Do People Make Decisions
Under Risk Based on
Ignorance?
An Empirical Test of the
Priority Heuristic against
Cumulative Prospect
Theory

Andreas Glöckner / Tilmann Betsch

MAX PLANCK SOCIETY

Do People Make Decisions Under Risk Based on Ignorance? An Empirical Test of the Priority Heuristic against Cumulative Prospect Theory

Andreas Glöckner / Tilmann Betsch

February 2008

Do People Make Decisions Under Risk Based on Ignorance?
An Empirical Test of the Priority Heuristic against Cumulative Prospect Theory[∗]

Andreas Glöckner[†] / Tilmann Betsch[‡]

Abstract

Brandstätter, Gigerenzer and Hertwig (2006) put forward the priority heuristic (PH) as a fast and frugal heuristic for decisions under risk. According to the PH, individuals do not make trade-offs between gains and probabilities, as proposed by expected utility models such as cumulative prospect theory (CPT), but use information in a non-compensatory manner and ignore information. We conducted three studies to test the PH empirically by analyzing individual choice patterns, decision times and information search parameters in diagnostic decision tasks. Results on all three dependent variables conflict with the predictions of the PH and can be better explained by the CPT. The predictive accuracy of the PH was high for decision tasks in which the predictions align with the predictions of the CPT but very low for decision tasks in which this was not the case. The findings indicate that earlier results supporting the PH might have been caused by the selection of decision tasks that were not diagnostic for the PH as compared to CPT.

Keywords: Decision Strategy, Fast and Frugal Heuristics, Bounded Rationality, Decision Latency, Process Tracing, Cumulative Prospect Theory

[†]   Address Correspondence to: Andreas Glöckner, Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, 53113 Bonn, Germany, Phone: +49(0)228–9141651, eMail: gloeckner@coll.mpg.de

[‡]   Department of Psychology, University of Erfurt, Germany.

Many everyday decisions have to be made under risk. Different options lead to outcomes that have different values and are realized with different probabilities. Such decision tasks can be conceptualized as choices between gambles (or prospects). A gamble $(x_1, p_1; \ldots ; x_n, p_n)$ leads to outcomes x_i with a probability of p_i. The probabilities p_i add up to 1, which means that all possible outcomes of the gamble are specified. One of the most fundamental assumptions in decision research is that individuals integrate outcomes and probabilities in a compensatory way and select the option with the highest weighted sum (von Neumann & Morgenstern, 1947; Edwards, 1954; Savage, 1954). In this tradition the (cumulative) prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) postulates that the expected utility V ("value") of gambles can be calculated by:

$$V = \sum_i \pi(p_i)v(x_i), \tag{1}$$

where the weighting function $\pi(p_i)$ is an inverse S-shaped function of the probabilities p_i and the value function $v(x_i)$ is an S-shaped function of the outcomes x_i. The cumulative prospect theory (CPT) proposes specific mathematical equations for both functions.[1]

Brandstätter, Gigerenzer, and Hertwig (2006) have recently argued that under certain conditions individuals use simple non-compensatory strategies (cf. Fishburn, 1974) instead of integrating information in a weighted compensatory manner, as advocated by CPT. These authors proposed the *priority heuristic* (PH) as an alternative process model for risky choices. According to the PH individuals do not carry out calculations of expected utilities (or values) but instead compare pairs of outcomes and probabilities using certain difference thresholds (cf. lexicographic semi-order; Tversky, 1969). In order to support the PH empirically, Brandstätter et al. report simulations in which the PH successfully predicted choices of the majority of participants. We argue that this data is not conclusive evidence in favor of the PH because (a) aggregated choice data, instead of individual choice data, have been analyzed and (b) the simulated choice sets did not include sufficient diagnostic decisions that could have allowed for a differentiation between the PH and the CPT. We carried out three experiments that eliminated these problems by using diagnostic decision tasks (i.e., decision tasks in which the predictions of the PH and the CPT differ) and analyzing patterns of individual choices.

We aim to show in this article that the proposed non-compensatory PH does not capture decision behavior of individuals concerning choices, decision times and information search, and that individual choices follow a weighted compensatory integration of values and probabilities.

---

[1]   According to the cumulative prospect theory (Tversky & Kahneman, 1992) the value function for non-negative outcomes is defined by $v(x) = x^{\alpha}$ and the weighting function is defined by $\pi(p) = p^{\gamma}/(p^{\gamma} + (1-p)^{\gamma})^{1/\gamma}$. The parameter α captures the shape of the value function and the parameter γ captures the shape of the weighting function.

The Priority Heuristic

Mathematical Formalization

The PH consists of two different process models for choices between gambles with all positive (or zero) outcomes (all $x_i \geq 0$) and choices between gambles with all negative (or zero) outcomes (all $x_i \leq 0$). The PH does not state a procedure for gambles with both positive and negative outcomes, which structurally also underlies many everyday decisions. In this paper, choices between two gambles A and B with two outcomes each ($n = 2$), all of which are positive or zero, will be investigated. For simplicity, we use a notation in which the outcomes for each gamble are ordered by their values from low to high and decisions have to be made between

$$\text{Gamble A} : (x_{MIN(A)}, p_{MIN(A)}; x_{MAX(A)}, p_{MAX(A)})$$
$$\text{Gamble B} : (x_{MIN(B)}, p_{MIN(B)}; x_{MAX(B)}, p_{MAX(B)}) \quad \text{and}$$

in which $x_{MIN(A)}$ is the minimum outcome of gamble A, $p_{MIN(A)}$ is the probability of the minimum gain, $x_{MAX(A)}$ is the maximum gain of gamble A, $p_{MAX(A)}$ is the probability of the maximum gain of gamble A. The same notation is used for gamble B. The expected value of the gambles can be computed by

$$EV_A = x_{MIN(A)} p_{MIN(A)} + x_{MAX(A)} p_{MAX(A)}$$
$$EV_B = x_{MIN(B)} p_{MIN(B)} + x_{MAX(B)} p_{MAX(B)} \qquad (2)$$

Furthermore, X_{MAX} is defined as the overall highest gain $X_{MAX} = \max(x_i)$, and the aspiration level X_{ASP} is defined as 1/10 of X_{MAX}, rounded to prominent numbers (i.e., 1, 2, 5, 10, 20, 50, 100, 200, 500, 1000, 2000; cf. Brandstätter et al., 2006)

$$X_{ASP} = round_{PROM}\left(\frac{X_{MAX}}{10}\right). \qquad (3)$$

The PH specifies three rules (Brandstätter et al., 2006, p. 413): (1) Search rule (priority rule) - Consider information in the following order: minimum outcome (x_{MIN}), probability of minimum outcome (p_{MIN}), maximum outcome (x_{MAX}). (2) Stopping rule - Stop the examination if the minimum outcomes differ by the aspiration level X_{ASP} or more; otherwise stop if the probabilities differ by .10 or more. (3) Decision rule - Choose (in this order) the gamble with the higher minimum outcome, the lower probability of the minimum outcome, or the higher maximum outcome. Furthermore, it is mentioned that individuals use an initial screening phase to determine whether the expected values of the gambles (EV_A and EV_B) are similar. If expected values differ substantially, the PH is not applied.

In sum, the PH postulates a non-compensatory, three-step decision strategy that can be formalized as follows:

Step 1:   Choose gamble A if $x_{MIN(A)} - x_{MIN(B)} \geq X_{ASP}$,

                  choose gamble B if $x_{MIN(B)} - x_{MIN(A)} \geq X_{ASP}$

                  otherwise continue with step 2.

Step 2: Choose gamble A if $p_{MIN(B)} - p_{MIN(A)} \geq 0.10$,

choose gamble B if $p_{MIN(A)} - p_{MIN(B)} \geq 0.10$

otherwise continue with step 3.

Step 3: Choose gamble A if $x_{MAX(A)} > x_{MAX(B)}$,

choose gamble B if $x_{MAX(B)} > x_{MAX(A)}$,

otherwise [not specified].

The initial screening phase could be considered a further step: it is determined whether the PH can be applied or not. It should be noted that Brandstätter et al. (2006) specified this step only vaguely: "One way would be to assume that people estimate the expected values, and if the ratio is smaller than two, they turn to the priority heuristic… Alternatively, people may first look at the three (four) reasons, and if no difference is markedly larger than the others, they apply the priority heuristic" (p. 426). For pragmatic reasons, we will mathematically specify the first mentioned alternative.

Step 0: Screen gambles to estimate expected values and X_{ASP},

if $\dfrac{1}{2} < \dfrac{EV_A}{EV_B} < 2$, continue with step 1,

otherwise [not specified].

To clarify this procedure, the PH will be explained using an example. Gamble A (€20, .80; €25, .20) pays €20 with a probability of .80, or €25 with a probability of .20 (EV_A = €21). Gamble B (€8, .90; €90, .10) pays €8 with a probability of .90, or €90 with a probability of .10 (EV_B = €16.20). An initial screening of the payoffs leads to the result that the expected values do not differ markedly and that the PH can be applied. In the screening phase the aspiration level (X_{ASP}) is determined to be €10. The decision maker would first compare the minimum outcomes of the alternatives ($x_{MIN(A)}$ = €20 and $x_{MIN(B)}$ = €8). The difference between these two payoffs is larger than the aspiration level. Thus, the inspection of further information will be stopped and the decision rule will be applied. Accordingly, it is assumed that the decision maker will choose the gamble with the higher minimum outcome (Gamble A) and ignore the remaining information.

As can be seen from the example, the PH is a simple heuristic. It simplifies the decision process by setting priorities. Probabilities and higher outcomes are neglected if the minimum outcomes of alternatives differ sufficiently. Processing of information primarily involves calculating differences but never integrating probabilities and values.

Empirical Evidence for the PH

Brandstätter et al. (2006) show that (a) the PH may account for well-known violations of rationality, (b) the PH may account for observed majority of choices (i.e., choices selected by the majority of participants) in a sample of existing data sets drawn from major publications in the field,

and (c) decision times are in line with the predictions of the PH. By contrast, Johnson, Schulte-Mecklenbeck, and Willemsen (2008) report data from process-tracing studies which indicate that individuals do not use the PH but instead apply compensatory integration models such as the CPT. In the following we will briefly discuss these findings.

Violations of rationality. Brandstätter et al. (2006) show that in certain gambles the PH can account for major violations of rationality, such as the Allais paradox (Allais, 1953), the reflection effect (Fishburn & Kochenberger, 1979), the certainty effect (Kahneman & Tversky, 1979), the possibility effect (Kahneman & Tversky, 1979), the fourfold pattern of risk-seeking and risk-aversion (Tversky & Fox, 1995), and intransitivity effects (Tversky, 1969). However, it should be noted that the reported examples do not necessarily prove that the PH can account for these violations in all possible gambles. Furthermore, it might be argued that other models can also account for most of these violations.

More importantly, but not mentioned by Brandstätter et al. (2006), the PH allows for differential predictions concerning the established violation of rationality, which we will discuss using the certainty effect as an example. Let us consider a choice between gambles A (€10, .60; €225, .40) and B (€0, .40; €160, .60). According to the certainty effect (and to CPT), people should choose gamble A (and avoid gamble B with the zero outcome option). According to the PH, people should choose gamble B because expected values do not differ (step 0), differences in minimum outcomes do not exceed the aspiration level (step 1), but the probability of the minimum outcome is lower in gamble B (step 2). This finding is of central importance for the current work because it allows us to differentiate between the CPT and the PH. The PH predicts that the certainty effect should diminish in gambles in which the differences between the minimum outcomes are low and the €0 outcome has a lower probability than the minimum outcome of the other gamble. Thus, such decisions are *diagnostic* for the use of the PH or the CPT, in that they allow us to identify if one or the other strategy has been used.

Simulations of majority choices. In a series of simulations, Brandstätter et al. (2006) compared the predictions of the PH and other decision strategies for a sample of existing data sets drawn from major publications in the field. Specifically, they tested whether model predictions converge with the choice of the *majority* of participants. As competitors to the PH, expected utility models (e.g., CPT) and prominent heuristics (e.g., equal weight strategy) were selected. The results revealed that the PH predictions match the choices of the majority quite well. In 87% of the 260 considered decision tasks, the PH correctly predicted if gamble A or B was chosen by the majority of participants. Thus, the hit rate of PH predictions was well above the level of chance (50%).

