic processing is less effortful and can include specific simple decision rules, schemata, or heuristics that mediate people’s attitudes such as, for example, the use of heuristic cues as source expertise (“experts’ statements can be trusted”), source likeability, message length, and consensus information. One basic assumption of the model (in persuasion settings) is that people must be motivated to engage in systematic processing, because people — as economy-minded souls — prefer less effortful to more effortful modes of information processing. But the “sufficiency principle” (the underlying motivation to hold accurate and valid attitudes) forces efficient information processors to strike a balance between minimizing their processing efforts and maximizing their judgmental confidence. The effected processing efforts (heuristic versus systematic) can therefore be viewed as a function of the discrepancy that exists between actual and desired levels of confidence (Eagly & Chaiken, 1993): “. . . that people will exert whatever level of effort is required to attain a sufficient degree of confidence that they have satisfactorily accomplished their processing goals” (Chaiken et al., 1989, p. 221). Although the underlying searching and stopping processes haven’t been tested empirically, Eagly, Chaiken, and their co-authors, have — within the field of social psychology — developed the idea of an individual confidence threshold (criterion point of sufficient or desired confidence) which can vary as a function of individual difference and situational factors.

In his article “Re-visions of rationality?”, Newell (2005) implicitly anticipates the advantages of such flexible threshold models. The repeatedly observed individual variability in decision making could be explained in one model, and a single threshold model could replace several discrete models (assuming, for example, that the threshold becomes higher the more important the decision is, or becomes lower the more exacting the time pressure is — down to consequent ORDM, or even guessing when time pressure is too exigent). In this respect, single heuristics in the adaptive toolbox (Gigerenzer et al., 1999) could be considered as a special case of a general threshold model (Lee & Cummins, 2004). Newell accurately expresses this when he says: “The ‘adjustable spanner’ perspective suggests that only one tool is used and that different

thresholds of accumulated evidence give rise to patterns of data that ‘mimic’ the stopping rule of the heuristics” (2005, p. 13). Hence, a confidence threshold would be the normal stopping rule, and the search behaviour observed would correspond to a specific single heuristic.

### 3.2.2 Alternative threshold models under uncertainty

Evidence accumulating approaches commonly try to model the individual decision process to include a stage of sequential gathering of information (search rule), the reaching of a threshold (stopping rule) and the decision for one of the options (a simple decision rule). For nearly all of these models, the termination of information search when reaching a threshold (stopping rule) is crucial. Apart from the two extremes (“search as much as you will find” and “search for one good reason”), many assumptions have been made and numerous investigations about different types of stopping thresholds have been conducted. In general, there could be more than one stopping mechanism (Newell’s “adjustable spanner” [2005] assumes a certain degree of flexibility), especially when considering that strategies are adapted to environments and circumstances. Time pressure (adhering to deadlines) can be mentioned as an external threshold, as well as cost arguments or other constraints (search costs, limited cognitive or material resources) (Simon, 1956).

### 3.2.3 Cost-benefit analysis as a stopping rule

The most explored threshold is probably that in cost-benefit models. The selection of different decision strategies has been seen as a result of a cost-benefit analysis in the way that people choose the strategy that requires the least investment for a satisfactory solution (Beach & Mitchell, 1978; Payne, Bettman & Johnson, 1988). Several authors have shown that information costs and awards do significantly affect depth, variability and latency of search (for example Lanzetta & Kanareff, 1962; Edwards & Slovic, 1965; Connolly & Serre, 1984; Gilliland, Schmitt & Wood, 1993; Saad & Russo, 1996; for an overview, see Payne, Bettman & Johnson, 1993). For example, lower search costs can lead to an extension of searched information, whereas higher search costs can lead to a restriction (for an overview, see Payne et al., 1993).

Gigerenzer, Todd, and the ABC Research Group argued against the rule “stop search when costs outweigh benefits” (listing it under “optimization under constraints”) because it would lead to an infinite calculating regress (1999, p. 11). Hausmann, Christen, and Läge (2006) proposed a simple mathematical model to calculate the economic value of the next obtainable cue only

(assuming people would be able to do this). In their experiment, they showed that people were cost-sensitive in principle, but they greatly overestimated the benefit of probabilistic information and spent — seen from a normative viewpoint — with increasing search cost, too much money for information. The authors interpreted this empirical fact as meaning that having useful (discriminating) information is valued higher than maximizing one's profit. Furthermore, people may collect information to avoid a pure guessing strategy. Other authors who have examined the decision process more closely came to a similar conclusion that individual information acquisition is more probably terminated by the principle of sufficient evidence (see for example Lee & Cummins, 2004).

