



# How the twain can meet: Prospect theory and models of heuristics in risky choice



Thorsten Pachur\*, Renata S. Suter, Ralph Hertwig

Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany

## ARTICLE INFO

### Article history:

Accepted 8 January 2017

### Keywords:

Risky choice  
Prospect theory  
Theory integration  
Heuristics  
Cognitive modeling

## ABSTRACT

Two influential approaches to modeling choice between risky options are algebraic models (which focus on predicting the overt decisions) and models of heuristics (which are also concerned with capturing the underlying cognitive process). Because they rest on fundamentally different assumptions and algorithms, the two approaches are usually treated as antithetical, or even incommensurable. Drawing on cumulative prospect theory (CPT; Tversky & Kahneman, 1992) as the currently most influential instance of a descriptive algebraic model, we demonstrate how the two modeling traditions can be linked. CPT's algebraic functions characterize choices in terms of psychophysical (diminishing sensitivity to probabilities and outcomes) as well as psychological (risk aversion and loss aversion) constructs. Models of heuristics characterize choices as rooted in simple information-processing principles such as lexicographic and limited search. In computer simulations, we estimated CPT's parameters for choices produced by various heuristics. The resulting CPT parameter profiles portray each of the choice-generating heuristics in psychologically meaningful ways—capturing, for instance, differences in how the heuristics process probability information. Furthermore, CPT parameters can reflect a key property of many heuristics, lexicographic search, and track the environment-dependent behavior of heuristics. Finally, we show, both in an empirical and a model recovery study, how CPT parameter profiles can be used to detect the operation of heuristics. We also address the limits of CPT's ability to capture choices produced by heuristics. Our results highlight an untapped potential of CPT as a measurement tool to characterize the information processing underlying risky choice.

© 2017 Elsevier Inc. All rights reserved.

## 1. Introduction

When in 2002 the Royal Swedish Academy of Sciences awarded Daniel Kahneman the Nobel Prize for Economic Sciences, it lauded two pillars of his joint work with Amos Tversky: prospect theory and the heuristics-and-biases program. Both lines of research investigate how human judgment and decision making systematically deviate from standard assumptions of rationality in economics. Yet there is also a fascinating disconnect between the two lines of research. Kahneman pointed out that when developing prospect theory, their approach “was as conservative as possible. [...] The goal [...] was to assemble the minimal set of modifications of expected utility theory that would provide a descriptive account of [...] choices between simple monetary gambles” (Kahneman & Tversky, 2000, p. x). The theoretical thrust of the heuristics-and-biases

\* Corresponding author at: Center for Adaptive Rationality, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany.  
E-mail address: [pachur@mpib-berlin.mpg.de](mailto:pachur@mpib-berlin.mpg.de) (T. Pachur).

program, in contrast, was anything but conservative; in fact, it “revolutionized academic research on human judgment” (Gilovich & Griffin, 2002, p. 1). It did so by proposing that people’s judgments under uncertainty do not involve extensive operations, but rest on simple mental processes—*heuristic in nature*—that curtail information search.

The disconnect between Kahneman and Tversky’s prospect theory and their work on heuristics (e.g., Tversky & Kahneman, 1974) echoes a more general trend in research on risk choice, namely the separation between algebraic models and models focused on the actual processes underlying choice (cf. Lopes, 1995). Algebraic models—with prospect theory and its extension, cumulative prospect theory (CPT; Tversky & Kahneman, 1992), being the most prominent instances—focus on the algebraic patterns of people’s risk preferences and thus on people’s ultimate choices but not on *how* the choices come about. They describe choices in terms of a multiplicative integration of (some function of) probabilities and outcomes, followed by maximization, and are rooted in the concept of mathematical expectation and its application to risky choice, first in terms of expected value theory and then in the early incarnation of expected utility theory (Bernoulli, 1738/1954). Like other algebraic models of risky choice (Birnbaum & Chavez, 1997; Lopes & Oden, 1999), CPT assumes that the evaluation of a risky option can be derived by combining mathematical expectation with a few psychological constructs (e.g., risk aversion, loss aversion, probability weighting), all of which are represented by parameterized functions.

The other approach to modeling risky choice takes Edwards’ (1953) question as its starting point: “[H]ow do people actually go about making decisions in gambling situations?” (p. 351). Informed by findings from process analyses (e.g., Lichtenstein, 1965; Payne & Braustein, 1978) and acknowledging cognition’s natural bounds of information and computational capacities (Simon, 1955), models of the choice process have often taken the form of heuristics. Unlike algebraic models, models of heuristics focus on the content (rather than just the result) of choice processes and are explicated in terms of the cognitive operations underlying a decision, such as search, stopping, and integration of information (Gigerenzer, Hertwig, & Pachur, 2011; Payne, Bettman, & Johnson, 1993). Being based on simple principles of information processing, models of heuristics typically forego the integration of probabilities and outcomes and ignore part of the available information (e.g., Payne et al., 1993; Thorngate, 1980; for an overview, see Pleskac, Diederich, & Wallsten, 2015). Models of heuristics for risky choice include the semiorde rule (Luce, 1956), elimination-by-aspects (Tversky, 1969, 1972), the minimax and maximax heuristics (e.g., Coombs, Dawes, & Tversky, 1970), the Pwin heuristic (Payne, 2005), the least-likely and the most-likely heuristics (Thorngate, 1980), the similarity heuristic (Leland, 1998; Rubinstein, 1988), and the priority heuristic (Brandstätter, Gigerenzer, & Hertwig, 2006). Although some of these models of heuristics have been formalized mathematically (e.g., Katsikopoulos & Gigerenzer, 2008; Tversky, 1972), we do not classify them as algebraic models here, thus following Lopes (1995). We refer to algebraic models as those that focus on predicting people’s overt choices (regardless of the psychological plausibility of the implied process), whereas models of heuristics also aim to specify how information is searched and when search is stopped.<sup>1</sup>

The two modeling approaches to risky choice exist side by side with little theoretical integration, and they have often been treated as antithetical (e.g., Brandstätter et al., 2006; Payne, 1973; Svenson, 1979). Lopes (1995, p. 177) described this separation thus: “Although one might suppose that [algebraic and process] accounts are alternate ways of describing the same thing—indeed, that one kind of model might eventually be reducible to the other—the approaches have tended to be disjoint.” The goal of this article is to investigate to what extent algebraic models and models of heuristics represent alternative ways of describing the same thing—or, at least, related things. We argue that CPT’s constructs—such as diminished sensitivity to outcome or probability information, risk aversion, and loss aversion, measured through parameters that govern the shape of CPT’s value and weighting functions (see below)—offer a useful framework for characterizing heuristic processing principles. Specifically, we demonstrate how differences in CPT parameters can index the use of different heuristic processes.

To avoid misunderstandings, let us clarify two issues right away. First, our goal is not to render one approach reducible to another (a possibility suggested in the quote from Lopes, 1995). Rather, we aim to work out how two possible but disjointed ways to describing risky choice could be brought together. We thus join forces with others who have contributed to overcome the fragmented state of psychological theory (see Dougherty, Gettys, & Ogden, 1999; Johnson & Meyer, 1984; Juslin, Olsson, & Olsson, 2003; Luan, Schooler, & Gigerenzer, 2011; Pleskac, 2007; Schooler & Hertwig, 2005). In a similar vein, Stewart, Chater, and Brown (2006) have demonstrated how the characteristic shapes of CPT’s value and weighting functions can arise from an interaction of simple cognitive comparison processes and particular structures of the environment, namely, the distribution of how frequently gains and losses of different magnitudes occur in the ecology. We complement this work by investigating how CPT’s shapes may arise from specific characteristics in the processing principles of heuristics, but without making particular assumptions about the choice ecology. Second, for the purpose of our analysis, we consider the models of heuristics investigated here as stylized descriptions of simple processing principles that may operate to varying degrees, depending on the decision context and the decision maker. We do not claim that any one of them offers a comprehensive and veridical account of risky choice (we turn to empirical tests of models of heuristics in the General Discussion).

In what follows, we review previous work connecting algebraic and heuristic models of decision making, and describe CPT’s parametric framework as well as several heuristics for risky choice proposed in the literature. We then discuss how choices generated by several heuristics assuming substantially different processes may be reflected in the parameters of

<sup>1</sup> Some of the models of heuristic investigated in our analyses, such as the minimax heuristic, were not originally proposed as process models (Savage, 1954/1972). However, for all models of heuristics, it is straightforward to derive predictions of how information is searched, when search is stopped, and how the choice is made based on the information—our key criteria for whether a model constitutes a process model.