Nonetheless, the conclusion that individuals have indeed used the PH might have been somewhat rushed. Most importantly, Brandstätter et al. (2006) report that up to 89% of the PH predictions dovetail with the predictions derived from CPT (using parameters employed by Erev et al., 2002; as cited in Brandstätter et al., 2006). In other words, the decision tasks used in the contests cannot be used to identify which of the two strategies was used by the participants. Most decision

tasks are simply not diagnostic for the considered decision strategies. On a more general level, a further criticism can be levied; namely, that the PH is conceptualized as a process model for individual choice behavior, and thus it should be able to predict individual choices instead of aggregated majority choices. In the analysis reported by Brandstätter et al. (2006), observed choices were aggregated over participants (e.g., 70% choices for gamble A) and dichotomized into majority choices (e.g., choice for gamble A if more than 50% of the participants choose gamble A). Obviously, such a data reduction makes it hard to draw sound conclusions about individual decision processes. Therefore, we re-analyzed the first data used by Brandstätter et al. (14 decision tasks taken from Kahneman & Tversky, 1979). First, we computed the success of the PH in predicting individual choices. Not surprisingly, the percentage of correct predictions dropped from 100% for majority choices to 77% for individual choices. This result might be partially explained by the fact that people made processing errors in applying their decision strategy. However, a closer look at the distribution of errors revealed that the drop was not likely to be caused by random processing errors only. The predictive accuracy of the PH for individual choices differed considerably between gambles, ranging from 58% to 92%. The CPT could explain these fluctuations by individual differences in the parameters of the weighting and the value function. However, such systematic fluctuations in error rates could not be explained by the PH.

Decision times. The PH predicts that decision time increases with the increasing numbers of computational steps needed to differentiate between gambles (cf. mathematical formalization) because in such cases more cognitive operations have to be carried out (see also Payne, Bettman, & Johnson, 1988). Brandstätter et al. (2006) could support this hypothesis in two experiments in which the number of necessary steps was systematically manipulated. However, the results should be interpreted with caution because, in the manipulation, other properties of the gambles that might also influence decision time, such as the ease of calculation and differences in the expected utilities of the gambles, have not been controlled for (cf. Cartwright, 1941; Cartwright & Festinger, 1943). Thus, the findings lend initial support for process predictions of the PH, but need further investigation.

Results from process tracing. According to the PH, individuals should search for information by comparing minimum outcomes, probabilities of minimum outcomes and maximum outcomes between gambles (e.g., moving from $x_{MIN(A)}$ to $x_{MIN(B)}$). Integration models like CPT would, by contrast, predict that information is searched within gambles (e.g., moving from $x_{MIN(A)}$ to $p_{MIN(A)}$). Johnson, Schulte-Mecklenbeck, and Willemsen (2008) tested this prediction of the PH using the computer based information search tracing tool MouselabWeb. Participants were presented with gambles for which outcomes and probabilities were initially hidden in a matrix of information cells. Individuals had to search for information by moving the mouse on the information cell. As soon as the mouse was moved over the information cell, its content was shown. The content was shown only as long as the mouse was placed on the information cell. Johnson et al. found that individuals mainly searched information within gambles. Comparisons between gambles are much less frequently observed. Thus, information search data speak against the view

that the PH is employed but support the view that integrative models such as the CPT are employed. However, it is not possible to rule out that the way the information was presented in the experiment might have influenced information search. In a different work we have shown that under certain conditions the mouselab indeed influences decision strategies by restricting information search (Glöckner & Betsch, 2007). In comparison to an open matrix presentation, which allows for information to be inspected without constraint, we found that the mouselab particularly hampers automatic processes of decision making that might be important in the screening phase (step 0) postulated by the PH. Thus, the information search data that speak against the view that the PH is employed should be interpreted with caution.

In sum, empirical results are equivocal and call for more thorough testing. We argue that this testing should consider three major methodological aspects. Most importantly, *diagnostic decision tasks* for which the primarily considered strategies PH and CPT make different predictions should be selected. Second, because the PH constitutes a process model, it cannot be cogently tested on aggregate choice data; *individual choice data* have to be analyzed. Third, in experiments that aim at investigating information search, the research method should *not hinder information search,* and open information presentation formats should be used.

Methodological Preliminaries

Diagnostic decision tasks. To test the PH against CPT it is necessary that both models make different predictions on the choices for some decision tasks and that a sufficient number of such decision tasks are included in the analysis. As already mentioned, a distinct prediction of the PH is that the certainty effect should disappear if the difference between minimum outcomes is low and the zero outcome is less probable. This can be formalized as follows: If in a decision task between two gambles with two positive outcomes and similar expected values (cf. step 0) the three conditions

$$(1)\ x_{MIN(A)} = 0 \,; x_{MIN(B)} > 0$$
$$(2)\ x_{MIN(B)} - x_{MIN(A)} < X_{ASP}$$
$$(3)\ p_{MIN(B)} - p_{MIN(A)} \geq 0.1$$

are fulfilled, the PH predicts that gamble A should be chosen, whereas CPT (with standard parameters;[2,] predicts that gamble B should be chosen. Note that if condition (1) is fulfilled and one of the other conditions is violated, the PH can predict that gamble B will be chosen. We will refer to decision tasks that fulfill all three conditions as CERT$_{CON}$ decision tasks (i.e., decision tasks in which the PH predictions are contrary to the certainty effect). We will call decision tasks that fulfill only the first condition CERT$_{PRO}$ decision tasks (i.e., decision tasks in which PH predictions are pro certainty effect). Obviously, such decision tasks can be used to identify individual decision strategies. A person who always chooses gamble B should be classified as a CPT

---

2   See footnote 1.

user, whereas a person who always chooses gamble A in CERT$_{CON}$ and gamble B in CERT$_{PRO}$ decision tasks can be classified as a PH user.

Further types of decision tasks that have the potential to differentiate between the PH and the CPT are decision tasks in which individual differences in the assessment of probabilities and values [i.e., in the curvature γ of the weighting function $\pi(p_i)$ and the curvature α of the value function $v(x_i)^3$] can lead to choices that differ from the PH. From the indefinite number of possible decision tasks, we selected three categories using a theory-driven approach. The selection was validated by analyzing choice predictions using CPT with two different sets of parameters measured by Erev et al. (2002; as cited in Brandstätter et al., 2006) and Tversky and Kahneman (1992).

First, we selected decision tasks for the category SIM, in which the expected utilities of both gambles were similar. Furthermore we selected decision tasks for the two categories MED$_{ALM\_CERT}$ (i.e., decision tasks in which a medium outcome is almost certain) and MED$_{CERT}$ (i.e., decision tasks in which a medium outcome is certain), which should also be sensitive to differences in the parameters of individual weighting and value function because they include a combination of high probabilities and high outcomes. We were able to confirm this theory-driven selection by calculating and comparing expected utilities according to CPT using the two above-mentioned sets of parameters. For the category SIM, both sets of parameters led to similar expected utilities of the gambles. In the two other categories, the manipulation led to preference reversals between gambles.[4]

SIM decision tasks consisted of two gambles with all similar outcomes and rather similar probabilities such as a decision between gambles A (€50, .6; €60, .4) and B (€40, .5; €69, .5). MED$_{ALM\_CERT}$ decision tasks are decision tasks in which a medium outcome is almost certain because, for both gamble A and gamble B, there is a large chance to win a medium outcome and a small chance to win a very high outcome or nothing [e.g., a decision between the gambles A (€100, .99; €0, .01) and B (€55, .90; €500, .1)]. MED$_{CERT}$ decision tasks are problems in which a medium outcome is certain in gamble A because both outcomes are medium, and in gamble B there is a high chance to get a lower outcome and a small chance to get a very high outcome [e.g., a decision between gambles A (€146, .98; €160, .02) and B (€67, .85; €600, .15)].

As intended by the selection of decision tasks, on an aggregated level the PH and CPT make completely different predictions concerning choices for the considered categories of decision tasks (Table 1). For simplicity, all decision tasks are coded in such a way that gamble A is the choice predicted by the PH. If individuals perfectly use the PH, all choices in all categories

---

3   See footnote 1.
4   Erev et al. (2002; as cited in Brandstätter et al., 2006) estimated the curvatures of the value function and the weighting function to be $\alpha = 0.33$ and $\gamma = 0.75$. Based on a different data set, Tversky and Kahneman (1992) estimated the parameters to be $\alpha = 0.88$ and $\gamma = 0.61$. We used both sets of parameters to calculate expected utilities (*V*) using Formula 1. As expected, for all SIM decision tasks used in Experiments 1 and 3 (see Appendix A) very similar expected utilities were found. For all MED$_{ALM\_CERT}$ and MED$_{CERT}$ the application of the parameters by Erev et al. (2002; as cited in Brandstätter et al., 2006) as compared to the parameters by Tversky and Kahneman (1992) led to the predicted preference reversals.

should be in favor of gamble A. Considering random processing errors in applying the PH, the probability of choices for gamble A should be $p = 1 - \varepsilon$. Under the assumption of random processing errors, it can be assumed that the error rate is equal in all categories: $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = \varepsilon_4 = \varepsilon_5$. Therefore, the probabilities of choices for gamble A should also be the same: $p_1 = p_2 = p_3 = p_4 = p_5 = 1 - \varepsilon$. In contrast, if all individuals use CPT with a constant error rate of ε_P, almost all participants ($p_1 = 1 - \varepsilon_P$) should choose gamble A in the CERT$_{PRO}$ decision tasks, whereas almost none of the participants ($p_2 = \varepsilon_P$) should choose gamble A in the CERT$_{CON}$ decision tasks. Medium levels of choices for gamble A are predicted for SIM, MED$_{ALM\_CERT}$ and MED$_{CERT}$ decision tasks because, depending on individual weighting and value functions, a substantial share of the participants (p_R) can be expected to choose gamble B (besides random processing errors).

	Proportion of Choices for Gamble A	
	Priority Heuristic	Cumulative Prospect Theory
CERT$_{PRO}$	very high ($p_1=1-\varepsilon_1$)	very high ($p_1=1-\varepsilon_P$)
CERT$_{CON}$	very high ($p_2=1-\varepsilon_2$)	very low ($p_2=\varepsilon_P$)
SIM	very high ($p_3=1-\varepsilon_3$)	medium ($p_3=1-p_{R1}-\varepsilon_P$)
MED$_{ALM\_CERT}$	very high ($p_4=1-\varepsilon_4$)	medium ($p_4=1-p_{R2}-\varepsilon_P$)
MED$_{CERT}$	very high ($p_5=1-\varepsilon_5$)	medium ($p_5=1-p_{R3}-\varepsilon_P$)

Table 1
Aggregated Choice Predictions for Decision Tasks

Analysis of individual choices. One of the challenges of behavioral decision research is to identify decision strategies on an individual level without influencing decision strategies by the research paradigm (Glöckner & Betsch, 2007). One method that has been repeatedly used (e.g., Bergert & Nosofsky, 2007; Bröder, 2003; Bröder & Schiffer, 2003a; Bröder & Schiffer, 2003b; Bröder, 2005; Glöckner, 2006) is a maximum-likelihood method for comparative model testing (cf. Wasserman, 2000). Thereby, individual choice patterns are compared with the differential predictions of a set of decision strategies to determine which strategy is most likely to account for the individual choice patterns (under the assumption of a fixed error probability). Specifically, in our experiments we compared the predictions of the PH against the predictions of CPT. The test of CPT requires that the parameters for the weighting and value functions be estimated (see Footnote 1 and Footnote 4). In line with Brandstätter et al. (2006), we aimed to avoid estimating the optimal parameters for the CPT in reference to the data set for which choices should be predicted; we used sets of parameters estimated in different earlier studies instead. Note that this method disfavors CPT (and thus makes classifications for the PH more likely) because in this way suboptimal parameters are considered for CPT. We nevertheless used this method for two reasons: first, to avoid the objection that CPT profits from the fact that it has more free pa-

rameters than the PH, and second, to replicate Brandstätter et al. (2006). For strategy classification we used two of the sets of parameters also applied by Brandstätter et al. (2006), which were estimated in earlier studies by Tversky and Kahneman (1992) and Erev et al. (2002; as cited in Brandstätter et al., 2006).[5]

Process measures. One widely used method to access correlates of decision processes consists in tracing information search. It can be used to validate findings from individual choice analyses. Johnson et al. (2008) used the MouselabWEB to present outcomes and probabilities of gambles and to measure information search by recording mouse movements. In recent research, Glöckner and Betsch (2007) have shown that the mouselab procedure itself influences decision making by hindering information search (cf. Lohse & Johnson, 1996). This might be particularly problematic for the PH which predicts that, before applying the heuristic, individuals engage in a quick screening of the payoffs to form an initial impression of the decision task (step 0). It is possible that the mouselab procedure hinders screening and changes the individuals' decision strategies.