### 3.2.4 Decision field theory as a threshold theory specific to preferential choice

A parallel approach, but explicitly related to preferential choice, is the decision field theory (DFT, see Busemeyer & Townsend, 1993; Busemeyer & Diederich, 2002). This model assumes a defined strength of preference at each stage of deliberating the different options. Should the individual express his or her preference, he or she would choose the one option with the highest strength of preference at that given moment. DFT defines a certain variation in the strength of preference for each of the options (and can therefore predict phenomena like preference reversal), but assumes only deliberation time and no active search for further information as a factor for this variability.

The DFT stopping rule is implemented either as fixed stopping time or as optional stopping time. The latter case shows a certain similarity with the desired level of confidence, because a critical strength of preference for one of the options is required to stop the process. However, each option in the preferential choice frame has a strength of preference independent of the others (they do not sum to 1.00 like the options in the horse race scenario presented above), so that no direct comparison between the stopping behaviour of the DFT (a critical strength of preference) and the DLC (a critical degree of certainty) can be made. The general conception in the theoretical framework, however, overlaps with the idea of a desired level of confidence by assuming a threshold that triggers the moment for the decision, and both models assume that no further information integration is then necessary to define the chosen option.

### 3.2.5 Sequential-sampling processes for psychophysical tasks

It is important to distinguish between conscious, strategic decisions (driven by reasoning and by emotions) and un-

conscious neural “decisions” on the level of psychophysiology. In the latter field, threshold models are very common for describing the process of distinction between a set of alternatives. Neurons seem to accumulate a sort of “evidence” before transmitting to the next neurons. Self-regulating Accumulator Models elaborated by Smith and Vickers (1988), and Vickers and Lee (1998, 2000), based on Vickers (1979), form an interesting approach even using terms like “level of confidence”. Since much of this work deals with lower-level (neural-based) analysis these models go beyond the scope of the current paper. Even though the work emphasizes that sequential evidence accumulation is a common procedure, these models cannot be referred to in detail within the framework of this paper.

## 3.3 Final conclusions

When carefully analysing all the models mentioned and other studies concerning evidence accumulation, one can conclude that several factors have an influence on the termination of information search, including information costs, amount of payoff, time pressure, complexity, importance, experience, and the level of confidence. There are several studies combining some of these factors, for example Lanzetta and Kanareff (1962; cost, payoff, aspiration), and Hlland and Kleinmuntz (1994; cost, time pressure, payoff, and experience). Other factors such as information costs, amount of payoff, time pressure, complexity, importance, experience, etc. are thought to be able to change the DLC in a specific decision task and can therefore indirectly influence the stopping of information search.

The concept of the desired level of confidence as a stopping rule in evidence accumulation tasks needs to be investigated in more detail, especially in other experimental settings. We are confident that further research in the field of sequential evidence accumulation, especially on the investigation of factors influencing the setting or the adjustment of the desired level of confidence, could be the key to helping us reveal and understand the complex connections contexts and mechanisms of decision behaviour under uncertainty.

## References

- Aschenbrenner, K. M., Albert, D., & Schmalhofer, F. (1984). Stochastic choice heuristics. *Acta Psychologica*, *56*, 153–166.
- Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *Academy of Management Review*, *3*, 439–449.
- Böckenholt, U., Albert, D., Aschenbrenner, M., & Schmalhofer, F. (1991). The effects of attractiveness,