CPT and test these predictions using computer simulations. Subsequently, we demonstrate how CPT reflects choices of a specific lexicographic strategy that adapts its information search depending on the structure of the choice problems, and that CPT can track how the behavior of a heuristic interacts with the structure of the environment. Finally, we examine, using a model recovery study, how well heuristics and CPT can be told apart under conditions of noise; and we demonstrate, both in an empirical study and another model recovery study, how the specific CPT parameter profiles of the individual heuristics can be used to infer the operation of the heuristics.

### 1.1. Algebraic models as measurement tools to reflect information processing

Others have previously explored the possibility that algebraic models can be used, by way of their adjustable parameters, as psychometric (i.e., measurement) tools to describe heuristic information processing. For instance, [Johnson and Meyer \(1984\)](#) investigated people's decisions in various multi-attribute choice sets that differed in the number of attributes per option. As the number of attributes per option increased, sequential, noncompensatory elimination strategies (as identified on the basis of verbal protocol analyses and process tracing) became more prevalent. The authors then fitted logit models (which are based on a compensatory algorithm) to the decisions, and showed that the shift toward noncompensatory strategies was captured in the model parameters by a more polarized distribution of attribute weights. Relatedly, [Bhatia \(2015\)](#) showed how various heuristics for multi-attribute choice (e.g., lexicographic, tallying) can be represented in multi-parameter accumulator models. Another example is the use of structural modeling analyses in social judgment theory (e.g., [Cooksey, 1996](#); [Hammond, Hursch, & Todd, 1964](#)). Here, researchers use multiple regression to measure the influence of cues on a person's judgment and, by extension, her judgment "policy," also permitting that policy to be a noncompensatory heuristic ([Gigerenzer & Kurz, 2001](#)).

In our work, we examine the relationship between algebraic models (using CPT as a paradigmatic case) and heuristic choice rules in the context of risky choice, a domain in which the potential of algebraic models as measurement tools for heuristic information processes has barely been examined.

### 1.2. The algebra of risky choice: cumulative prospect theory

CPT assumes that choices between risky options, each having outcomes  $x_m > \dots > x_1 \geq 0 > y_1 > \dots > y_n$  and corresponding probabilities  $p_m \dots p_1$  and  $q_1 \dots q_n$ , are based on a person's subjective valuation of these options, followed by maximization. In CPT, the overall valuation  $V$  of an option  $A$  is defined as

$$V(A) = \sum_{i=1}^m v(x_i) \pi_i^+ + \sum_{j=1}^n v(y_j) \pi_j^- \quad (1)$$

The subjective value,  $v$ , of an outcome is determined by the *value function*:

$$\begin{aligned} v(x) &= x^{\alpha^+}, \\ v(y) &= -\lambda(-y)^{\alpha^-} \end{aligned} \quad (2)$$

where  $\alpha^+$  and  $\alpha^-$  reflect the sensitivity to differences in positive (gains) and negative (losses) outcomes, respectively. Both parameters are assumed to lie in the interval between 0 and 1, yielding a concave and convex value function for gains and losses, respectively. The more concave (convex) the value function, the less sensitive choices are to differences in outcomes. The parameter  $\lambda$  in Eq. (2) indicates the relative weight of gains and losses. It is assumed that a loss of a certain magnitude has a greater psychological impact on a choice than does a gain of the same magnitude. This *loss aversion* occurs if  $\lambda > 1$ , leading to a steeper value function for losses than for gains.

Further, CPT assumes a rank-dependent transformation of the outcomes' probabilities into decision weights. More specifically, the weight  $\pi^+$  ( $\pi^-$ ) given to a positive (negative) outcome is the difference between the probability of receiving an outcome at least as good (bad) as  $x$  ( $y$ ) and the probability of receiving an outcome better (worse) than  $x$  ( $y$ ):

$$\begin{aligned} \pi_m^+ &= w^+(p_m) \\ \pi_n^- &= w^-(q_n) \\ \pi_i^+ &= w^+(p_i + \dots + p_m) - w^+(p_{i+1} + \dots + p_m) \quad \text{for } 1 \leq i < m \\ \pi_j^- &= w^-(q_j + \dots + q_n) - w^-(q_{j+1} + \dots + q_n) \quad \text{for } 1 \leq j < n \end{aligned} \quad (3)$$

The probability *weighting functions* for gains and losses— $w^+$  and  $w^-$ , respectively—are assumed to have an inverse S-shaped curvature, embodying the overweighting of rare events and the underweighting of common events. Several types of weighting functions have been proposed (e.g., [Lattimore, Baker, & Witte, 1992](#); [Prelec, 1998](#); [Tversky & Kahneman, 1992](#); for an overview, see [Stott, 2006](#)). We will use a two-parameter version that separates the curvature of the weighting function from its elevation (e.g., [Goldstein & Einhorn, 1987](#); [Gonzalez & Wu, 1999](#)):

$$w^+ = \frac{\delta^+ p^{\gamma^+}}{\delta^+ p^{\gamma^+} + (1-p)^{\gamma^+}} \quad \text{for } x$$

$$w^- = \frac{\delta^- p^{\gamma^-}}{\delta^- p^{\gamma^-} + (1-p)^{\gamma^-}} \quad \text{for } y$$
(4)

The parameters  $\gamma^+$  and  $\gamma^-$  (both ranging between 0 and 1) govern the function's curvature in the gain and loss domains, respectively, and indicate how sensitive choices are to differences in probability (with smaller values of  $\gamma < 1$  reflecting lower sensitivity). The elevation of the weighting function is controlled by the parameters  $\delta^+$  and  $\delta^-$  (both  $> 0$ ), respectively. As highlighted by [Gonzalez and Wu \(1999\)](#), the elevation reflects the degree of risk aversion (traditionally assumed to be captured by the curvature of the value function; but see [Lopes, 1987, 1995; Wakker, 2010](#)), with a lower (higher) elevation in the gain (loss) domain indicating higher risk aversion.

CPT can account for systematic violations of expected utility theory—such as the Allais paradox, the certainty effect, the reflection effect, and the fourfold pattern of risk attitudes (for an overview, see [Wu, Zhang, & Gonzalez, 2004](#)). It has also successfully been used to both describe and predict risky choice in diverse sets of monetary choice problems (e.g., [Glöckner & Pachur, 2012](#); but see [Birnbbaum, 2004; Brandstätter et al., 2006](#)). As process models, however, algebraic models such as CPT are rarely invoked. Rather, they are often depicted as taking an agnostic stance with regard to the cognitive operations underlying people's choices ([Hoffman, 1960; Johnson & Ratcliff, 2014](#)). Describing the process of choosing has, however, been a key concern for models of heuristics.<sup>2</sup>

### 1.3. Process models of risky choice: boundedly rational heuristics

Many descriptive models of risky choice in the expectation tradition implicitly assume “that the decision maker is capable of evaluating all the available alternatives and examining all the states and the outcomes. In most decisions, however, this assumption is unrealistic. [Instead] different approaches to decision theory, based on some simplification scheme, are called for” ([Coombs et al., 1970](#), pp. 140–141). One breakthrough in the inquiry into how people simplify choices came when [Slovic and Lichtenstein \(1968\)](#) proposed that risky options could be characterized as multidimensional stimuli (see also [Payne, 1973](#)). For instance, a two-outcome gamble offering either a gain or a loss can be described on four basic dimensions (or reasons): the probability of winning, the amount to be won, the probability of losing, and the amount to be lost. Breaking down options into these components raises the possibility that not all reasons receive the same amount of attention, or indeed that some do not even enter the choice process. Consequently, [Tversky \(1969, 1972\)](#), for instance, argued that people may compare the options in a reason-wise fashion, inspecting the reasons sequentially and terminating the process as soon as a choice can be made. In economics, [Rubinstein \(1988\)](#) proposed that when two options are similar in terms of one reason (e.g., probability of winning), attention is shifted to other reasons (e.g., amount to be won).

A frequent feature of models of heuristics is limited search, and many heuristics are therefore *noncompensatory* ([Einhorn, 1970](#)). Limited search is implemented differently in two classes of heuristics. One class consists of strategies that rely on one reason only. For instance, the *minimax heuristic* chooses the option whose worst outcome is better, ignoring both the options' probabilities and the better outcomes ([Savage, 1954/1972](#)). The second class consists of lexicographic choice heuristics. They proceed through several reasons sequentially until they encounter the first one that enables a choice to be made ([Fishburn, 1974; Gigerenzer, Hertwig, & Pachur, 2011; Thorngate, 1980](#)). With lexicographic heuristics, the number of reasons inspected thus depends on the structure of the options (i.e., in terms of which reason the options differ, or differ enough).