This problem can be avoided by applying an *open mouselab method* (Glöckner & Betsch, 2007). Participants should be given the opportunity to quickly screen the entire set of information, and thereafter information search is able to be recorded by tracing participants actions after instructing them to move the mouse to the piece of information they have just considered. Such a procedure decreases the likelihood that the information presentation format will hinder the application of the PH because it makes fast screening possible. We used such an open mouselab method in the third experiment.

Overview of Experiments

We tested choice predictions and process predictions of the PH in three experiments. All studies applied diagnostic decision tasks. Choice data was supplemented with complementary data obtained from process measures, particularly decision time data (Experiment 1 and 3) and information search parameters (Experiment 3). In Experiment 1 we tested whether the PH can predict individual choices and decision times when diagnostic decision tasks are used and choices are analyzed on an aggregated and on an individual level. Specifically, we aimed to test the choice predictions of the PH and the CPT. In Experiment 2 we aimed to replicate these results and to investigate the more general hypothesis behind the PH that information is integrated in a non-compensatory manner, rather than a compensatory one. Thereby, we took into account the possibility that individuals might use different aspiration levels. In the third experiment, based on an open mouselab method, we tested the process predictions of the PH that information is mainly searched by a process in which outcomes between gambles are compared.

---

5   The parameters can be found in footnote 4 above.

Gamble 1		
A	50%	100€
B	50%	50€

Gamble 2		
A	65%	40€
B	35%	140€

Figure 1.

Presentation format of decision tasks used in Experiment 1 and 2

Experiment 1

Method

Participants and design. Eighty-two undergraduate students with different majors from the University of Erfurt participated in the experiment, which was part of a one-hour experimental battery. They were paid €6 and completed 40 decision tasks. Decision tasks were manipulated using a within-subjects design in which each respondent makes 8 similar decisions in each of the 5 categories of decision tasks.

Procedure and materials. The experiment was computerized. First, participants were familiarized with the decision task using the following instruction:

> You will be repeatedly presented with pairs of gambles. You should decide which of the gambles you would prefer to play. Each gamble has exactly two possible outcomes, A and B, which have some given probability of coming about. Each outcome leads to a certain monetary gain, which is given in Euro. [Example omitted] This means that in gamble 1 you have 30% chance of winning €10 and a 70% chance of winning €20, and so on. Your task is to select the gamble you prefer. Press "y" to select gamble 1 and press "m" to select gamble 2. [page break] If you have any questions, please ask the experimenter; otherwise press continue [page break] Please note: try to make accurate decisions and to proceed as quickly as possible. Put the left finger on "y" and the right finger on "m" and press both keys to start the decision phase.

The time pressure instruction was used to encourage the application of simple heuristics like the PH. Gambles were presented on the left and right sides of the screen (Figure 1), and participants were to select between the gambles by pressing a key on the left or right side of the keyboard (please note that the position of the keys "y" and "z" are reversed in German keyboards). Choices and decision times were recorded as dependent variables.

Forty decision tasks (see Appendix A) were shown in individual randomized order. The presentation order of the gambles (i.e., gamble A first or gamble B first) and the order of outcomes within each gamble (i.e., low outcome first vs. high outcome first) were changed between trials to prevent gamble A from always being the gamble predicted by the PH. The 40 presented decision tasks were eight different versions of the five categories of decision tasks introduced above

(see Table 1): $CERT_{PRO}$, $CERT_{CON}$, SIM, MED_{ALM_CERT}[6] and MED_{CERT}. Each gamble had two non-negative outcomes. The expected value for each pair of gambles was approximately equal.[7] As such, the conditions for the application of the PH were optimal: individuals were instructed to decide fast, expected values did not differ between gambles, and it was possible to quickly screen the information.

For convenience, we will shortly recapitulate the predictions of CPT concerning choices: in $CERT_{PRO}$ decision tasks, the predictions of the PH are in line with the predictions of the CPT; thus a high percentage of choices for gamble A (i.e., the gamble that is predicted by the PH) is expected. In $CERT_{CON}$ decision tasks, the PH predicts against CPT and a low percentage of choices for gamble A is expected. In SIM decision tasks, expected values are rather similar; therefore a medium level of choices for gamble A is expected. In MED_{ALM_CERT} and MED_{CERT} decision tasks, again a medium frequency of choices for gamble A is expected because individual differences in value and weighting function can lead to choices for gamble A and B. In contrast, the PH predicts an equally high proportion of choices for gamble A in all five categories.

As outlined above, the predictions of the PH were compared with the predictions of CPT with two sets of parameters derived by Erev et al. (2002; as cited in Brandstätter et al., 2006) and Tversky and Kahneman (1992). As intended by the authors, there was little overlap among the three considered decision strategies on choice predictions in the 40 decision tasks selected for the experiment. In 62% of the decision tasks the PH and the *CPT-Erev* (i.e., the CPT with the parameters derived by Erev et al., 2002; as cited in Brandstätter et al., 2006) made the same predictions. For PH and *CPT-T&K* (i.e., CPT with the parameters derived by Tversky & Kahneman, 1992), the overlap was at 45%. The overlap between CPT-Erev and CPT-T&K was 48%. Thus, the overlap in predictions was considerably lower than in the decision tasks used by Brandstätter et al. (2006), which should allow for a reliable classification of decision strategies.

According to the PH, decision times should be the same for the categories $CERT_{PRO}$, SIM, MED_{ALM_CERT} and MED_{CERT} because in all cases, at least the aspiration level, the minimum outcomes of both gambles differ. Thus, for all four categories of decision tasks, equivalent decision times can be expected. Only in $CERT_{CON}$ decision tasks the probability of the minimum gain has to be considered, and a higher decision time is predicted by the PH.

We expected that individuals will not use the PH and that the results reported by Brandstätter et al. (2006) are caused by the fact that insufficient diagnostic decision tasks were used and data was analyzed on an aggregated level only. Thus, we tested the following hypotheses derived from the PH and expected all of them to be falsified:

---

6   Due to rounding to prominent numbers, the difference in minimum gains lay below the rounded aspiration level in decision task 4 of the category MED_{ALM_CERT}, which was not intended by the authors. Additional analysis showed that omitting decision task 4 from analysis would not substantially influence the results.

7   In some decision tasks, expected values of the gambles differed slightly due to rounding errors. The differences were in each case lower than €1.

H_1: The proportions of choices for gamble A do not differ between the categories CERT$_{PRO}$, CERT$_{CON}$, SIM, MED$_{ALM\_CERT}$ and MED$_{CERT}$.

H_2: The majority of individual choice patterns are in line with the predictions of the PH.

H_3: Decision times are higher in CERT$_{CON}$ decision tasks than in the remaining decision tasks.

Results and Interpretation

Aggregated choices. As mentioned above, all choices were coded so that the PH predicts choices for gamble A only. To test the first hypothesis – i.e., that choice probabilities for gamble A and B do not differ between categories – a χ^2-test was conducted on the independence between choices (gamble A, gamble B) and the categories of decision tasks (CERT$_{PRO}$, CERT$_{CON}$, SIM, MED$_{ALM\_CERT}$, MED$_{CERT}$). The test turned out to be highly significant, χ^2 (4, N=3280) = 593.0, p < .001. Thus, hypothesis H_1 was rejected. Figure 2 indicates that the specific prediction of the PH – namely that the certainty effect would disappear if the minimum gains did not differ considerably – was not supported by the data. In CERT$_{CON}$ decision tasks, almost all participants selected the option predicted by CPT and the certainty effect (i.e., gamble B) instead of the option predicted by the PH.
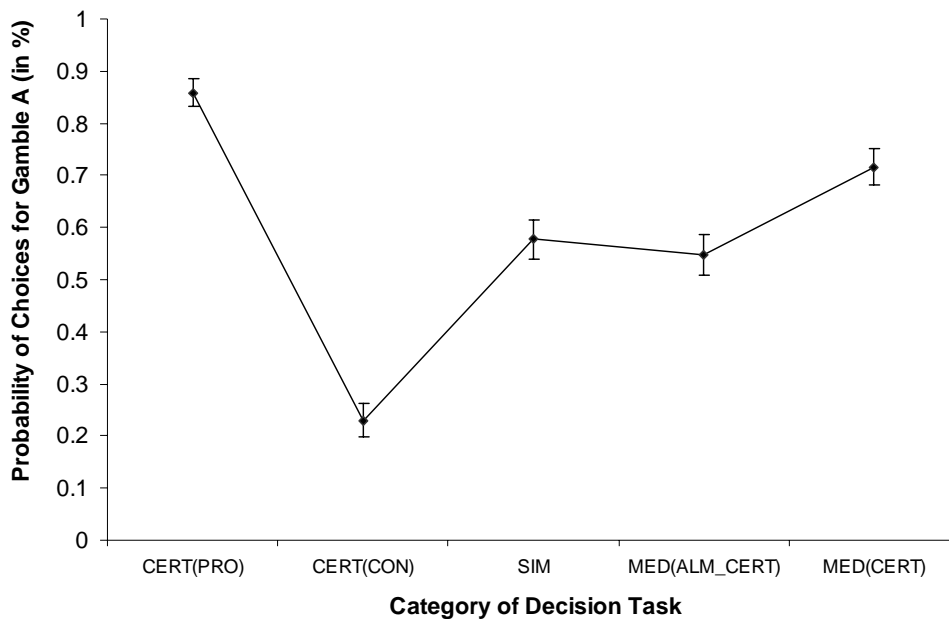

Figure 2.
Choices for gamble A (i.e., PH prediction) in Experiment 1, with error bars indicating 95% confidence intervals for proportions.

Note that aggregated choice proportions for the categories descriptively follow the predictions of the CPT (cf. Table 1). A high proportion of choices for gamble A was observed in CERT$_{PRO}$ de-

13

cision tasks; a low proportion was observed in CERT$_{CON}$ decision tasks; and the proportions in the other categories lay in between.

Individual analysis. Overall, only 59% of the individual choices were in line with the predictions of the PH. Note that the predictive accuracy of the PH is therefore just slightly above the level of chance (i.e., 50% in decision tasks between two gambles). However, it would be incorrect to conclude from this proportion that none of the participants used PH, as it would also be wrong to conclude that 59% of the participants did use this strategy. Therefore an analysis was run on an individual level to determine if individual choices can be better explained by the PH, the CPT-Erev or the CPT-T&K.

The model fit for each of the three considered decision strategies was estimated using the maximum-likelihood method (also called *Bayesian strategy classification method*; Bröder & Schiffer, 2003a) outlined above. The method is based on the estimation of the conditional likelihood that a specific individual choice pattern is observed given a certain decision strategy and a constant error rate. The specific details of the strategy classification method are described in Appendix B.

The results of the strategy classification are depicted in Table 2. For about 78% of the participants, choices can be best explained by one of the two versions of CPT. Only for 9% of the participants did the PH offer a better account of the observed choice patterns. Thus, hypothesis H$_2$, which proposed that the majority of individual choice patterns are in line with the PH, is also not supported by the data.