- dominance, and attribute differences on information acquisition in multiattribute binary choice. *Organizational Behavior and Human Decision Processes*, 49, 258–281.
- Bröder, A. (2000). 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*, 26, 1332–1346.
- Bröder, A. (2005). *Entscheiden mit der “adaptiven Werkzeugkiste”*: Ein empirisches Forschungsprogramm [Decision making with the adaptive toolbox: An empirical research programme]. Lengerich: Pabst Science Publishers.
- Bröder, A. & Schiffer, S. (2003). “Take The Best” versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General*, 132, 277–293.
- Browne, G. J., & Pitts, M. G. (2004). Stopping rule use during information search in design problems. *Organizational Behavior and Human Decision Processes*, 95, 208–224.
- Busemeyer, J. R., & Diederich, A. (2002). Survey of decision field theory. *Mathematical Social Sciences*, 43, 345–370.
- Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, 100, 432–459.
- Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion process. In J. S. Uleman, & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). New York: Guilford Press.
- Connolly, T., & Serre, P. (1984). Information search in judgement tasks: The effects of unequal cue validity and cost. *Organizational Behavior and Human Performance*, 34, 387–401.
- Dieckmann, A., & Todd, P. M. (2005). Simple ways to construct search orders. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society* (pp. 309–314). Mahwah, NJ: Erlbaum.
- Doyle, J. (1999). Rational decision making. In R. A. Wilson & F. C. Keil (Eds.), *MIT encyclopedia of the cognitive sciences* (pp. 701–703). Cambridge, MA: MIT Press.
- Eagly, A. H., & Chaiken, S. (1993). *The Psychology of Attitudes*. San Diego, CA: Harcourt Brace Jovanovich.
- Edwards, W., & Slovic, P. (1965). Seeking information to reduce the risk of decision. *American Journal of Psychology*, 78, 188–197.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650–669.
- Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The Take The Best heuristic. In G. Gigerenzer, P. M. Todd & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 75–95). New York: Oxford University Press.
- Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: the adaptive toolbox. In G. Gigerenzer, P. M. Todd & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 3–34). New York: Oxford University Press.
- Gigerenzer, G., Czerlinski, J., & Martignon, L. (1999). *How good are fast and frugal heuristics?* In J. Shanteau, B. Mellers, & D. Schum (Eds.), *Decision science and technology: Reflections on the contributions of Ward Edwards* (pp. 81–103). Norwell: Kluwer Academic Publishers.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 506–528.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Gilliland, S.W., Schmitt, N., & Wood, L. (1993). Cost-benefit determinants of decision process and accuracy. *Organizational Behavior and Human Decision Processes*, 56, 308–330.
- Hausmann, D., Christen, S. & Läge, D. (2005). *Und sie lernen es doch! Das Erlernen von probabilistischen Cues und das Anwenden des Gelernten in einer konkreten Entscheidungssituation.* [Learning of probabilistic cue features and their use in a concrete decision making situation.] Forschungsberichte aus der Angewandten Kognitionsspsychologie, Nr. 08. Retrieved October 17, 2007, from Zurich University, Institute of Psychology Web site: <http://www.allgpsy.unizh.ch/AKZ-Forschungsberichte/articles/08-AKZ.htm>.
- Hausmann, D., Christen, S. & Läge, D. (2006). Urteils-Sicherheit um jeden Preis? Verletzung der Stopregel in Abhängigkeit von der Höhe der Informationssuchkosten [Judgment certainty at all costs? Violation of the stopping rule as a function of the amount of the information search costs]. In B. Gula, R. Alexandrowicz, S. Strauss, E. Brunner, B. Jenull-Schiefer, & O. Vitouch (Hrsg.), *Perspektiven psychologischer Forschung in Österreich. Proceedings zur 7. Wissenschaftlichen Tagung der Österreichischen Gesellschaft für Psychologie* (pp. 45–51). Lengerich, Pabst.
- Hausmann, D., & Läge, D. (2005). How the desired level of confidence influences the stopping rule of information search. In K. Opwis & I.-K. Penner (Eds.), *Pro-*