Models of heuristics commonly rely on simplifying and psychologically plausible principles of information processing—such as reason-wise, sequential, and limited search, and aspiration levels. Evidence for the operation of many of these principles has been obtained in process-tracing analyses (e.g., [Payne & Braunstein, 1978; Russo & Doshier, 1983; Su et al., 2013](#)), although tests of specific models of heuristics have yielded little support for them as general models of decision making under risk (e.g., [Birnbbaum, 2008; Fiedler, 2010; Glöckner & Pachur, 2012; Johnson, Schulte-Mecklenbeck, & Willemsen, 2008; Stewart, Hermens, & Matthews, 2016](#)). The degree to which heuristic processing principles are triggered seems to depend on context factors (e.g., number of reasons and options, time pressure, affect; [Pachur, Hertwig, & Wolkewitz, 2014; Payne et al., 1993](#)) and can also vary across trials (e.g., [Venkatraman, Payne, Bettman, Luce, & Huettel, 2009](#)).

### 1.4. Heuristics and their alter egos in CPT

To what extent can CPT represent and decode the use of simplifying principles of information processing? CPT's potential to capture heuristic choices may be modest for several reasons. First, CPT and heuristics differ in both the amount of information they consider and how they integrate it. For instance, CPT determines the valuation of each option by integrating all outcomes, each weighted by some function of their probabilities, whereas the minimax heuristic focuses on a single reason, namely its worst outcomes, and ignores all other outcome and probability information. Second, whereas CPT determines

<sup>2</sup> There are also other approaches that focus on modeling the cognitive process leading up to a choice (including processes of search and stopping), such as evidence accumulation models (e.g., [Busemeyer & Townsend, 1993](#); see [Johnson & Ratcliff, 2014](#), for an overview). Although some of these approaches have been successfully applied to risky choice (e.g., [Bhatia, 2014; Glöckner & Herbold, 2011](#)), they are not the topic of this article. These models, while permitting the derivation of process predictions such as response times, often predict the same choices as expectation models.

each option's valuation separately and independently of other options, heuristics rely on reason-wise comparisons between options and gauge the relative rather than absolute attractiveness of an option. Third, lexicographic heuristics assume a sequential processing of the reasons, where the consideration of a specific reason (e.g., the probability of an outcome) depends on another reason (e.g., the magnitude of the outcome). In CPT, in contrast, the consideration of probability information is independent of the associated outcome. Fourth, although some of CPT's constructs, such as probability or outcome sensitivity, seem directly open to a processing-based interpretation—that is, the extent to which a heuristic processes an attribute (e.g., probabilities)—such a link has not yet been directly demonstrated. Finally, CPT is not indefinitely flexible in modeling choice patterns but has various formal and conceptual constraints: Both the value and weighting function must be monotonic. Moreover, the value function is assumed to be concave for gains and convex for losses, cannot be steeper for gains than for losses, and the weighting function is assumed to have an inverse S-shaped curvature. (In the General Discussion, we will discuss results from an implementation of CPT that lifts some of these common constraints.) These properties of CPT may bias which choice patterns it can accommodate and thus limit its ability to represent heuristic choices.

Notwithstanding these arguments, it has been shown that multiplicative and additive information integration can to some extent mimic each other (e.g., Massaro & Friedman, 1990; Stewart, 2011) and that some algebraic models are sufficiently flexible to capture noncompensatory choice rules (e.g., Bhatia, 2015; Johnson & Meyer, 1984). How well can CPT, which is based on a compensatory calculus, represent the choices of noncompensatory heuristics, and how does its parametric framework characterize them? To address these questions, we examined five heuristics that embody categorically different ways of considering probability information: minimax, maximax (Coombs et al., 1970), the priority heuristic (Brandstätter et al., 2006), the least-likely heuristic, and the most-likely heuristic (Payne et al., 1993; Thorngate, 1980).<sup>3</sup> Minimax and maximax completely ignore probabilities; the priority heuristic sometimes considers them, depending on the structure of the choice problem; the least-likely and most-likely heuristics always consider probabilities. Detailed descriptions of the heuristics, and thus also differences in how outcomes are treated, are provided in Table 1. Note that we selected these heuristics on purely conceptual grounds; as stated earlier, we do not claim that any of them is able to provide a comprehensive account of choice under risk.

In the following, we derive predictions concerning how the heuristics' different policies may translate into differences in the curvature of CPT's weighting function as well as in other constructs represented by CPT's parameters, namely, outcome sensitivity, risk aversion, and loss aversion (see Table 2 for a summary of the predictions). Note that although some of CPT's parameter constraints (which implement these constructs) could in principle be lifted—for instance, by also permitting an S-shaped weighting function, a convex (concave) value function for gains (losses), and gain seeking (the opposite of loss aversion)—we will, for conceptual clarity, retain the usual parameter constraints here (cf. Fox & Poldrack, 2014). This has the additional advantage that our results are comparable with existing modeling work on CPT, which has assumed these constraints (e.g., Glöckner & Pachur, 2012; Nilsson, Rieskamp, & Wagenmakers, 2011; Rieskamp, 2008; Suter, Pachur, & Hertwig, 2016). The parameter constraints limit the range of choice behavior that CPT can represent—and as such hamper its flexibility as a measurement tool; in that sense, CPT does not represent a bias-free measurement model. The constraints, however, embody the essence of CPT. Moreover, even with its parameter constraints, CPT is a flexible model that can accommodate a wide range of choice patterns. In the General Discussion, we address how CPT represents the heuristics' choices when parameter constraints are lifted.

#### 1.4.1. Probability sensitivity

In CPT, the degree to which choices are sensitive to differences in probabilities is assumed to be represented by the  $\gamma$  parameter (Eq. (4)), with lower values indicating lower sensitivity. When CPT is fitted to choices produced by the minimax or maximax heuristics, both of which ignore probabilities, the resulting  $\gamma$  parameters should be lower—yielding a more strongly curved weighting function and indicating lower probability sensitivity—than when fitted to choices produced by the priority heuristic. Depending on the choice problem, the latter sometimes takes probabilities into account. By contrast, the least-likely and the most-likely heuristics, both of which always attend to probabilities, should yield values of the  $\gamma$  parameter close to 1, reflecting high sensitivity to probabilities (for a more detailed discussion of these predictions, see Appendix A).

#### 1.4.2. Outcome sensitivity

In CPT, the degree to which choices are sensitive to differences in the magnitude of outcomes is represented by the  $\alpha$  parameter of the value function (Eq. (2)), with lower values indicating lower sensitivity. Although the five heuristics described above were selected because they differ in the processing of probabilities, they also differ in their concern for outcomes. A heuristic that chooses solely on the basis of outcomes (such as minimax or maximax) should display greater outcome sensitivity than one that bases its choices on probabilities (e.g., the least-likely heuristic). Outcomes always enter the process of the minimax, maximax, and most-likely heuristics; the priority heuristic decides based on minimum outcomes as long as they discriminate between options (and otherwise turns to probability information); and the least-likely heuristic never bases a choice on outcome information (although outcomes are considered in an initial step of the decision process;

<sup>3</sup> The most-likely heuristic corresponds to the first step of the lexicographic heuristic (Payne et al., 1993). For the choice problems investigated here, both heuristics almost always make the same predictions. We therefore report the results of the former only.



**Table 1**  
Heuristics tested in the simulations.

Heuristic	Description	Consideration of probabilities
Minimax	Choose the gamble with the highest minimum outcome <sup>a</sup>	Never
Maximax	Choose the gamble with the highest outcome <sup>a</sup>	Never
Priority heuristic	Go through reasons in the following order: minimum gain, probability of minimum gain, and maximum gain. Stop examination if the minimum gains differ by 1/10 (or more) of the maximum gain; otherwise, stop examination if probabilities differ by 1/10 (or more) of the probability scale. Choose the gamble with the more attractive gain (probability). For gambles with more than two outcomes, the search rule is identical, apart from the addition of a fourth reason: probability of maximum gain. For loss gambles, the heuristic remains the same except that “gains” are replaced by “losses.” For mixed gambles, the heuristic remains the same except that “gains” are replaced by “outcomes.”	Sometimes
Least-likely	Identify each gamble’s worst outcome. Then choose the gamble with the lowest probability of the worst outcome <sup>a</sup>	Always
Most-likely	Identify each gamble’s most likely outcome. Then choose the gamble with the highest, most likely outcome <sup>a</sup>	Always

<sup>a</sup> If the decisive reason did not discriminate, the heuristic chose randomly. Heuristics are from Thorngate (1980) and Brandstätter et al. (2006; see also Coombs et al., 1970).