	Strategy Classification			
	n	%	Prediction Rate	*Md(L-Ratio)*
PH	7	.09	.67	2
CPT (Erev et al., 2002)	40	.49	.75	18
CPT (Tversky & Kahneman, 1992)	23	.28	.77	39
Not classified	12	.15		

Table 2
Individuals' Decision Strategies in Experiment 1

Note. The table shows the results of the individual strategy classification for the 82 participants of Experiment 1; *n* refers to the number of participants classified for the respective strategy. The column *Prediction Rate* shows the proportions of correct predictions of the respective strategy for the participants classified as users of this strategy. The far right column shows the median of the likelihood ratios for the strategy classification. The likelihood ratio is computed by dividing the likelihood for the classified strategy by the sum of the likelihoods of the remaining strategies. Higher values indicate that strategy classification is more reliable.

The median of the likelihood ratios, which are computed by dividing the likelihood value for the best strategy by the sum of the likelihood values for the remaining strategies for each participant, is given in the far right column of Table 2. The likelihood ratio is a measure for the reliability of the classification, with high numbers indicating high reliability. It can be seen that classifications for CPT were rather reliable, whereas classifications for the PH were less reliable. Furthermore,

the third column of Table 2 shows the proportion of correct predictions of each strategy for the participants classified as users of this strategy. In line with the results concerning likelihood ratios, both versions of the CPT had a higher correct prediction rate than the PH.

Decision times. First, decision times were screened for extreme outliers (M +/- 3 SD) and very low decision times ($t < 500$ms), which were deleted (2.3% of the data points). The average decision time for the remaining data points was 5.5 seconds. Decision times were analyzed using a repeated measurement analysis of covariance with the five categories of decision tasks as the within-subjects factor, and with the order of the presentation as a covariate. The main effect of the factor category turned out to be highly significant, $F(4, 326) = 22.6$, $p < .001$, $\eta^2 = .22$, indicating substantial decision time differences between categories. Average decision times and 95% confidence intervals are depicted in Figure 3. It can be seen that the highest decision time is observed for SIM. To test the third hypothesis – namely, that decision times are higher for CERT$_{CON}$ than for the remaining decision tasks – a planned contrast was computed, comparing decision times of CERT$_{CON}$ against the mean. The contrast turned out to be significant, $z = 2.48$, $p < .05$, but in the opposite direction. Decision times in the CERT$_{CON}$ decision tasks were significantly lower than the mean of the remaining decision tasks. Thus, hypothesis H$_3$, which was derived from the PH, was rejected as well.
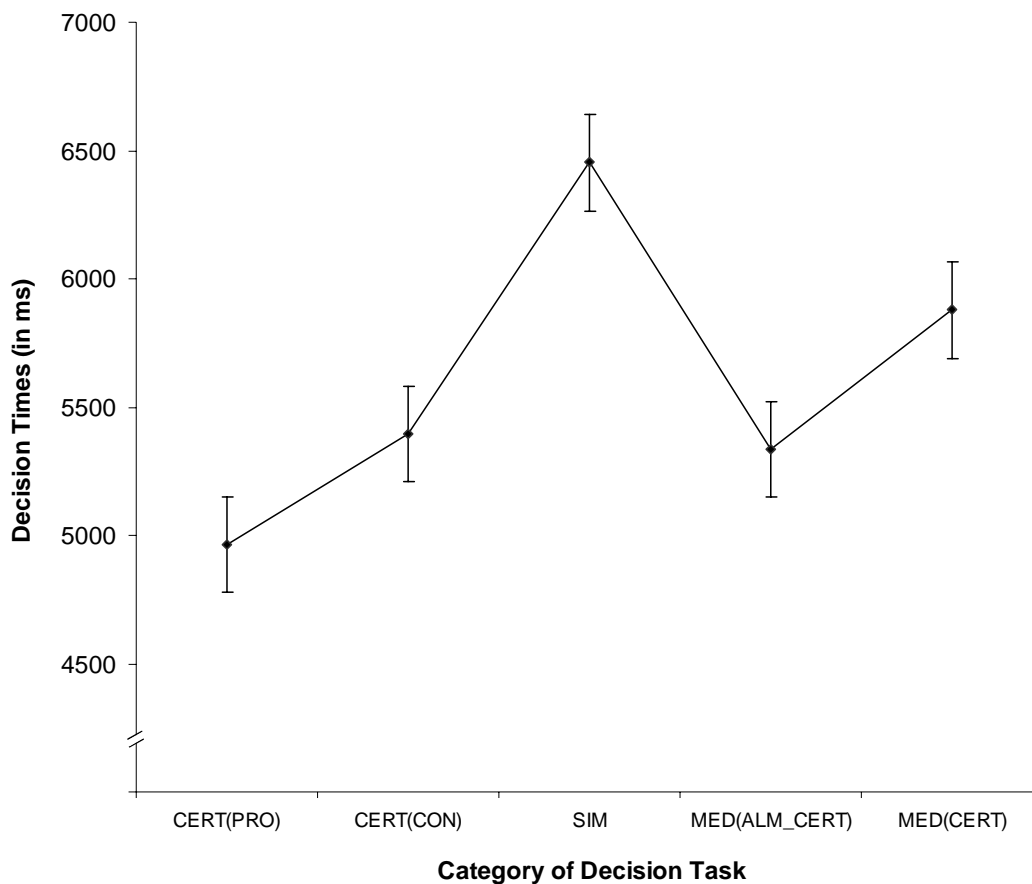

Figure 3.
Decision times in Experiment 1 with error bars indicating 95% confidence intervals.

15

Discussion

The aim of the first experiment was to provide a critical empirical test of the PH as compared to CPT, using diagnostic decision tasks and an individual-level analysis of decision strategies. Based on our findings, three hypotheses that were derived from the PH had to be rejected. First, it was shown on an aggregated level that the proportion of choices for gamble A (i.e., PH predictions) differs significantly among categories. In the categories that contain mainly un-diagnostic gambles (i.e., gambles in which the predictions of the PH and CPT dovetail; e.g., $CERT_{PRO}$), a high proportion of choices for gamble A was observed, whereas choices for gamble A were considerably lower in categories with diagnostic gambles (e.g., $CERT_{CON}$). The data did not support the PH prediction that the certainty effect should disappear under specific conditions. In sum, aggregated choices were in line with the predictions of CPT and strongly conflicted with the predictions of the PH (cf. Table 1 and Figure 2). Second, it was found that the large majority of individual choice patterns can be better explained by CPT than by the PH. There were only very few individuals for whom choices were roughly in line with the predictions of the PH, and the mean reliability of these classifications was low. Finally, it could be shown that decision times were not in line with the prediction of the PH. No increased decision time for decision tasks in the category $CERT_{CON}$ were observed; on the contrary, decision time was even significantly lower in this category compared to the mean of the decision time in the remaining categories.

In our experiment, the observed rate of choices that could be explained by the PH was considerably lower than in the simulations reported by Brandstätter et al. (2006) (59% vs. 87%), and it was marginally above chance level. According to our data, the major reason for this difference (besides aggregation effects) is that Brandstätter et al. used decision tasks that were not diagnostic for the decision strategies PH and CPT. In categories of decision tasks in which there was a high overlap between the choice predictions from the PH with CPT, a high accuracy rate for the PH was observed, whereas the prediction rate was substantially lower in categories with low overlap. In the decision tasks used in the simulations of Brandstätter et al., choice predictions of the PH and the CPT-Erev had an overlap of 89%; respectively, the PH was observed to lead to 87% correct predictions. In our experiment the overlap was 62%, the PH was observed to lead to 59% correct predictions.

Decision time results sharply conflict with the previous findings (Brandstätter et al., 2006, pp. 423). Individual decision times did not increase in those tasks for which the PH predicted an increase in the depth of processing. In line with classic (Cartwright, 1941; Cartwright & Festinger, 1943; Festinger, 1943a; Festinger, 1943b) and more recent (Bergert & Nosofsky, 2007; Glöckner, 2006) findings, decision times were highest for "hard" decision tasks in which the (subjective) expected utility of both options was almost equal (i.e., SIM decision tasks).

In Experiment 2, we wanted to replicate the findings of Experiment 1 using different diagnostic decision tasks. In addition we wanted to examine whether the predictive performance of the PH increases if aspiration levels of different sizes are considered (cf. Brandstätter et al., 2006, p. 425). A generalized version of the PH that was based on relaxed assumptions about aspiration

levels was tested. To carry out the analysis, we introduced a generalized lexicographic (*genLEX*) strategy, which is a generalization of the PH that allows for aspiration levels ranging from 1/10 to 1/3 (for minimum gains and probabilities of minimum gains). The upper limit of aspiration levels (i.e., 1/3) was selected based on the authors' estimation that gains (and probabilities) that differ by more than 1/3 (or .33) should usually not be considered to be equal. Specifically, we tested the following hypothesis derived by genLEX and expected that this hypothesis would be rejected:

H_4: Manipulations of minimum gains and probabilities of minimum gains below the aspiration level of 1/10 and above the aspiration level of 1/3 do not influence choices.

Additionally, in an exploratory account we aimed to investigate whether individuals indeed use aspiration levels for differences in minimum gains or minimum probabilities when they compare gambles. Therefore, in five sets of decision tasks, differences of minimum gains and probabilities of minimum gains were systematically manipulated.

Experiment 2

Method

Participants and design. One hundred and fifty-four students from different majors took part in the study, which was part of a 1.5-hour experimental battery. They were rewarded with €9. Decision tasks were manipulated within-subjects between five *sets*. Each set consisted of six to twelve decision tasks in which one fixed reference gamble was compared with different competing gambles. Competing gambles were manipulated within sets so that the differences in minimum gains or probability of minimum gains were systematically varied, resulting in the further within-subjects factor *decision tasks*, which was nested within the factor sets.

Procedure and materials. Experiment 2 used essentially the same procedure as in the previous experiment, but with different decision tasks. Forty decisions between two gambles with two outcomes and a constant expected value of €100 were used. Five sets of decision tasks were constructed, in which one gamble was held constant and one pay-off (minimum gain or probability of minimum gain) for the other gamble was varied. To hold the expected value of the manipulated gamble constant, a complementary parameter of the same gamble was changed accordingly (see Appendix C).

In the first set, gamble A (€90, .50; €110, .50) was compared with gamble B, for which the minimum gain was varied on 12 levels from €85 to €0 (while the maximum gain was increased from €115 to €200 respectively). Thus, gamble B was manipulated from

$$(€85, \quad .50_{[fix]}; \quad €115, \quad .50_{[fix]}) \text{ to}$$

$$(€0, \quad .50_{[fix]}; \quad €200, \quad .50_{[fix]}).$$

17

Elements with the subscript "[fix]" were held constant. According to the genLEX, there should be a flip in choice proportions from gamble B to gamble A after the aspiration level for differences in minimum gains is reached. The inflection point of the choice function would indicate the aspiration level that is used by the majority of the participants.

In sets 2 and 3, the minimum gains of both gambles did not differ by more than 1/10 (set 2), or the minimum gains were exactly equal (set 3). The probabilities of the minimum gains were manipulated on 8 levels ranging from high to low probabilities. In both sets the fixed gamble was gamble A (€10, .90; €910, .10). In set 2, the probabilities and outcomes of gamble B were manipulated from

$$(€0_{[fix]}, .99; \qquad €10000, .01) \text{ to}$$

$$(€0_{[fix]}, .20; \qquad €125, \quad .80).$$

In set 3, the probabilities and outcomes of gamble B were manipulated from

$$(€10_{[fix]}, .85; \qquad €610, \quad .15) \text{ to}$$

$$(€10_{[fix]}, .10; \qquad €110, \quad .90).$$

According to the genLEX, in set 2 all choices should be for gamble A (a flip of choice proportions would only be expected if the aspiration level of probabilities is below .10). Choice proportions in set 3 should flip from gamble A to gamble B after the difference in probabilities of minimum gains reaches the aspiration level for probabilities of minimum gains.