- ceedings of KogWis05. The German Cognitive Science Conference 2005* (pp. 71–76). Basel, Schwabe.
- Holland, J. S., & Kleinmuntz, D. N. (1994). Factors influencing the use of internal summary evaluations versus external information in choice. *Journal of Behavioural Decision Making, 7*, 79–102.
- Läge, D., & Daub, S. (2006). *Entscheiden mit mehreren Cue-Informationen: Die Entscheidungsregel am Beispiel des Städteparadigmas von Gigerenzer & Goldstein (1996)*. Forschungsberichte aus der Angewandten Kognitionspsychologie, Nr. 35. Retrieved October 17, 2007, from Zurich University, Institute of Psychology Web site: <http://www.allgpsy.unizh.ch/AKZ-Forschungsberichte/articles/35-AKZ.htm>
- Läge, D., Hausmann, D., Christen, S., & Daub, S. (2005). *Was macht einen guten Cue aus? Strategien der Informationssuche beim heuristischen Entscheiden unter Unsicherheit [What constitutes a good cue? Strategies for information search in decision making under uncertainty]*. Forschungsberichte aus der Angewandten Kognitionspsychologie, Nr. 05. Retrieved October 17, 2007, from Zurich University, Institute of Psychology Web site: <http://www.allgpsy.unizh.ch/AKZ-Forschungsberichte/articles/05-AKZ.htm>
- Lanzetta, J. T., & Driscoll, J. M. (1968). Effects of uncertainty and importance on information search in decision making. *Journal of Personality and Social Psychology, 10*, 479–486.
- Lanzetta, J. T., & Kanareff, V. T. (1962). Information cost, amount of payoff, and level of aspiration as determinants of information seeking in decision making. *Behavioral Science, 7*, 459–473.
- Lee, M. D., & Cummins, T. D. R. (2004). Evidence accumulation in decision making: Unifying the “take the best” and the “rational” models. *Psychonomic Bulletin and Review, 11*, 343–352.
- Leth-Steensen, C., & Marley, A. A. (2000). A model of response time effects in symbolic comparison. *Psychological Review, 107*, 62–100.
- Martignon, L., & Hoffrage, U. (1999). Why does one-reason decision making work? A case study in ecological rationality. In G. Gigerenzer, P. M. Todd & the ABC Research Group (Eds.), *Simple heuristics that make us smart*. (pp. 119–140). New York: Oxford University Press.
- Newell, B. R. (2005). Re-visions of rationality? *Trends in Cognitive Sciences, 9*, 11–15.
- Newell, B. R., Rakow, T., Weston, N. J., & Shanks, D. R. (2004). Search strategies in decision making: The success of “success”. *Journal of Behavioral Decision Making, 17*, 117–137.
- Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing “one-reason” decision-making. *Journal of Experimental Psychology: Learning, Memory & Cognition, 29*, 53–65.
- Newell, B. R., Weston, N. J., & Shanks, D. R. (2003). Empirical tests of a fast and frugal heuristic: not everyone “takes-the-best”. *Organizational Behavior and Human Decision Processes, 91*, 82–96.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 14*, 534–552.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press.
- Rouder, J. N. (2001). Testing evidence accrual models by manipulating stimulus onset. *Journal of Mathematical Psychology, 45*, 334–354.
- Saad, G., & Russo, J. E. (1996). Stopping criteria in sequential choice. *Organizational Behavior and Human Decision Processes, 67*, 258–270.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69*, 99–118.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*, 129–138.
- Smith, P. L., & Vickers, D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology, 32*, 135–168.
- Todd, P. M., & Dieckmann, A. (2005). Heuristics for ordering cue search in decision making. In L. K. Saul, Y. Weiss, and L. Bottou (Eds.), *Advances in neural information processing systems 17* (pp. 1393–1400). Cambridge, MA: MIT Press.
- Vickers, D. (1979). *Decision processes in visual perception*. New York: Academic Press.
- Vickers, D., & Lee, M. D. (1998). Dynamic models of simple judgments: I. Properties of a self-regulating accumulator module. *Nonlinear Dynamics, Psychology, and Life Sciences, 2*, 169–194.
- Vickers, D., & Lee, M. D. (2000). Dynamic models of simple judgments: II. Properties of a self-organizing PAGAN (parallel, adaptive, generalized accumulator network) model for multi-choice tasks. *Nonlinear Dynamics, Psychology, and Life Sciences, 4*, 1–31.

## Appendix: Stopping behavior in Experiment 2.

The table on the next page shows participants' stopping behaviour in Experiment 2 ( $n = 36$ ) in the Minimalist condition (uncovered cue validities were randomly distributed) and the TTB condition (cue validities were structured in a decreasing order). For both conditions the number of ORDM (information search stopped after the first uncovered cue) and MRDM (search continued after the first uncovered cue) is shown (amounting to a total of 29 trials for each condition). The desired level of confidence (DLC) was calculated for both conditions separately (level with the minimal sum of error values; for details see Experiment 1 or Hausmann & Läge, 2005). The number of DLC hits was calculated from the number of correctly used ORDM (when the first uncovered cue validity was equal to or higher than the calculated DLC) and the number of correctly used MRDM (when the first uncovered cue validity was lower than the calculated DLC). The classification of a DLC-stopping strategy was deduced if the number of DLC hits was equal to or larger than 21 (binomial test,  $\alpha = .50$ ); a non-significant number ( $< 21$ ) was classified as "none" strategy. A classification of an MRDM-stopping strategy was deduced if the number of MRDM was 29 (100% of the trials), and an ORDM-stopping strategy would have been resolved if the number of ORDM was 0.