**Table 2**  
Summary of the predictions for CPT’s constructs when accommodating heuristic choices.

Domain		Heuristic				
		Minimax	Maximax	Priority heuristic <sup>a</sup>	Least-likely	Most-likely
Decisive reason	Gain	Minimum outcome	Maximum outcome	Minimum outcome	Probability of worst outcome	Most likely outcome
	Loss	Minimum outcome	Maximum outcome	Maximum outcome	Probability of worst outcome	Most likely outcome
	Mixed	Minimum outcome	Maximum outcome	Minimum outcome	Probability of worst outcome	Most likely outcome
Prediction	Probability sensitivity	Low sensitivity (low $\gamma$ )	Low sensitivity (low $\gamma$ )	Medium sensitivity (medium $\gamma$ )	High sensitivity (high $\gamma$ )	High sensitivity (high $\gamma$ )
	Risk aversion	Risk averse in gains (low $\delta^-$ )	Risk seeking in gains (high $\delta^+$ )	Risk averse in gains (low $\delta^+$ )	Risk averse in gains (low $\delta^-$ )	–
		Risk averse in losses (high $\delta^-$ )	Risk seeking in losses (low $\delta^-$ )	Risk seeking in losses (low $\delta^-$ )	Risk averse in losses (high $\delta^-$ )	–
	Outcome sensitivity <sup>b</sup>	High sensitivity (high $\alpha$ )	High sensitivity (high $\alpha$ )	Medium/high sensitivity (medium/high $\alpha$ )	Low sensitivity (low $\alpha$ )	High sensitivity, (high $\alpha$ )
Loss aversion <sup>c</sup>	Loss averse (high $\lambda$ )	Gain seeking (low $\lambda$ )	Loss averse (high $\lambda$ )	Loss averse (high $\lambda$ )	–	

<sup>a</sup> We list the reason that receives highest priority because it is examined first.

<sup>b</sup> In the parameter estimation, a common  $\alpha$  and  $\gamma$  for the gain and loss domains was estimated.

<sup>c</sup> An asymmetry between losses and gains—as embodied in the notion of loss aversion—can also be represented by the elevation of the weighting function (i.e.,  $\delta^- > \delta^+$ ); see text.

Table 1). Therefore, to the extent that concern for outcomes is reflected in CPT’s value function, minimax, maximax, and the most-likely heuristic should result in the highest estimates of the  $\alpha$  parameter (indicating high outcome sensitivity), followed by the priority heuristic; the least-likely heuristic should result in the lowest  $\alpha$  parameter estimates (indicating low outcome sensitivity).

### 1.4.3. Risk aversion

As highlighted by Lopes (1995), one factor influencing a decision maker’s risk aversion is the amount of attention given to the worst outcomes (see also Abdellaoui, L’Haridon, & Zank, 2010; Birnbaum, 2008; Brandstätter et al., 2006; Wakker, 2001). Accordingly, the minimax heuristic can be expected to be risk averse, because it assumes an exclusive concern with the options’ worst outcomes (in gain and loss domains). Conversely, the maximax heuristic should be risk seeking, because it exclusively focuses on the options’ best outcomes (Brandstätter et al., 2006; Lopes, 1990). The priority heuristic’s risk aversion depends on the domain: It produces risk-averse choices in the gain domain and risk-seeking choices in the loss domain (a pattern called the reflection effect; Kahneman & Tversky, 1979). Because the least-likely heuristic chooses the option with the lowest probability of the worst outcome, it should be risk averse in both domains.

Predicting the most-likely heuristic’s degree of risk aversion is less straightforward. As the heuristic chooses the option with the most attractive most likely outcome, in multi-outcome options it might consider the worst, the best, or one of the intermediate outcomes (depending on their probability); it is therefore unclear whether the heuristic’s behavior will be risk averse, risk seeking, or risk neutral.

Although risk attitude is often explained based on the curvature of the utility/value function, since Yaari (1987) it has increasingly been recognized that different degrees of risk aversion can also be modeled by the elevation of a probability

weighting function—especially when they arise from differential attention to extreme outcomes (e.g., Lopes, 1995, p. 206). From this perspective, minimax and the least-likely heuristic might lead to low (high) values on the  $\delta$  parameter for gains (losses), maximax to high (low) values of  $\delta$  for gains (losses), and the priority heuristic to low values of  $\delta$  (for gains and losses equally).

#### 1.4.4. Loss aversion

The degree of loss aversion in CPT is modeled by the  $\lambda$  parameter in the value function (Eq. (2)), with  $\lambda > 1$ . Because the outputs of the value and weighting functions are combined multiplicatively, however, decisions reflecting a larger weight for losses than gains can, in principle, also be modeled by a more elevated weighting function for the loss than for the gain domain (i.e.,  $\delta^- > \delta^+$ ; Pachur & Kellen, 2013; Schmidt & Zank, 2005).<sup>4</sup> Which factors might influence the extent to which a heuristic leads to loss-averse choices?

One obvious candidate is the degree to which a heuristic prioritizes information about possible losses over information about possible gains. Consider, for instance, the minimax heuristic. This heuristic predicts the “choice of such an act that the greatest loss that can possibly accrue to it shall be as small as possible” (Savage, 1954/1972, p. 164). Other things being equal, minimax’s concern with losses should result in stronger loss aversion than maximax, whose concern lies with gains. That is, minimax should lead either to relatively high values of the  $\lambda$  parameter or to a relatively large discrepancy between  $\delta^-$  and  $\delta^+$  (with  $\delta^- > \delta^+$ ). Similarly, the other heuristics concerned with losses in mixed gambles—the least-likely and priority heuristics—might also lead to pronounced loss aversion. In contrast, the maximax heuristic, whose aspiration is to chase gains, might not produce loss-averse choices but instead more loss-neutral choices. Again, predicting the most-likely heuristic’s degree of loss aversion is less straightforward.

## 2. How and how well does CPT reflect heuristic processing principles?

We tested these predictions using computer simulations. This test also contributes to a validation of the—to our knowledge yet untested—assumptions that the weighting and value functions reflect probability and outcome sensitivity, respectively. For instance, it is currently unclear whether the functions’ curvatures are indicative of differences between choice mechanisms in the degree to which they attend to probability and outcome information. The heuristics included in our analyses tend to focus on the extreme outcomes (and their probabilities) offered by an option and to ignore less extreme outcomes (and their probabilities). Therefore, the larger the number of outcomes an option has, the more information will be ignored. To test the extent to which CPT’s ability to represent choices generated by heuristics depends on the number of outcomes, we conducted the analyses separately for problem sets with two-, three-, and five-outcome gambles. Each set consisted of a total of 180 two-option choice problems, with 60 each involving gains, losses, and gains *and* losses (i.e., mixed gambles). All problems included pairs of options with similar expected values (EVs; specifically, the ratio of the absolute differences of the options’ EVs to the absolute value of the option with the smaller expected value was smaller than 1; for an analysis with pairs of options differing widely in expected value, see Appendix B). To avoid introducing specific regularities (to which the different heuristics might respond differently), we constructed the problems randomly, in line with the approach taken in many empirical studies (e.g., Erev, Roth, Slonim, & Barron, 2002; Glöckner & Pachur, 2012; Jarvstad, Hahn, Rushton, & Warren, 2013; Rieskamp, 2008; Scheibehenne, Rieskamp, & González-Vallejo, 2009); further, probability levels covered the entire range of 0 to 1, and the outcome levels were selected randomly from the range of 0 to 100 and –100 to 0 in the gain and loss domain, respectively. The construction procedure is described in more detail in Appendix C. For each problem, we determined the option that a given heuristic would choose and then estimated the parameters of CPT’s weighting and value functions for those simulated choices using a Bayesian approach (e.g., Lee & Wagenmakers, 2013; see Appendix D for more details).

### 2.1. Model fit

In our implementation of CPT, we set  $\alpha^+ = \alpha^-$  because, as shown by Nilsson et al. (2011), estimating separate exponents of the value function for gains and losses can lead to misestimates of  $\lambda$  (see also Wakker, 2010). We also set the probability sensitivity parameter  $\gamma$  to be equal across gains and losses, because the heuristics under investigation treat probabilities equally across the gain and loss domains. The elevation parameter  $\delta$ , by contrast, was estimated separately for gains and losses. The reason is that, as described above, some heuristics differ starkly with regard to risk aversion, and separate elevation parameters for gains and losses afford an alternative way to model loss aversion (Pachur & Kellen, 2013). Overall, as implemented here, CPT encompasses six adjustable parameters:  $\alpha$  and  $\lambda$  for the value function;  $\gamma$ ,  $\delta^+$ , and  $\delta^-$  for the weighting function; as well as a  $\phi$  parameter for the choice rule needed to derive predicted choice probabilities (see Appendix D). To reflect CPT’s main assumptions, we restricted the parameter values as follows in the parameter estimation (see Rieskamp, 2008):  $0 \leq \alpha \leq 1$ ;  $1 \leq \lambda \leq 5$ ;  $0 \leq \gamma \leq 1$ ;  $0 \leq \delta^\pm \leq 5$ ;  $0 \leq \phi \leq 5$  (for more information on the prior distributions used in the parameter estimation, see Appendix D).