In sets 4 and 5, again, the difference in minimum gains between the gambles was manipulated. In both sets, gambles were compared with a fix gamble A (€70, .70; €170, .30), and the minimum gain was decreased in 6 levels, starting from close to €70. The crucial difference between the two sets concerned the probabilities of the minimum gains. In set 4, the probabilities of the minimum gains were higher for the fix gamble A than for the varied gamble B, whereas in set 5, the probabilities of minimum gains for gamble B were higher. In set 4, the outcomes and probabilities of gamble B were manipulated from

$$(€65, \quad .50[fix]; \qquad €135, .50[fix]) \text{ to}$$

$$(€40, \quad .50[fix]; \qquad €160, .50[fix]).$$

In set 5, the outcomes and probabilities of gamble B were manipulated from

$$(€60, \quad .90[fix]; \qquad €460, \quad .10[fix]) \text{ to}$$

$$(€0, \quad .90[fix]; \qquad €1000, .10[fix]).$$

In set 4, according to the genLEX, a flip in choice proportions from gamble B to gamble A would be expected when the aspiration level used by the majority of individuals is reached. In set

5, no changes in choice proportions would be predicted by the genLEX because minimum gains as well as the probability of minimum gains point towards gamble A in all decision tasks.

In sum, sets 1 and 4 aimed to determine if there is a constant aspiration level for the first reason postulated by the PH (and the genLEX), "minimum gain". If the argumentation of Brandstätter et al. (2006) is correct, this aspiration level should be around 1/10 of the maximum gain, which should result in a flip of choice proportions after this level is reached. In sets 2 and 3, aspiration levels for the second reason, "probability of minimum gain," are measured. If the PH holds true, the expected inflection point for choice proportions should be around .10. For both reasons, observed inflection points could be used to determine the aspiration level mainly used by individuals.

To replicate findings from Experiment 1, the decision tasks of the five sets were again selected to be diagnostic for the decision strategies PH, CPT-Erev and CPT-T&K. There was an overlap in predictions of .67 between the PH and CPT-Erev, and an overlap of .47 between the PH and CPT-T&K. Both CPT versions shared .55 of the choice predictions. Note that the overlap of the strategies with the genLEX cannot be computed because choice predictions differ for different aspiration levels.

Results and Interpretation

Individual choice analysis. First, the proportion of individual choices in line with the predictions of the PH was computed per participant. The overall mean was .53 ($SE = .009$). Thus, the results of Experiment 1 could be replicated with a different set of diagnostic decision tasks. The predictive accuracy of the PH was again just slightly above the chance level. Second, individual choice patterns were analyzed using the maximum likelihood method described in Appendix B (sets were used as categories of decision tasks). The results replicated the findings of Experiment 1. For the large majority of participants, choices could be best explained by the versions of CPT (Table 3). Only for very few participants was the PH a better model for explaining their choices. Again, classifications for the PH were rather unreliable (i.e., lower median of likelihood ratios), and even for the eight participants classified as PH users, the PH made only .60 correct predictions, which is again just slightly above chance level.

	Strategy Classification			
	n	%	Prediction Rate	*Md(L-Ratio)*
PH	8	.05	.60	2
CPT (Erev et al., 2002)	75	.49	.77	18
CPT (Tversky & Kahneman, 1992)	31	.20	.68	5
Not classified	38	.25		

Table 3
Individuals' Decision Strategies in Experiment 2

Analysis of aspiration levels. Figures 4a to 4e depict the choice functions, with error bars indicating 95% confidence intervals of proportions for the decision tasks of sets 1 to 5. In sets 1, 3 and 4, two dotted lines mark the positions in which the limits of aspiration levels of 1/10 (left line) and of 1/3 (right line) are reached. If individuals use differences in minimum gains or probabilities of minimum gains in a non-compensatory manner and (plausible) aspiration levels from 1/10 to 1/3 (i.e., genLEX), a flip of choice proportions within the area marked by the dotted lines would be expected. Outside the dotted lines, no differences would be expected.
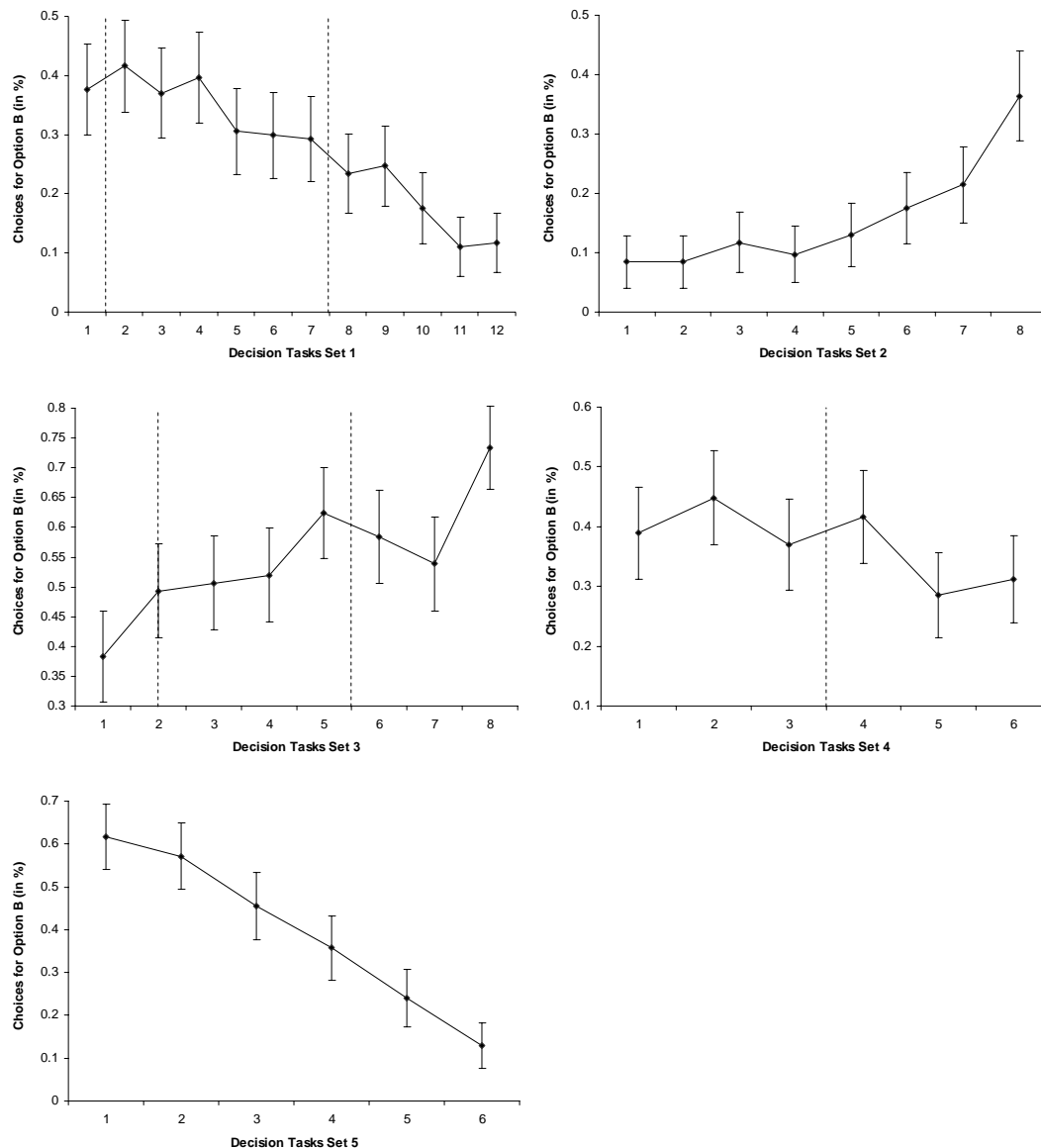

Figure 4a-e.
Choices in Experiment 2, with error bars indicating 95% confidence intervals for proportions. The two dotted lines mark when the aspiration levels of 1/10 (left line) and of 1/3 (right line) are reached.

We investigated H_4 by testing separately against the null hypotheses a) that manipulations of differences below 1/10 do not influence choices, and b) that manipulations of differences above

1/3 do not influence choices; χ^2-test of independence between choices for gambles A and B, and the respective decision tasks within each set were computed. The compared decision tasks and the results of the tests are summarized in Table 4.

	Manipulation	Changes in Choices for	Analyzed Decisions	Strategy Classification			
				N	df	χ^2	p
Set 1	$x_{MIN(B)}$	Differences>1/3	Tasks 8 to 12	770	4	17.09	<.01
Set 2	$p_{MIN(B)}$	Differences<1/10	Tasks 1 to 3	462	2	1.52	n.s.
		Differences>1/3	Tasks 7 to 8	308	1	11.13	<.01
Set 3	$p_{MIN(B)}$	Differences>1/3	Tasks 6 to 8	462	2	13.57	<.01
Set 4	$x_{MIN(B)}$	Differences<1/10	Tasks 1 to 3	462	2	2.11	n.s.
Set 5	$x_{MIN(B)}$	Differences<1/10	Tasks 1 to 6	924	5	115.72	<.001

Table 4

Tests against Equal Distributions of Choices in Experiment 2

Note. The table presents the results of tests of the H_4 – i.e., that manipulations of differences in minimum gains (x_{MIN}) or differences in probabilities in minimum gains (p_{MIN}) below 1/10 or above 1/3 should not influence choices. The three left columns indicate the set within which the manipulation was conducted, the variable which was manipulated within the set and the section of differences that was tested (i.e., the question is whether manipulations of differences below 1/10 aspiration levels influence choices or whether manipulations of differences above 1/3 aspiration levels influence choices). The fourth column indicates the decision tasks within the set, for which an equal distribution is predicted if the manipulation outside the aspiration levels (i.e., below 1/10 or above 1/3) does not influence choices. χ^2 tests of independence between choices for option A and B and the respective decision tasks (i.e., fourth column) were conducted to test the hypotheses. The number of observations, degrees of freedom, the χ^2 values and the p values of these significance tests are depicted in the four right columns. Significant values indicate that H_4 was rejected.

All three tests that investigated whether changes in differences above the maximal considered aspiration level (1/3) influence choices turned out significant. Furthermore, one of the three tests that investigated whether changes in differences below the minimal considered aspiration level (1/10) influence choices was significant. Thus, H_4, which postulates that changes of minimum gains and probabilities of minimum gains below an aspiration level of 1/10 and above an aspiration level of 1/3 do not influence choices, was rejected. These results speak against the usage of aspiration levels as hypothesized by the genLEX and the PH.

Further evidence against the application of a genLEX strategy is provided by a closer investigation of set 5. In set 5, genLEX users with any aspiration level for differences in minimum gains should choose gamble A in all six decision tasks, because the first and the second reasons (minimum gain and probability of minimum gain) speak for gamble A. Only people for whom the first two reasons do not differentiate (i.e., people with particularly high aspiration levels for the probability of minimum gains) select gamble B in all six decision tasks. Note, however, that

any distribution of intra-individually stable aspiration levels would lead to a constant rate of choices for gamble A (and gamble B) in all six decision tasks. In contrast to this prediction, choice proportions for gamble B significantly decreased between decision tasks 1 and 6 (cf. Table 4; last row). Therefore, set 5 provides the most clear-cut evidence against the genLEX and against the idea that information is used in a non-compensatory manner, as suggested by Brandstätter et al. (2006).

Discussion

The findings of Experiment 2 replicate the results of the previous experiment, in that the PH cannot account for individuals' choices for sets of decision tasks that contain diagnostic decisions. Again, CPT seems to be the more appropriate model for describing choices. Additionally, choices seem to be influenced by the manipulation of the minimum gains and the probabilities of minimum gains that were below and above reasonable aspiration levels (i.e., 1/10 and 1/3). This finding speaks against the fact that the reasons "minimum gain" and "probability of minimum gain" are used in a non-compensatory manner. Thus, the general principle underlying the PH − that individuals apply simple non-compensatory heuristics (Gigerenzer et al., 1999) in risky choices − has to be questioned.

Evidence of both reported experiments speaks clearly against the application of the PH and also against the general idea that the reasons postulated by Brandstätter et al. (2006) are considered in a non-compensatory manner based on aspiration levels. Thus, our findings are in line with the information-search results of Johnson et al. (2008), which indicate that information is not searched in line with the prediction of the PH. Nevertheless, as pointed out above, the results by Johnson et al. (2008) could be criticized for methodological reasons. It cannot be ruled out that the mouselab method used by Johnson et al. hindered information search and the application of the PH. To rule out this possible objection and to strengthen our reported results, a further experiment was conducted in which the decision tasks of Experiment 1 were presented in an open mouselab format so that information was instantly available to the participants. To bolster the application of the PH, individuals were explicitly instructed to screen all information before starting the decision process. Individuals were instructed to move the mouse to the information they were currently considering. This movement was used to record the information search. We tested the following hypotheses derived by the PH and expected to reject them:

H_5: Individuals inspect only a part of the information (as proposed by the stopping rule of the PH).