## Notes on the table:

S = Participant number

Seq. = Sequence of the minimalist and TTB-condition (Min-TTB or TTB-Min) Minimalist condition: Column 3 to 6 contains the number (#) of trials following one of the models (ORDM, MRDM), the calculation (calc.) of the DLC, and the number of DLC hits.

TTB condition: Column 7 to 10 contains the number (#) of trials following one of the models (ORDM, MRDM), the calculation (calc.) of the DLC, and the number of DLC hits.

The last two columns indicate classification in the Minimalist and the TTB condition (DLC, ORDM, MRDM or none).

Stopping behavior in Experiment 2 (see last page).

S	Seq.	Minimalist condition				TTB condition				Classification	
		# of ORDM	# of MRDM	Calc. DLC	# of DLC hits	# of ORDM	# of MRDM	Calc. DLC	# of DLC hits	Minimalist condition	TTB condition
8	Min-TTB	0	29	1.00	29	0	29	1.00	29	MRDM	MRDM
14	Min-TTB	3	26	.91	27	0	29	1.00	29	DLC	MRDM
1	Min-TTB	4	25	.89	26	0	29	1.00	29	DLC	MRDM
3	Min-TTB	2	27	.89	26	0	29	1.00	29	DLC	MRDM
32	TTB-Min	6	23	.81	22	0	29	1.00	29	DLC	MRDM
23	TTB-Min	4	25	.89	29	2	27	.92	27	DLC	DLC
33	TTB-Min	4	25	.89	29	5	24	.89	28	DLC	DLC
27	TTB-Min	4	25	.88	27	4	25	.89	27	DLC	DLC
34	TTB-Min	5	24	.87	27	1	28	.94	27	DLC	DLC
6	Min-TTB	5	24	.87	27	10	19	.85	25	DLC	DLC
35	TTB-Min	4	25	.86	26	7	22	.85	21	DLC	DLC
13	Min-TTB	4	25	.85	22	1	28	.94	27	DLC	DLC
11	Min-TTB	5	24	.84	25	2	27	.93	26	DLC	DLC
5	Min-TTB	7	22	.84	26	6	23	.88	24	DLC	DLC
7	Min-TTB	7	22	.84	26	7	22	.83	25	DLC	DLC
15	Min-TTB	7	22	.82	29	4	25	.89	26	DLC	DLC
16	Min-TTB	7	22	.82	24	13	16	.74	24	DLC	DLC
29	TTB-Min	7	22	.81	27	15	14	.70	25	DLC	DLC
2	Min-TTB	7	22	.80	27	9	20	.84	28	DLC	DLC
10	Min-TTB	8	21	.80	27	12	17	.79	29	DLC	DLC
9	Min-TTB	8	21	.80	27	22	7	.64	23	DLC	DLC
20	TTB-Min	10	19	.78	27	17	12	.70	28	DLC	DLC
19	TTB-Min	11	18	.76	25	17	12	.68	22	DLC	DLC
25	TTB-Min	10	19	.76	24	18	11	.67	23	DLC	DLC
28	TTB-Min	13	16	.75	27	15	14	.71	25	DLC	DLC
17	TTB-Min	13	16	.75	29	17	12	.68	26	DLC	DLC
30	TTB-Min	16	13	.72	24	11	18	.80	27	DLC	DLC
26	TTB-Min	16	13	.72	28	16	13	.71	29	DLC	DLC
31	TTB-Min	16	13	.72	29	16	13	.71	28	DLC	DLC
22	TTB-Min	18	11	.72	23	23	6	.63	25	DLC	DLC
18	TTB-Min	21	8	.67	28	25	4	.61	24	DLC	DLC
12	Min-TTB	4	25	.89	29	21	8	.66	16	DLC	none
36	TTB-Min	3	26	.85	23	17	12	.69	19	DLC	none
24	TTB-Min	8	21	.81	26	20	9	.71	17	DLC	none
4	Min-TTB	10	19	.78	29	11	18	.77	19	DLC	none
21	TTB-Min	16	13	.72	23	18	11	.73	19	DLC	none