<sup>4</sup> As described in more detail below, in the analyses we set  $\gamma^+ = \gamma^-$  and  $\alpha^+ = \alpha^-$ .

To what extent does CPT, which embodies a compensatory, multiplicative calculus, capture choices generated by heuristics that embody a noncompensatory calculus? To answer this question, we first determined posterior predictives, a common way of assessing model fit in Bayesian statistics. Posterior predictives are synthetic, model-generated data sets that are produced by parameters values from the posterior distributions. Specifically, we determined, based on the posterior distributions of the CPT parameters, the proportion of times the choice predicted by CPT coincided with the choice predicted by the respective heuristic. As a second index of how well CPT captures the choices in the respective choice set, we determined the deviance information criterion (DIC; Spiegelhalter, Best, Carlin, & van der Linde, 2002), with a lower DIC indicating a better correspondence between the model and the data.

As shown in Table E1 (Appendix E), CPT captured choices generated by the heuristics rather well, in most cases mimicking them more than 70% of the time (range 57–95%). Similarly, the DICs for each heuristic's choices were considerably better than expected under chance (i.e., DIC = 249.53; DIC differences larger than 10 are considered as strongly favoring the better model; see Spiegelhalter et al., 2002, p. 613). However, CPT did not capture all heuristics equally well. Interestingly, the choices of the heuristics that ignore probabilities (i.e., minimax and maximax) were captured better than those of the priority heuristic (which sometimes considers probabilities), as reflected by a higher percentage of overlapping choices and lower DIC values. The worst fit was obtained for those heuristics that consider both outcomes and probabilities (least-likely and most-likely). It was not the case that CPT had a worse model fit when the gambles had a higher number of outcomes. Overall, CPT was able to capture the heuristics' choices rather well, even though some of the estimates showed high mass at the boundaries of the parameter range (see Appendix F). We return to the differences between the heuristics and CPT parameter estimates with an extended parameter range in the General Discussion.

## 2.2. Heuristics' alter egos in CPT

How do CPT's parameters characterize the different heuristics in terms of probability and outcome sensitivity as well as risk and loss aversion? Fig. 1 depicts the weighting and value functions for all five heuristics in the set of problems with two-outcome gambles graphically. The functions are based on the means of the posterior distributions of the parameter estimates (the mean values for all CPT parameters and for all three sets of choice problems can be found in Table E1 in Appendix E).

### 2.2.1. Probability sensitivity

We hypothesized that differences between the heuristics in terms of their consideration of probabilities would be reflected in the  $\gamma$  parameter of CPT's weighting function. As Fig. 1 shows, CPT indeed yielded weighting functions with very distinct shapes. As predicted, minimax and maximax resulted in low values of the  $\gamma$  parameter—see Table E1 of Appendix E—that is, in very strongly curved weighting functions; the priority heuristic resulted in a somewhat higher value of  $\gamma$  and therefore a less strongly curved weighting function than minimax or maximax; the least-likely and most-likely heuristics resulted in values of  $\gamma$  close to 1 and therefore the least strongly curved weighting functions. CPT's  $\gamma$  parameter thus veridically reflected differences in the heuristics' concern for probabilities. This pattern of results held across the problems with two-, three-, and five-outcome gambles (see Table E1 in Appendix E).

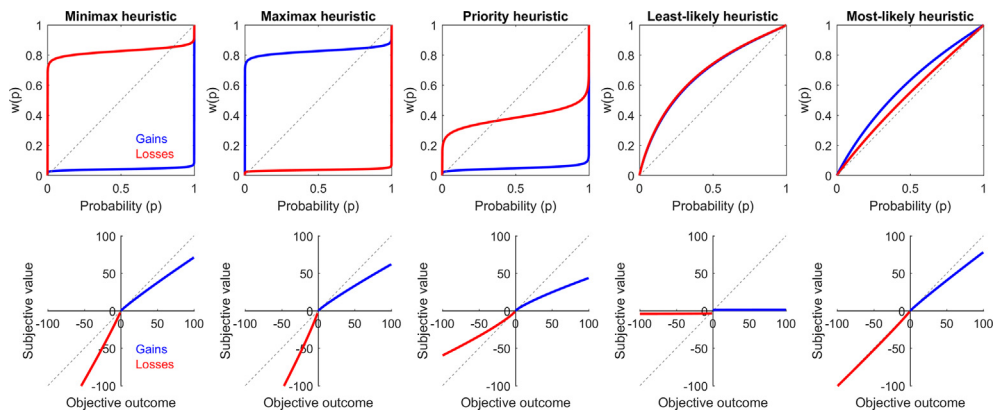
### 2.2.2. Outcome sensitivity

CPT also veridically reflected differences between the heuristics in terms of their sensitivity to outcome information, as indicated by the shapes of the value function estimated for the different heuristics' choices. Recall that values of  $\alpha < 1$  indicate a more concave (convex) value function in the gain (loss) domain and thus a lower outcome sensitivity. As Fig. 1 (lower row) shows, for the heuristics that make a choice merely on the basis of outcome information or that consider this information first—minimax, maximax, the priority heuristic, and the most-likely heuristic—the value function was approximately linear (i.e.,  $\alpha$  was rather high), indicating high sensitivity to outcomes. The least-likely heuristic, by contrast, which chooses on the basis of probability information, yielded a very low value of  $\alpha$  and a very strongly curved value function. This pattern held consistently across the problems with two-, three-, and five-outcome gambles (see Table E1 in Appendix E).

### 2.2.3. Risk aversion

Due to their differing concerns for extreme outcomes, the heuristics' policies may result in differences in the elevation of CPT's weighting function, indicating differences in risk aversion. In this context, it is important to distinguish between the gain and loss domain. In the gain domain, high risk aversion is indicated by a *low* elevation; in the loss domain, high risk aversion is indicated by a *high* elevation. A heuristic such as minimax that invariably focuses on the worst outcomes in both gains and losses may therefore yield a weighting function with a low elevation for gains (i.e., a low value of  $\delta^+$ ) and a high elevation for losses (i.e., a high value of  $\delta^-$ ). The opposite pattern is predicted for a heuristic such as maximax that bets on the best outcomes and is thus risk seeking. The estimated parameters confirmed this prediction, and this held across all sets of choice problems (see Table E1 in Appendix E). Minimax resulted in a weakly (strongly) elevated weighting function for gains (losses), whereas maximax resulted in a strongly (weakly) elevated weighting function for gains (losses). For the choices generated by the priority heuristic, CPT yielded low values of the  $\delta^+$  and  $\delta^-$  parameters, indicating risk aversion in the gain domain and risk seeking in the loss domain (i.e., the reflection effect). The most-likely heuristic yielded estimates for  $\delta^+$





**Fig. 1.** CPT's weighting function (upper row) and value functions (lower row) estimated for the choices of all five heuristics for problems with two-outcome gambles (see Table E1 in Appendix E for the parameter values).

and  $\delta^-$  that had intermediate values (for both gains and losses), suggesting risk neutrality. Finally, the least-likely heuristic resulted in a weakly (strongly) elevated weighting function for gains (losses), suggesting a high level of risk aversion.<sup>5</sup>

#### 2.2.4. Loss aversion

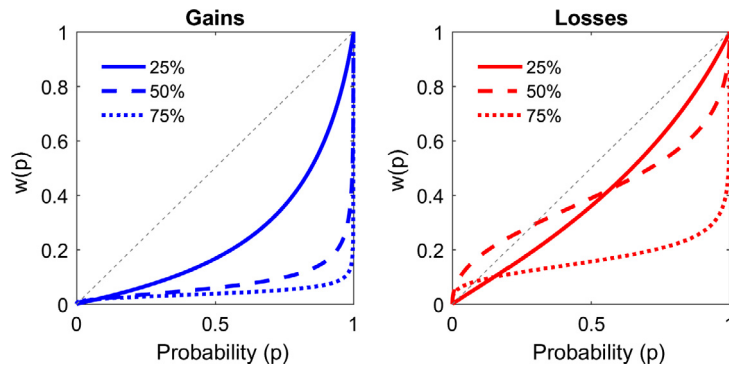
Our tentative hypothesis was that the focus of minimax, the least-likely heuristic, and the priority heuristic in mixed gambles on the worst outcomes (losses) would lead to greater loss aversion than shown by heuristics that focus on the best outcomes (maximax). Concerning the estimated values of the  $\lambda$  parameter, the results showed a mixed picture for all five heuristics (the parameter estimates are reported in Table E1 in Appendix E): For some heuristics,  $\lambda$  increased with the number of outcomes in the gamble; for others, it decreased. Minimax, for instance, showed a strong degree of loss aversion in problems with two-outcome gambles but less pronounced loss aversion in problems with three- or five-outcome gambles. In addition, recall that asymmetries in the weighting of gain and loss outcomes can also be represented in CPT's weighting function, namely by assuming a higher elevation for loss probabilities than for gain probabilities (i.e.,  $\delta^- > \delta^+$ ; Pachur & Kellen, 2013; Schmidt & Zank, 2005). Minimax indeed showed a gain–loss asymmetry in the elevation, as did the priority heuristic and, to some extent, the least-likely heuristic, both of which focus on losses in mixed gambles. For maximax, which focuses on gains, the pattern was reversed (i.e.,  $\delta^+ > \delta^-$ ), suggesting a higher weighting of gains than losses. For the most-likely heuristic, the pattern was less consistent. For the problems with two-outcome and five-outcome gambles, the asymmetry between  $\delta^+$  and  $\delta^-$  indicated a small degree of higher weighting of gains over losses; for problems with three-outcome gambles, in contrast, the asymmetry indicated loss aversion.