H_6: Individuals show information search transitions mainly between gambles (i.e., by comparing gains or probabilities between gambles) and fewer transitions within gambles.

Experiment 3

Method

Participants and design. Thirty-three undergraduate students of different majors from the University of Erfurt participated in the experiment, which was part of a one-hour experimental battery. They were paid €6. In line with Experiment 1, decision tasks were manipulated using a within-subjects design in which each participant makes 8 decisions in each of the 5 categories of decision tasks.

Procedure and materials. Experiment 3 employed essentially the same general instructions and the same decision tasks of Experiment 1. However, in contrast to Experiment 1, information was presented in an open mouselab (Figure 5) and participants were instructed to first screen information and then to move the mouse to the information currently being considered. The following additional instruction was used:

The information [about the gambles] will be presented in a circle format. On the left side you see the two possible outcomes of gamble 1, and on the right side you see the two possible outcomes of gamble 2. [example omitted] We would like to record your information search; to facilitate this, please try to follow your eye movements with the mouse. Please move the mouse to the point on the screen you are looking at. [example omitted] Try first to get an overview of the decision situation by screening all values. Then click on the start button in the middle of the screen to start the decision. Move the mouse to the information (gains or probabilities) you are currently inspecting. If you have come to a decision, click on the button of the gamble. [example omitted]
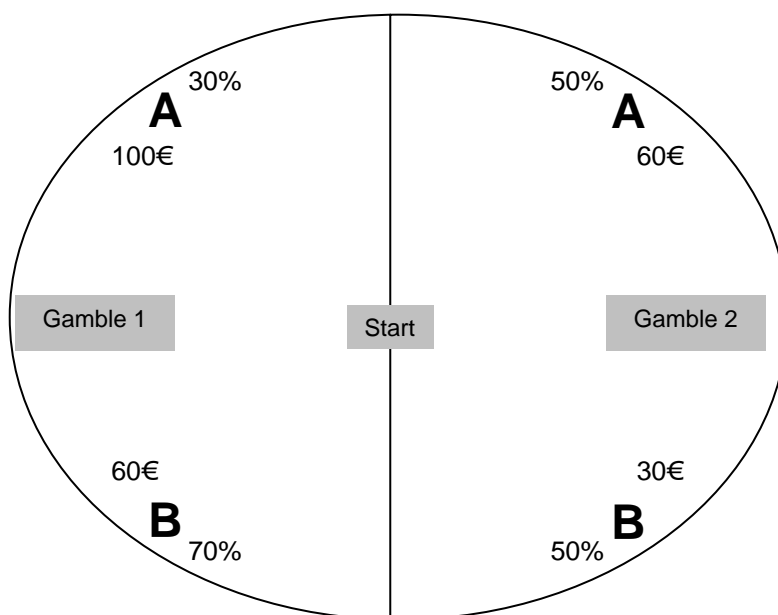

Figure 5.
Open mouselab presentation of gambles in Experiment 3.

The start button in the middle was used to center the mouse on the screen after the initial information screening so that the distance to all pieces of information was approximately equal. In the decision phase, information about what the mouse was currently pointing at was highlighted with a light-green background color. After each decision, an instruction screen was shown. The next decision task was started by mouse-click. In an attempt to force individuals to obey the instructions and to follow the eye movements with the mouse, it was not possible to select an option before at least two pieces of information had been inspected (and touched with the mouse). If individuals tried to select an option before that, a message box informed them that they should first inspect information and that they should always move the mouse to the inspected information so that it is highlighted.

After one learning trial with instruction and one learning trial without instruction, the 40 decision tasks of Experiment 1 were presented (Appendix A) in an individually randomized order. Note that the fourth decision task of the category MED_{ALM_CERT} was modified to correct for the rounding error discussed in Footnote 3. Thus, the fourth MED_{ALM_CERT} decision task in Experiment 3 was a choice between gambles A (€217, .92; €1800, .08) and B (€0, .02; €350, .98). Choices and information search were recorded. Decision times were measured from the pressing of the start button to the selection of an option.
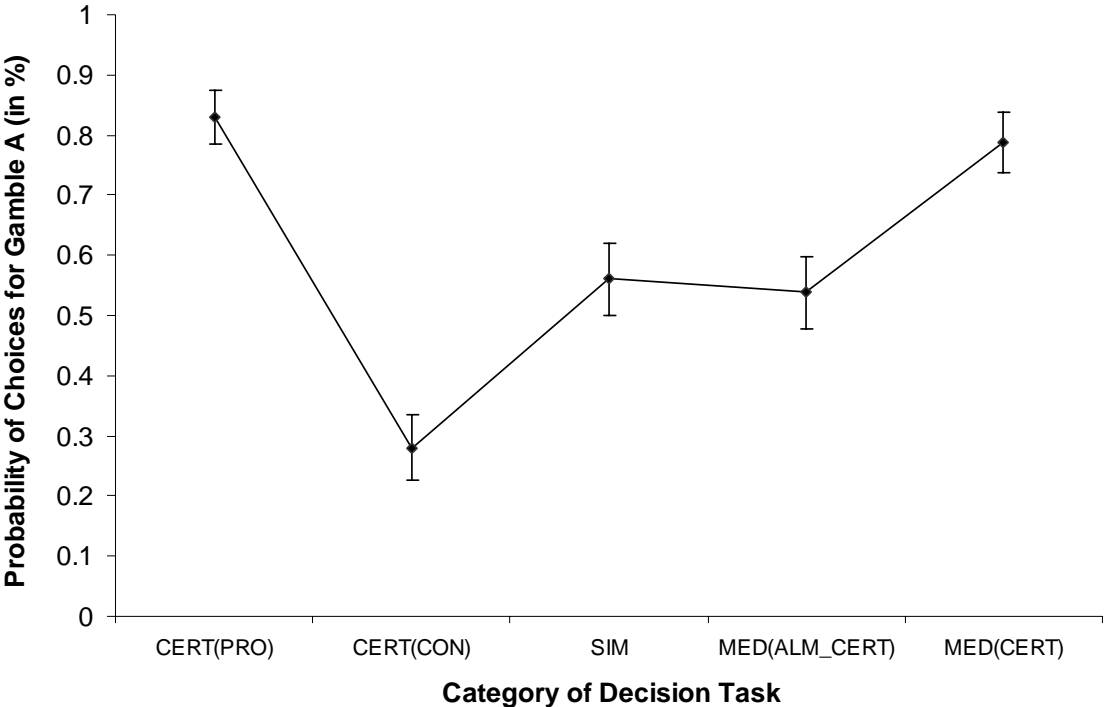

Figure 6.
Choices for gamble A (i.e., PH prediction) in Experiment 3 with error bars indicating 95% confidence intervals for proportions.

Results

Choices and decision time. The choice distributions almost exactly replicated the findings of Experiment 1 (Figure 6), indicating that the differences in choice proportions between categories are rather stable and that the open mouselab presentation method does not influence choices. Overall, again only 60 percent of the individual choices were in line with the predictions of the PH.

For the analysis of decision times, again extreme outliers (*M* +/- 3 *SD*) and decision times below 500ms were deleted; this was necessary for 1.8 percent of the data points. Decision times were overall slower than in Experiment 1, but they did replicate the general pattern (Figure 7). Again, decisions were made particularly slower in the SIM decision tasks and faster in the other decision tasks.
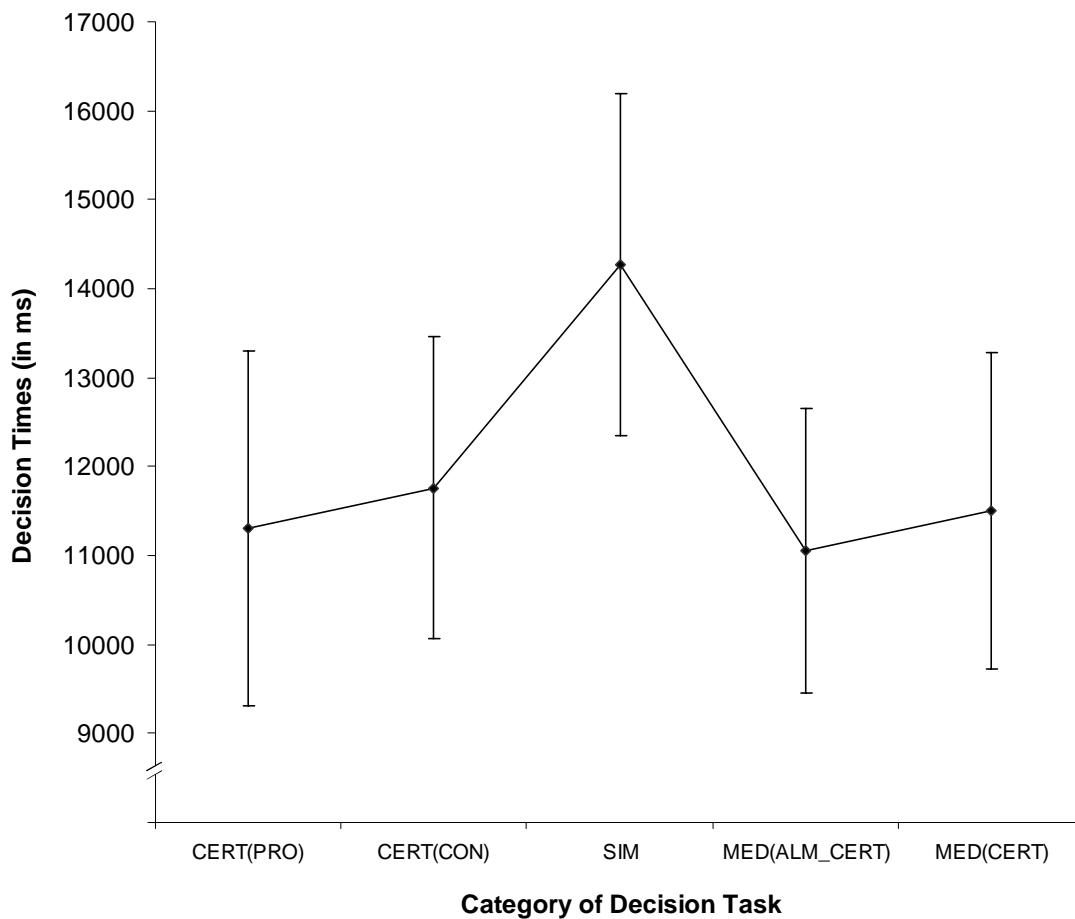
Figure 7.
Decision times in Experiment 3, with error bars indicating 95% confidence intervals.

Information search. On average, the participants investigated 9.2 pieces of information (*SE* = .46) in the decision phase of each task. Note that only eight pieces of information were available. Thus, on average, each of the pieces of information was inspected more than once. This finding is in line with the prediction of CPT that all information is taken into account, but the finding

conflicts with the prediction of the PH that a part of the information is ignored. According to the PH, it would be expected that only the two minimum gains are inspected in the categories $CERT_{PRO}$, SIM, MED_{ALM_CERT} and MED_{CERT} (i.e., 2 pieces of information), and in the category $CERT_{CON}$ minimum gains and probabilities of minimum gains (i.e., 4 pieces of information) are looked up. Therefore, the PH predicts an average information acquisition rate of 3.2. A t test revealed that the overall information acquisition rate was significantly higher than this predicted acquisition rate, $t(31) = 13.0$, $p < .001$. Thus, the H_5 – namely that individuals use a frugal strategy of information search as suggested by the PH – was rejected.

According to the PH, probabilities should only be inspected if gains do not differ (i.e., in category $CERT_{CON}$), which should result in a much lower overall inspection rate for probabilities than for gains, whereas CPT would predict approximately equal inspection rates. The observed inspection rates were only slightly higher for gains ($M=5.0$) than for probabilities ($M=4.2$), which is more in line with the prediction of CPT.

To analyze the direction of information search and to test the H_6, which proposes that an information search should be carried out mainly between gambles and less often within gambles, a CPT-transition index and a PH-transition index were computed. The CPT-transition index was computed by counting all information search transitions within gambles, and the PH-transition index was computed by counting all direct transitions between gains or probabilities of both gambles. A multivariate analysis of variance with the two scores as dependent variables revealed significant differences between both scores, *Pillais V* = .91, $F(1, 32) = 332.9$, $p < .001$. In line with findings by Johnson et al. (2008), individuals mainly searched information within gambles ($M_{CPT-Index} = 4.80$, $SE = 0.26$) and only very few transitions between gambles were observed ($M_{PH-Index} = 0.49$, $SE = 0.05$). Thus, the H_6, which proposes that individuals compare information between gambles by carrying out information search transitions between gambles, has to be clearly rejected. In line with the predictions of the CPT and in line with earlier findings, information is mainly searched within gambles.

Discussion

The third experiment aimed at testing the PH by investigating hypotheses concerning information search behavior. The analysis showed that information search was not frugal but exhaustive and that information searches mainly occur within gambles and not in a direct comparison of gains and losses between gambles. Both findings are in line with the predictions of CPT. The two hypotheses derived from the PH had to be rejected. Individuals do not use frugal information search strategies that ignore large parts of the information (H_5), and information search is not carried out by comparing gains and probabilities between gambles (H_6).

The results strengthen findings by Johnson et al. (2008) that individuals' information search does not align with the predictions of the PH. In our experiment some methodological caveats of the experiments by Johnson et al. could be ruled out by three methodological modifications. First,

we used an additional explicit screening instruction to allow individuals to quickly look at the available information and to estimate aspiration levels (cf. step 0 of the PH). This could make it easier to apply the PH. Second, we introduced a more clear-cut differentiation between the screening and decision phases, by asking people to click on the start button after screening the information. This circumvents the somewhat artificial split of information search used by Johnson et al. in a "reading phase" and a "choice phase," based on individually observed data. Johnson et al. considered the information search to belong to a reading phase until all pieces of information were at least inspected once; thereafter, information search was considered to belong to the choice phase. If individuals, for instance, were not interested in one piece of information, all information search movements were considered to belong to the reading phase. Third, the application of the open mouselab method reduced the likelihood that the decision strategy would be influenced by the research method, as compared to the standard mouselab used by Johnson et al. (cf. Glöckner & Betsch, 2007). The convergence of the results of Experiment 3 with the results of Experiment 1 indicates that decision strategies were not severely influenced by the open mouselab method.

However, taking a methodologically critical perspective, in our experiment we cannot rule out the possibility that participants did not obey the instructions and that information search did not correspond to the recorded mouse movements. Note, however, that the main focus of the study was to test the hypotheses derived by the PH, and (aside from converging evidence of choices and decision times) it seems rather unlikely that individuals first applied a frugal information search strategy and afterwards moved the mouse to all available pieces of information. Furthermore, we cannot rule out that the instruction to start decision making with a screening of all information induced exhaustive instead of limited information search and thus led to a rejection of H_5. Note, however, that the high resemblance of choice data and decision times with Experiment 1 speaks against a substantial influence of the additional instruction on choice behavior.

General Discussion

In this paper, we first discussed empirical findings concerning the Priority Heuristic that has been proposed as a fast-and-frugal heuristic for risky choices (Brandstätter et al., 2006) and came to the conclusion that the evidence was equivocal. A closer investigation of the simulations used by Brandstätter et al. (2006) led us to suspect that the major findings supporting the PH are based on two facts: first, that decision tasks used in the simulations were not diagnostic for the PH as compared to CPT; and second, that only aggregated choices were analyzed. Therefore, we selected diagnostic decision tasks in our experiments and analyzed individual choices. We derived six specific hypotheses from the PH to conduct a critical empirical test of the model against CPT. All six hypotheses were rejected, and in all cases decision behavior was more aligned with the predictions of CPT than the PH.

It was possible to demonstrate that aggregated choice data deviated significantly from the predictions of the PH if diagnostic decision tasks are included. Particularly, the first PH hypothesis was

rejected, which proposed that the proportion of choices in line with the predictions of the PH should be approximately equal for different decision tasks. In contrast, it was found that choices were mainly in line with the PH for those categories of decision tasks in which the predictions of the PH were aligned with the predictions of CPT, whereas low proportions of choices aligned with the predictions of the PH if this was not the case. For instance, it could be shown that in the category of decision tasks in which CPT predicts a certainty effect but the PH predicts that the certainty effect should disappear (i.e., $CERT_{CON}$), choice proportions in line with the PH were very low, whereas these proportions were very high in cases in which the predictions of CPT and the PH aligned (i.e., $CERT_{PRO}$).

The second PH hypothesis stated that individual choices can be best explained by the PH. In contrast to this hypothesis, it was found that for the majority of the participants individual choice patterns can be best explained by CPT (i.e., Exp. 1: 78%, Exp. 2: 69%). Only for a small minority of the participants (i.e., Exp. 1: 9%, Exp. 2: 5%) were individual choices roughly in line with the predictions of the PH. Overall, it was found that the predictive accuracy of the PH mainly depended on the overlap in choice predictions with CPT. We reduced this overlap by using diagnostic decision tasks and found that the predictive accuracy of the PH for individual choices dropped almost to the level of chance (i.e., Exp. 1: 59%, Exp. 2: 53%, Exp. 3: 60%).

The third PH hypothesis stated that decision times should increase with an increasing number of reasons (i.e., steps) that have to be considered when applying the PH. This hypothesis also had to be rejected because decision time did not depend on this factor. In line with classic (Cartwright, 1941; Cartwright & Festinger, 1943; Festinger, 1943a; 1943b) and more recent findings (Bergert & Nosofsky, 2007; Glöckner & Betsch, 2007; Glöckner, 2007), decision times seem to increase with the increase in the similarity of the expected utilities of the gambles. These findings conflict with the findings reported by Brandstätter et al. (2006). As discussed above, we suspect that the fact that Brandstätter et al. (2006) did not control for the similarity of expected utilities in their experiment could have caused the different results.

The fourth PH hypothesis postulated that changes in minimum gains and probabilities of minimum gains that are below or above reasonable aspiration levels (i.e., 1/10 and 1/3) should not influence choices. Based on the findings of Experiment 2, this hypothesis was also rejected. The manipulation of differences in minimum gains and probabilities of minimum gains influenced choice probabilities also outside of the considered aspiration levels. In line with the predictions of CPT, choice functions were descriptively more aligned with a linear than with a step function. Thus, overall, evidence indicated that the reasons postulated by the PH are not considered in a non-compensatory manner using aspiration levels.

The fifth PH hypothesis was that information search should be frugal, in that a large part of the information should be ignored. Experiment 3 showed that this hypothesis is not supported by the data. In line with experiments on probabilistic inference tasks (Glöckner, 2006; Glöckner & Betsch, 2007; Newell, Weston, & Shanks, 2003; Newell & Shanks, 2003; cf. Bröder, 2003), on

average individuals inspected all pieces of information at least once and did not truncate the information search according to the stopping rule proposed by the PH.

The sixth PH hypothesis proposed that individuals should mainly search information by directly comparing gains and probabilities between gambles rather than by carrying out information search within each gamble. This hypothesis was also clearly rejected. In line with the predictions of CPT, information search was dominantly carried out within gambles and only a few transitions between gambles were observed.

In sum, the findings concerning all six hypotheses converge in showing that individuals do not apply the PH, and decision behavior concerning choices, decision times and information search can be better explained by CPT. In contrast to the provoking proposition by Brandstätter et al. (2006) – that individuals use gains and probabilities in risky decisions in a non-compensatory manner – the majority of individuals seem to make trade-offs between probabilities and gains as proposed by CPT.

The results highlight one fundamental risk of the simulation methods often cited to support fast-and-frugal heuristics (e.g., Czerlinsky, Gigerenzer, & Goldstein, 1999; Gigerenzer & Goldstein, 1996). As we tried to demonstrate in this paper, it could be easily misleading to argue that a fast-and-frugal heuristic is a better model for individuals' decision making than a more complex compensatory model merely based on the facts that (a) the heuristic is able to predict choices in a limited sample of decision tasks equally well and (b) that the decision strategy has fewer adjustable parameters and therefore receives more support than the more complex model (cf. Occam's Razor argument; Brandstätter et al., 2006). A sound test of heuristics has to include sufficient diagnostic decision tasks that allow differentiation between decision strategies.

The results reported in this paper are in line with recent findings in research on probabilistic decision tasks that individuals have the computational capacity quickly to make decisions that approximate weighted additive information integration (Glöckner & Betsch, 2007) and even use this as a default strategy (Bröder, 2003). In Experiment 1, most of the participants integrated information in line with CPT within 5.5 seconds (on average). Note that the deliberate calculation of choices according to CPT would necessitate that the participants carry out two multiplications of probabilities and gains for each gamble, add the products up for each gamble and compare the sums. Considering the fact that many of the gains and probabilities were no simple multiples of 10 or .10, it is rather unlikely that these calculations could have been carried out deliberately in such a short time. As outlined in more detail elsewhere (Glöckner & Betsch, in press; Glöckner, 2007), we argue that individuals apply decision strategies that are based on automatic processes and are part of the intuitive system instead (Kahneman & Frederick, 2002; cf. Epstein, 1990; Hogarth, 2001; Petty & Cacioppo, 1986; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977; Sloman, 2002; Strack & Deutsch, 2004). We put forward that these automatic processes enable individuals in risky decisions and in probabilistic inference decisions to quickly come to

choices that approximate weighted compensatory information integration (cf. Hammond, Hamm, Grassia, & Pearson, 1987).[8]

Our results indicate that CPT is a well-suited paramorphic model (cf. Hoffman, 1960; also called "as-if model"; Brandstätter et al., 2006) that describes individuals' choices in gambling decision tasks, even without fitting parameters to individually observed data. Nevertheless, we agree with Brandstätter et al. (2006) that there is a need for more precise process models. However, according to our data, non-compensatory heuristics like the PH are not good candidates for explaining the processes underlying risky decisions. Considering the findings that individuals are able quickly to integrate information in a weighted compensatory manner and that decision time increases with the increasing similarity of the expected utilities of the gambles, evidence accumulation models (Busemeyer & Townsend, 1993; Busemeyer & Johnson, 2004; Gold & Shadlen, 2007; Lee & Cummins, 2005; Newell, 2005) and parallel constraint satisfaction network models (Holyoak & Simon, 1999; Glöckner, 2006; Glöckner & Betsch, in press; Simon, Krawczyk, & Holyoak, 2004; Thagard & Millgram, 1995) might be good candidates for explaining these processes. However, further research will be needed to improve our understanding of them.

---

8   Converging evidence for this view is provided by research that shows that even lower animals that are not able to use deliberate processing in decision making (i.e., sticklebacks), select mating partners based on a weighted compensatory information integration (Künzler & Bakker, 2001; see also Glimcher, Dorris, Bayer, 2005, Real, 1991, and Glöckner, in press).

References

Allais, M. (1953/1979). La psychologie de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine. *Econometrica, 21*, 503-46. Translated and reprinted in Allais & Hagen, 1979.

Bergert, F. B., & Nosofsky, R. M. (2007). A Response-Time Approach to Comparing Generalized Rational and Take-the-Best Models of Decision Making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(1), 107-129.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The Priority Heuristic: Making Choices Without Trade-Offs. *Psychological Review, 113*(2), 409-432.

Bröder, A. (2003). Decision making with the "adaptive toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 611-625.

Bröder, A., & Schiffer, S. (2003). Bayesian strategy assessment in multi-attribute decision making. *Journal of Behavioral Decision Making, 16*(3), 193-213.

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*(2), 277-293.

Bröder, A. (2005). *Entscheiden mit der „Adaptiven Werkzeugkiste": Ein empirisches Forschungsprogramm* [Decision making with the adaptive toolbox: An empirical research program]. Lengerich, Germany: Pabst Science Publications.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100*(3), 432-459.

Busemeyer, J. R., & Johnson, J. G. (2004). Computational models of decision making. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 133-154). Malden, MA: Blackwell Publishing.