### 3. How does CPT reflect lexicographic search?

Whereas heuristics such as minimax and maximax focus on a single reason, lexicographic strategies such as the priority heuristic (Brandstätter et al., 2006; see also Payne et al., 1993) assume that multiple reasons are examined sequentially, and that search is stopped once a reason is found that discriminates between the options. Which reason determines the choice and how many reasons are inspected by lexicographic strategies depends on the characteristics of the choice problem, namely which reason is the first to discriminate between the options. Because of this dependency, the outcome or probability sensitivity of choices following from a lexicographic heuristic can be expected to vary, depending on the properties of the choice problems.

Using the priority heuristic as an example of a lexicographic strategy, we next demonstrate that CPT can track this dependency. An analysis of the priority heuristic is also conceptually interesting, because it integrates several other heuristics (Table 1) into one. For gains, the heuristic first considers the same information on which minimax focuses exclusively (minimum outcomes). However, it makes a choice only if the difference between the minimum outcomes exceeds the aspiration level (Table 1). If not, it turns to the probability of the minimum outcomes (again using an aspiration level). This, in turn, coincides with the decision rule of the least-likely heuristic. If the probabilities do not differ, the policy of the maximax

<sup>5</sup> Overall, these results were robust across the problems with two-outcome, three-outcome, and five-outcome gambles—with one exception. Specifically, whereas in problems with three- and five-outcome gambles the elevation of the least-likely heuristic's weighting function indicated risk aversion for gains—consistent with the heuristic's focus on the worst outcome—in problems with two-outcome gambles the elevation indicated risk seeking. Why? The least-likely heuristic chooses the option whose worst outcome occurs with lower probability. In problems with two-outcome gambles, this option is automatically also the one with the higher probability of the best outcome—and thus the one that a risk-seeking strategy would choose. This regularity may explain why the heuristic resulted in a highly elevated weighting function for gains in problems with two-outcome gambles.



**Fig. 2.** CPT's weighting functions estimated for the choices of the priority heuristic in problems with two-outcome gambles, separately for problems where the decision was based on the top-ranked reason (probability of minimum gain/loss) in 25%, 50%, or 75% of cases, respectively (the choice was otherwise based on the second reason: the probability of the minimum gain/loss). See Table E2 in Appendix E for the parameter values.

heuristic will ultimately be invoked (i.e., focus on maximum outcomes). An analysis of the priority heuristic with CPT across different sets of choice problems can thus also be viewed as an analysis of how CPT captures the contingent use of various choice strategies, dependent on the structure of the problems.

We estimated CPT's parameters for choices generated by the priority heuristic in three sets of problems, in which the minimum gains (losses) discriminated (and determined the choice) in 75%, 50%, or 25% of cases and probabilities were, in turn, consulted (and determined the choice) in 25%, 50%, or 75% of cases. We predicted that the CPT parameter  $\gamma$  would increase across these sets, indicating increasing probability sensitivity, because of the increasing consideration of probabilities. We created three sets of two-outcome problems, each including 180 randomly generated choice problems—60 gain, 60 loss, and 60 mixed problems—with similar EVs. The problems were constructed according to the same principles as in our first analysis (see Appendix C for details).

The weighting functions based on the mean of the posterior distributions of the  $\gamma$  parameter showed the predicted pattern across the three problem sets (Table E2 in Appendix E): When choices were based predominantly on outcome information (75%),  $\gamma$  was relatively low, indicating low sensitivity to probabilities. When choices were based on probability information in 50% of cases, the  $\gamma$  parameter was considerably higher. Finally, when choices relied predominantly on probability information (25%),  $\gamma$  had the highest value. Fig. 2 plots the corresponding weighting functions for the three problem sets, separately for gains and losses. As can be seen, the curvature of the weighting function is more pronounced for those sets in which probabilities determine the choice less frequently. The model fit of CPT varied with the percentage of cases in which the first reason discriminated (see Table E2 in Appendix E), an issue to which we return in the General Discussion. In the Supplemental Material, we also report an analysis for a modified version of the priority heuristic, in which the order in which the probability and the magnitude of the minimum outcome are inspected is reversed (Rieskamp, 2008).

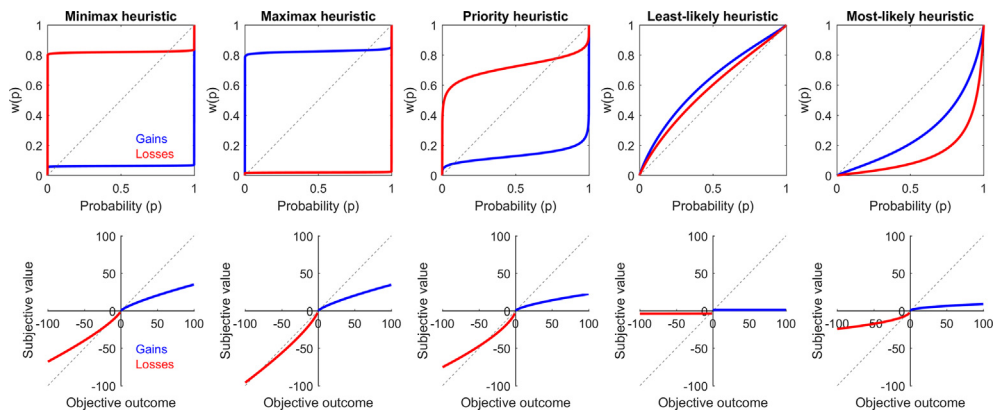
To recap, the curvature of CPT's weighting function veridically reflected differences in how a lexicographic strategy—the priority heuristic—attended to probabilities across sets of choice problems that differed in the frequency with which choices were based on probability information. That is, CPT's parameters can be harnessed to characterize and differentiate lexicographic search orders and search truncation.

#### 4. Using CPT to characterize environment-dependent behavior of heuristics

Our analyses thus far have shown that CPT can be used to characterize the choices of heuristics with respect to probability sensitivity, outcome sensitivity, and risk aversion. The extent to which a heuristic shows risk aversion may depend on the structure of the choice environment. The choice problems we have used thus far were constructed such that their probabilities and outcomes were independent (as is often the case in empirical investigations of decision making under risk; see Pleskac & Hertwig, 2014, Fig. 9). Outside the laboratory, however, there is often a strong negative relationship between payoffs and probabilities (Pleskac & Hertwig, 2014, Figs. 3 and 4).

When large rewards and losses tend to be less likely than smaller rewards and losses, how risk averse are the choices produced by heuristics? In such an environment, the most-likely heuristic, which focuses on the most likely outcomes, will choose the option with the smaller gains and losses; it should therefore result in risk-averse choices for gains and risk-seeking choices for losses. Note that this behavior contrasts with the most-likely heuristic's risk-neutral choices when probabilities and outcomes are uncorrelated (Fig. 1). For the other heuristics, a negative correlation between probabilities and outcomes should not influence their risk attitude, as they either base their choices solely on outcome information (minimax and maximax) or consider outcomes first and the associated probabilities later (priority heuristic and least-likely heuristic).

To test this prediction, we repeated the procedure of our computer simulation (with choice problems consisting of gambles with similar EVs) in every respect, except that outcomes (i.e., their absolute magnitudes) and probabilities were now



**Fig. 3.** CPT's weighting function (upper row) and value functions (lower row) estimated for the choices of all five heuristics for problems with two-outcome gambles in which outcomes and probabilities were negatively correlated (see Table E3 in Appendix E for the parameter values).

negatively correlated, thus creating high conflict between outcomes and probabilities.<sup>6</sup> As before, we determined each heuristic's choices in each of the three new problem sets and estimated CPT's parameters for these choices (using the same Bayesian estimation procedure as above). The model fit for CPT, as indicated by posterior predictive checks and DICs, was rather good for most heuristics (Table E3 of Appendix E). For the choices of the most-likely heuristic, CPT's fit was much better than in the problems with uncorrelated outcomes and probabilities. For the least-likely heuristic, however, CPT's fit was again poor (we will discuss this point later).