Cartwright, D. (1941). Decision-time in relation to the differentiation of the phenomenal field. *Psychological Review, 48*(5), 425-442.

Cartwright, D., & Festinger, L. (1943). A quantitative theory of decision. *Psychological Review, 50*(6), 595-621.

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In *Simple heuristics that make us smart* (pp. 97-118). New York, NY: Oxford University Press.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 51*(4), 380-417.

Erev, I., Roth, A. E., Slonim, R. L., & Barron, G. (2002). *Combining a theoretical prediction with experimental evidence to yield a new prediction: An experimental design with a random sample of tasks.* Unpublished manuscript, Columbia University and Faculty of Industrial Engineering and Management, Techion, Haifa, Israel.

Epstein, S. (1990). Cognitive-experiential self-theory. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 165-192). New York, NY: Guilford Press.

Festinger, L. (1943). Studies in decision: I. Decision-time, relative frequency of judgment and subjective confidence as related to physical stimulus difference. *Journal of Experimental Psychology, 32*(4), 291-306.

Festinger, L. (1943). Studies in decision. II. An empirical test of a quantitative theory of decision. *Journal of Experimental Psychology, 32*(5), 411-423.

Fishburn, P. C. (1974). Lexicographic orders, utilities, and decision rules: A survey. *Management Science, 20*, 1442-1472.

Fishburn, P. C., & Kochenberger, G. A. (1979). Two-piece from von Neumann-Morgenstern utility function. *Decision Science, 10,* 503-518.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review, 103*(3), 592-596.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*(4), 650-669.

Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart. Evolution and cognition*. Simple heuristics that make us smart. xv, 416 pp. New York, NY: Oxford University Press.

Glimcher, P. W., Dorris, M. C., & Bayer, H. M. (2005). Physiological utility theory and the neuroeconomics of choice. *Games and Economic Behavior, 52*(2), 213-256.

Glöckner, A. (2006). *Automatische Prozesse bei Entscheidungen* [Automatic processes in decision making]. Hamburg: Kovac.

Glöckner, A., & Betsch, T. (2007). Multiple-reason decision making based on automatic processing. *Manuscript submitted for publication*.

Glöckner, A., & Betsch, T. (in press). Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making. *Judgment and Decision Making*.

Glöckner, A. (2007). Does intuition beat fast and frugal heuristics? A systematic empirical analysis. In H. Plessner, C. Betsch & T. Betsch (Eds.), *Intuition in judgment and decision making* (pp. 309–325). Mahwah, NJ: Lawrence Erlbaum.

Glöckner, A. (in press). How evolution outwits bounded rationality: The efficient interaction of automatic and deliberate processes in decision making and implications for institutions. In C. Engel & W. Singer (Eds.), *Better Than Conscious? Implications for Performance and Institutional Analysis. Strüngmann Forum Report 1*. Cambridge, MA: MIT Press.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 50,* 535-74.

Hammond, K. R., Hamm, R. M., Grassia, J., & Pearson, T. (1987). Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. *IEEE Transactions on Systems, Man, & Cybernetics, 17*(5), 753-770.

Hogarth, R. M. (2001). *Educating intuition.* Educating intuition. xii, 335 pp. Chicago, IL: University of Chicago Press.

Hoffman, P. J. (1960). The paramorphic representation of clinical judgment. *Psychological Bulletin, 57*(2), 116-131.

Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General, 128*(1), 3-31.

Johnson, E. J., Schulte-Mecklenbeck, M., & Willemsen, M. C. (2008). Process models deserve process data: Comment on Brandstatter, Gigerenzer, and Hertwig (2006). *Psychological Review, 115*(1), 263-272.

Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263-292.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49-81). New York, NY: Cambridge University Press.

Künzler, R., & Bakker, T. C. M. (2001). Female preferences for single and combined traits in computer animated stickleback males. *Behavioral Ecology, 12*(6), 681-685.

Lee, M. D., & Cummins, T. D. R. (2004). Evidence accumulation in decision making: Unifying the "take the best" and the "rational" models. *Psychonomic Bulletin & Review, 11*(2), 343-352.

Lohse, G. L., & Johnson, E. J. (1996). A comparison of two process tracing methods for choice tasks. *Organizational Behavior and Human Decision Processes, 68*(1), 28-43.

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, and Cognition, 29*(1), 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*(1), 82-96.

Newell, B. R. (2005). Re-Visions of rationality? *Trends in Cognitive Sciences, 9*(1), 11-15.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 534-552.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 123-205.

Real, L. A. (1991). Animal choice behavior and the evolution of cognitive architecture. *Science, 253,* 980-886.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*(1), 1-66.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*(2), 127-190.

Simon, D., Krawczyk, D. C., & Holyoak, K. J. (2004). Construction of Preferences by Constraint Satisfaction. *Psychological Science, 15*(5), 331-336.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69*, 99-118.

Sloman, S. A. (2002). Two systems of reasoning. In T. Gilovich, D. Griffin & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 379-396). New York, NY: Cambridge University Press ISBN 0-521-79260-6 (hardcover); 0-521-79679-2 (paperback) Cambridge University Press Print.

Strack, F., & Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior. *Personality and Social Psychology Review, 8*(3), 220-247.

Thagard, P. & Millgram, E. (1995). Inference to the best plan: A coherence theory of decision. In A. Ram & D. B. Leake (Eds.), *Goal-driven learning* (pp. 439-454). Cambridge, MA: MIT Press.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review, 76*(1), 31-48.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5,* 297–323.

Tversky, A., & Fox, C. R. (1995). Weighing risk and uncertainty. *Psychological Review, 102*(2), 269-283.

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior.* Princeton, NJ: Princeton University Press.

Savage, L. J. (1954). *The foundations of statistics* (2nd ed.). New York: Dover.

Wasserman, L. (2000). Bayesian model selection and model averaging. *Journal of Mathematical Psychology, 44*(1), 92-107.

Appendix A

Decision Tasks Experiment 1 and 3

The following table shows the 40 decision tasks used in Experiments 1 and 3. Each decision task consists of two gambles A and B (columns), two possible outcomes (Out 1 and Out 2) and their probabilities (rows). The five main columns represent the different categories of decision task. For each category the eight different decision tasks are presented in the rows. The expected value for each pair of gambles is approximately the same. All gambles are coded, so that gamble A is predicted by the PH and, in each gamble, outcome 1 is lower than outcome 2.

Category of Decision Tasks

	CERT_PRO Gamble A		Gamble B		CERT_CON Gamble A		Gamble B		SIM Gamble A		Gamble B		MED_ALM_CERT Gamble A		Gamble B		MED_CERT Gamble A		Gamble B	
Out 1	€40	.40	€0	.60	€0	.40	€10	.60	€50	.60	€40	.50	€55	.90	€0	.01	€146	.98	€67	.85
Out 2	€80	.60	€160	.40	€160	.60	€225	.40	€60	.40	€69	.50	€500	.10	€100	.99	€160	.02	€600	.15
Out 1	€30	.50	€0	.30	€0	.30	€5	.50	€30	.70	€20	.60	€67	.85	€0	.02	€83	.98	€38	.85
Out 2	€40	.50	€50	.70	€90	.70	€120	.50	€40	.30	€54	.40	€600	.15	€150	.98	€100	.02	€340	.15
Out 1	€10	.45	€0	.40	€0	.40	€8	.55	€60	.55	€76	.40	€27	.90	€0	.01	€150	.98	€72	.85
Out 2	€40	.55	€44	.60	€100	.60	€125	.45	€100	.45	€80	.60	€250	.10	€50	.99	€200	.02	€600	.15
Out 1	€40	.30	€0	.20	€0	.20	€6	.30	€40	.25	€60	.50	€181	.90	€0	.02	€55	.99	€28	.80
Out 2	€100	.70	€103	.80	€103	.80	€115	.70	€80	.75	€80	.50	€1800	.10	€350	.98	€100	.01	€165	.20
Out 1	€40	.50	€0	.25	€0	.25	€6	.50	€75	.50	€50	.35	€35	.85	€0	.02	€215	.98	€50	.85
Out 2	€70	.50	€73	.75	€90	.75	€130	.50	€89	.50	€100	.65	€340	.15	€83	.98	€230	.02	€1150	.15
Out 1	€30	.40	€0	.30	€0	.30	€4	.60	€50	.50	€40	.65	€67	.85	€5	.02	€300	.99	€50	.85
Out 2	€50	.60	€60	.70	€45	.70	€72	.40	€100	.50	€140	.35	€600	.15	€150	.98	€320	.01	€1720	.15
Out 1	€20	.65	€0	.50	€0	.50	€7	.65	€30	.70	€20	.60	€28	.90	€2	.01	€460	.99	€60	.80
Out 2	€70	.35	€75	.50	€80	.50	€100	.35	€40	.30	€54	.40	€250	.10	€50	.99	€500	.01	€2060	.20
Out 1	€45	.60	€0	.33	€0	.33	€5	.60	€10	.50	€5	.60	€111	.90	€0	.02	€750	.99	€90	.80
Out 2	€65	.40	€80	.67	€90	.67	€145	.40	€33	.50	€47	.40	€1100	.10	€215	.98	€800	.01	€3390	.20

Appendix B

The conditional likelihood L_k for a data set (individual set of choices) under the condition that a certain decision strategy k was applied with a constant error rate ε_k is given by

$$L_k = p(n_{jk}|k, \varepsilon_k) = \prod_{j=1}^{J} \binom{n_j}{n_{jk}} (1-\varepsilon_k)^{n_{jk}} \varepsilon_k^{(n_j - n_{jk})},$$

with J being the total number of categories of decision tasks, n_j being the number of choices in category j, and n_{jk} being the number of choices in line with the prediction of strategy k in category j. The error rate ε_k for each decision strategy is estimated by

$$\hat{\varepsilon}_k = \left[\sum_{j=1}^{J}(n_j - n_{jk})\right] \div \left[\sum_{j=1}^{J} n_j\right].$$

For each participant, the conditional likelihoods L_k for all strategies were computed and the resulting likelihood values were compared. Individuals were classified as users of a certain strategy if the conditional likelihood for this strategy was higher than the conditional likelihood for the remaining strategies. Furthermore, it was necessary that the conditional likelihood of the best strategy was at least larger than the sum of the conditional likelihoods for the remaining strategies. If this was not the case, classification was considered unreliable, and no classification was carried out. If the best strategy had an error rate larger than .50 (i.e., when the predictive accuracy was below chance level), no classification was carried out either.

Appendix C

Decisions Tasks Experiment 2

Decision tasks in Experiment 2 were categorized in five sets (main rows). In each set, the gambles A (left main column) were held constant. The gambles B were manipulated on six to twelve levels (right main column). Each of the columns depicts a version of gamble B that was paired with the respective fix gamble A of the set. All gambles have expected values of €100.

	Manipulation of Gamble B											
Gamble A	1	2	3	4	5	6	7	8	9	10	11	12
Set 1												
.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50
€90	€85	€80	€75	€70	€65	€60	€40	€20	€10	€5	€1	€0
.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50
€110	€115	€120	€125	€130	€135	€140	€160	€180	€190	€195	€199	€200
Set 2												
.90	.99	.98	.95	.90	.80	.60	.50	.20				
€10	€0	€0	€0	€0	€0	€0	€0	€0				
.10	.01	.02	.05	.10	.20	.40	.50	.80				
€910	€10000	€5000	€2000	€1000	€500	€250	€200	€125				
Set 3												
.90	.85	.80	.75	.70	.60	.50	.40	.10				
€10	€10	€10	€10	€10	€10	€10	€10	€10				
.10	.15	.20	.25	.30	.40	.50	.60	.90				
€910	€610	€460	€370	€310	€235	€190	€160	€110				
Set 4												
.70	.50	.50	.50	.50	.50	.50						
€70	€65	€60	€55	€50	€45	€40						
.30	.50	.50	.50	.50	.50	.50						
€170	€135	€140	€145	€150	€155	€160						
Set 5												
.70	.90	.90	.90	.90	.90	.90						
€70	€60	€50	€40	€20	€10	€0						
.30	.10	.10	.10	.10	.10	.10						
€170	€460	€550	€640	€820	€910	€1000						