Overall, as shown in Fig. 3, the different heuristics produced similar weighting and value functions in CPT as in the uncorrelated environment. As predicted, however, the one exception was the most-likely heuristic, which yielded a very different pattern on the elevation parameter  $\delta$  in the correlated environment than it did in the uncorrelated environment: The heuristic's ignorance of rare but large gains and rare but large losses now led to risk aversion for gains and to risk seeking for losses (as indicated by a pattern of pronounced underweighting in the resulting weighting function)—in contrast to the risk-neutral pattern obtained in the uncorrelated environment. To illustrate, in the problems with two-outcome gambles, the parameter values were  $\delta^+ = 0.27$  and  $\delta^- = 0.08$ , indicating risk aversion for gains and risk seeking for losses, as opposed to  $\delta^+ = 1.72$  and  $\delta^- = 1.23$  in the uncorrelated environment. This pattern held across all three sets of choice problems (see Table E3 of Appendix E).

Moreover, there seemed to be a trend for the probability sensitivity parameter  $\gamma$  to be lower in the correlated than in the uncorrelated environment (see Table E3 of Appendix E). This finding can be explained as follows: In the negatively correlated environment, the most probable outcome is always the smallest gain or the smallest loss. As the most-likely heuristic focuses exclusively on these outcomes, its choices in such an environment are comparable to those of minimax for gains and maximax for losses, which have a low probability sensitivity. The other parameters were mostly unchanged for this heuristic. For minimax, maximax, the priority heuristic, and the least-likely heuristic, the resulting parameter profiles were highly similar to those obtained in the uncorrelated environment. These results illustrate how CPT's parametric framework can help reveal a choice policy's sensitivity (or lack thereof) to changes in the structure of the environment (e.g., Simon, 1956; Todd, Gigerenzer, & the ABC Research Group, 2012).

## 5. Model recovery: how well can CPT and the heuristics be distinguished?

Several empirical investigations have compared CPT and heuristics directly as models predicting observed choices (Glöckner & Pachur, 2012; Pachur, Hertwig, Gigerenzer, & Brandstätter, 2013; Rieskamp, 2008; Su et al., 2013). Our finding that CPT is able to accommodate heuristic choices highlights a potential problem with these model comparisons: If CPT can mimic choices generated by heuristics, such choices could erroneously be attributed to a multiplicative mechanism (as embodied by CPT), and vice versa. To examine this possibility, we conducted a model recovery study (e.g., Pitt & Myung, 2002). Specifically, drawing on the set of 180 two-outcome choice problems used in our first computer simulation (and which has also been used in previous empirical studies; e.g., Glöckner & Pachur, 2012; Rieskamp, 2008), we simulated the choices of heuristics as well as the choices of CPT (based on a published set of parameter values); using strategy

<sup>6</sup> The average correlation between the probabilities and the outcomes across all problems was strongly negative, with  $r = -0.70$ ,  $t(178) = 12.96$ ,  $p < 0.001$ ,  $r = -0.81$ ,  $t(178) = 18.11$ ,  $p < 0.001$ , and  $r = -0.85$ ,  $t(178) = 21.60$ ,  $p < 0.001$ , for the problems with two-outcome, three-outcome, and five-outcome gambles, respectively (one-sample *t*-test testing the mean Fisher's *z*-transformed correlation against zero). The correlation between the outcomes and the corresponding probabilities was determined separately for each problem. To illustrate, for a problem with two-outcome gambles, the correlation of the four outcomes of gambles A and B and their respective probabilities was computed. We then averaged the *z*-transformed correlation coefficients across all 180 problems. In the earlier problem sets, outcomes and probabilities had an average correlation of  $r = 0.09$ ,  $t(178) = 1.23$ ,  $p = 0.22$ ,  $r = -0.1$ ,  $t(178) = 1.38$ ,  $p = 0.17$ , and  $r = -0.007$ ,  $t(178) = 0.09$ ,  $p = 0.93$  for the problem with two-outcome, three-outcome, and five-outcome gambles, respectively.

classification, we then inferred which of these models generated a given set of choices. To examine how unsystematic noise may increase the risk of confusing CPT and heuristics, we varied the amount of noise in the data.

### 5.1. Data generation

We employed the same 180 two-outcome problems (60 gain, 60 loss, and 60 mixed problems) as in the simulation with uncorrelated outcomes and probabilities. In a first step, choices were generated for six models: the minimax heuristic, the maximax heuristic, the priority heuristic, the least-likely heuristic, the most-likely heuristic (Table 1), and CPT using parameter values reported by Glöckner and Pachur (2012;  $\alpha = 0.68$ ;  $\lambda = 1.52$ ;  $\gamma = 0.73$ ;  $\delta = 0.99$ ). In order to examine recovery performance as a function of different degrees of noise, we varied the percentage of (randomly selected) problems in which the prediction of a given model was reversed. That is, instead of selecting the option predicted by a model, the opposite one was chosen (for other studies using this approach, see, e.g., Cavagnaro, Gonzalez, Myung, & Pitt, 2013; Yechiam & Busemeyer, 2005). Preliminary analyses showed that the largest changes in recovery performance occurred with error rates between 35% (63 of the 180 choices reversed) and 43.8% (79 of the 180 choices reversed). We therefore report the recovery rates mainly within this interval. As an additional benchmark, we also generated choice sets in which all responses were generated at random.<sup>7</sup>

### 5.2. Model recovery

Using a Bayesian latent mixture approach (e.g., Lee, 2016; Zeigenfuse & Lee, 2010), we classified each choice vector (consisting of 180 synthetic choices) to one of the six generating mechanisms or to a “guessing” mechanism. The key assumption of this approach is that the observed behavioral data arise from different processes (e.g., cognitive mechanisms). A categorical latent mixture variable  $z$  indicates which mechanism generated the data. When modeling a choice, the probability that option A is chosen over option B in a particular choice problem corresponds to the predicted choice probability of the mechanism to which the  $z$  variable points. To derive this probability for minimax, maximax, the priority heuristic, the least-likely heuristic, and the most-likely heuristic, we estimated an application error parameter  $\phi$  for each mechanism that was assumed to be constant across choice problems. For CPT, we estimated  $\alpha$ ,  $\lambda$ ,  $\gamma$ ,  $\delta^+$ , and  $\delta^-$  and an application error parameter  $\phi$ . For comparability with the heuristics, this parameter was assumed to be constant across choice problems. For the “guessing” mechanism,  $\phi$  was fixed at 0.5. The key result of a latent mixture analysis consists in the posterior distributions of the latent mixture variable  $z$  for each choice vector. For each tested mechanism, the distribution gives the posterior probability that it generated the data. One key advantage of the posterior distribution is that it automatically balances goodness-of-fit with all forms of model complexity. Each choice vector was classified to one of the six generating models or to the “guessing” mechanism, depending on to which the posterior distribution of the latent mixture variable  $z$  indicated the largest probability mass. Ideally, the recovered model was the one that actually generated the choices. For each error rate, the procedure was repeated 1000 times for each generating model.

### 5.3. Results

Table 3 shows, for different error rates and each generating mechanism, the proportion of choice vectors (averaged across the 1000 repetitions) that were classified as having been generated by one of the six mechanisms or the “guessing” mechanism, with the proportion of correct recoveries indicated in bold. Cells off the diagonal represent incorrect recoveries, that is, cases in which a choice vector was classified as having been generated by one mechanism when in fact another had generated it. With an error rate of 0%, recovery performance was perfect, and it was perfect for error rates up to 33.3% (results are therefore only shown in Table 3 from an error rate of 35%). Even with an error rate of 42.2%, the generating mechanism was still correctly recovered in the majority of cases. This held for all mechanisms except for CPT, for which from an error rate of 40% the majority of cases were classified as guessing. This suggests that for a data set with high noise, the complexity of CPT is not matched with a corresponding increase in ability to fit the data. There was a slight trend for choices produced by CPT to be incorrectly recovered more frequently by the priority heuristic than by the other heuristics. This may be due to the fact that the priority heuristic implies several of the choice regularities that shaped the development of CPT (e.g., reflection effect, Allais paradox, fourfold pattern; Katsikopoulos & Gigerenzer, 2008). With noise levels higher than 43.3% for all mechanisms the majority of cases were classified as guessing. Overall, these results suggest that although CPT is able to accommodate choices generated by heuristics, CPT and heuristics can be distinguished with high accuracy, unless when choice generation is very noisy.<sup>8</sup>

<sup>7</sup> Importantly, note that the model recovery results for a complete set of randomly generated choices might differ from those for a set with a 50% error rate. If the predictions of a given strategy are reversed in 50% of cases, it will never recover its own choices better than chance. All other strategies, however, could coincidentally predict these choices better than chance. Therefore, the strategy that never recovers choices better than chance will be penalized relative to the remaining strategies. When all choices are random, however, all strategies share the same handicap.

<sup>8</sup> We obtained similar results when the model parameters were estimated with a maximum-likelihood approach and the simulated choice vectors were classified based on the Bayesian Information Criterion (BIC; Schwarz, 1979). The main difference was that, with the maximum-likelihood approach, under high levels of noise (43% or higher) CPT was incorrectly classified as the generating mechanism in a disproportionate number of cases. A potential explanation for this discrepancy is that, to control for model complexity, BIC takes into account only a model's number of free parameters, whereas the Bayesian latent mixture approach is also sensitive to a model's functional form. Therefore, the identifiability of choices produced by CPT versus heuristics also depends on the method of model selection.

**Table 3**

Proportion of classifications as a function of generating mechanism and error rate in the model recovery study. Correct recoveries are in bold.

Error rate	Generating mechanism	Recovered mechanism						
		MI	MA	PH	LL	ML	CPT	Guessing
0%	MI	<b>1</b>	0	0	0	0	0	0
	MA	0	<b>1</b>	0	0	0	0	0
	PH	0	0	<b>1</b>	0	0	0	0
	LL	0	0	0	<b>1</b>	0	0	0
	ML	0	0	0	0	<b>1</b>	0	0
	CPT	0	0	0	0	0	<b>1</b>	0
35%	MI	<b>0.991</b>	0	0.005	0	0	0.004	0
	MA	0	<b>0.999</b>	0	0	0	0.001	0
	PH	0.004	0	<b>0.996</b>	0	0	0	0
	LL	0	0	0	<b>0.999</b>	0	0.001	0
	ML	0	0	0	0	<b>0.991</b>	0.009	0
	CPT	0.007	0.008	0.116	0.021	0.006	<b>0.792</b>	0.050
38.3%	MI	<b>0.962</b>	0.001	0.017	0	0.002	0.018	0
	MA	0	<b>0.980</b>	0.001	0	0.003	0.016	0
	PH	0.021	0.002	<b>0.972</b>	0	0.001	0.004	0
	LL	0	0	0	<b>0.995</b>	0.002	0.003	0
	ML	0.001	0.002	0.001	0	<b>0.959</b>	0.037	0
	CPT	0.035	0.025	0.108	0.040	0.043	<b>0.406</b>	0.343
40%	MI	<b>0.919</b>	0.001	0.040	0	0.006	0.034	0
	MA	0	<b>0.955</b>	0.009	0	0.010	0.026	0
	PH	0.032	0.015	<b>0.937</b>	0	0.005	0.011	0
	LL	0	0	0.001	<b>0.990</b>	0.001	0.008	0
	ML	0.002	0.006	0.003	0.007	<b>0.914</b>	0.068	0
	CPT	0.023	0.025	0.124	0.054	0.034	<b>0.282</b>	0.458
42.2%	MI	<b>0.795</b>	0.002	0.096	0	0.025	0.079	0.003
	MA	0.005	<b>0.863</b>	0.028	0.001	0.034	0.069	0
	PH	0.079	0.040	<b>0.827</b>	0.005	0.023	0.026	0
	LL	0	0.002	0.004	<b>0.949</b>	0.016	0.029	0
	ML	0.015	0.028	0.012	0.016	<b>0.805</b>	0.124	0
	CPT	0.041	0.038	0.100	0.055	0.043	<b>0.156</b>	0.567
43.3%	MI	<b>0.305</b>	0.007	0.092	0.001	0.027	0.082	0.486
	MA	0.011	<b>0.153</b>	0.056	0	0.034	0.085	0.661
	PH	0.116	0.056	<b>0</b>	0.006	0.030	0.047	0.745
	LL	0	0.007	0.005	<b>0</b>	0.028	0.056	0.904
	ML	0.028	0.043	0.033	0.034	<b>0.144</b>	0.122	0.596
	CPT	0.030	0.044	0.085	0.058	0.035	<b>0.091</b>	0.657
43.8%	MI	<b>0.065</b>	0.013	0.094	0.001	0.020	0.071	0.736
	MA	0.008	<b>0</b>	0.051	0.005	0.051	0.078	0.807
	PH	0.109	0.040	<b>0</b>	0.013	0.012	0.039	0.787
	LL	0	0.003	0.008	<b>0</b>	0.029	0.036	0.924
	ML	0.022	0.049	0.022	0.040	<b>0</b>	0.115	0.752
	CPT	0.032	0.042	0.054	0.048	0.040	<b>0.103</b>	0.681
50%	MI	<b>0</b>	0.017	0.021	0.010	0.027	0.026	0.899
	MA	0.038	<b>0</b>	0.030	0.023	0.022	0.011	0.876
	PH	0.029	0.022	<b>0</b>	0.027	0.024	0.034	0.864
	LL	0.015	0.028	0.018	<b>0</b>	0.021	0.028	0.890
	ML	0.035	0.026	0.036	0.022	<b>0</b>	0.018	0.863
	CPT	0.034	0.022	0.021	0.020	0.030	<b>0.020</b>	0.853
Random choice		0.033	0.028	0.026	0.030	0.027	0.027	<b>0.830</b>

Note. MI = minimax heuristic, MA = maximax heuristic, PH = priority heuristic, LL = least-likely heuristic, ML = most-likely heuristic.

## 6. Using CPT to track heuristic choice in an empirical data set

Our computer simulation analyses demonstrated that each heuristic gives rise to a specific parameter profile that reflects the heuristic's processing properties. Next, we illustrate with an empirical study how CPT profiles can be used to infer the use of a heuristic. In this study, we had participants make (hypothetical) risky choices both between pairs of monetary gambles and between pairs of medications. In the former, participants risked losing an amount of money with some probability; in the latter, they risked experiencing a side effect with some probability. Importantly, the choice problems in both domains were matched monetarily (see below for details). Several studies have found systematically different patterns of choice in monetary versus medical problems (e.g., Lejarraga, Pachur, Frey, & Hertwig, 2016; Pachur & Galesic, 2013). These differences have been attributed to side effects triggering stronger affect than monetary losses (Pachur et al., 2014; Suter, Pachur,



Hertwig, Endestad, & Biele, 2015; Suter et al., 2016; see also Rottenstreich & Hsee, 2001). Might people rely on different strategies in the different domains?

Our analyses proceeded as follows: We first analyzed the monetary and medical choices separately using CPT. For the medical choices, we obtained a CPT parameter profile that resembled that of the minimax heuristic (as identified in our computer simulations; see Fig. 1). Based on this result, we conducted a strategy classification, in which individual participants were classified as using either a compensatory choice strategy (EV strategy) or the noncompensatory minimax heuristic. We then examined whether this classification was consistent with the CPT analysis—in other words, whether in the medical choices people were more likely to rely on the minimax heuristic.

## 6.1. Method

### 6.1.1. Participants

Eighty students (51 women, aged 17–55 years,  $M = 25.5$ ) participated for course credit in the study, which was conducted at the University of Basel. Twenty participants were randomly assigned to each of four between-subjects conditions (see below).

### 6.1.2. Materials

As in Suter et al. (2016; see also Pachur et al., 2014), we used the following 12 side effects as medical outcomes: fatigue, flatulence, diarrhea, fever, itching, trembling, dizziness, insomnia, depression, speech disorder, hallucinations, and memory loss; monetary losses were used as monetary outcomes. Participants completed four tasks: a monetary evaluation task, two choice tasks (monetary and medical choice problems), and an affective evaluation task.

In the *monetary evaluation task*, participants indicated their willingness-to-pay (WTP; in Swiss Francs) to avoid each of the side effects. Specifically, they were asked to imagine that they were suffering from an illness (not further specified) and had to take medication for one week. Two medications are available. Both treat the illness equally well but one exacts a particular side effect with certainty (e.g., fatigue), whereas the other has no side effects. Participants indicated the amount of money they would be willing to pay to receive the medication without the side effect rather than the medication with the side effect.

Next, participants were presented with *choice tasks*. In one, they chose between two medications, each of which could lead to a particular side effect with some probability. In the other, they chose between two monetary gambles, each of which could lead to the loss of an amount of money with some probability. We constructed these problems using the individual-specific WTPs garnered in the monetary evaluation task and replacing the side effects with the respective WTPs. As a result, the monetary choice problems matched the medical ones monetarily. Using the entire range of probability levels and taking advantage of the median WTPs for side effects obtained in Pachur et al. (2014), we constructed 44 medical choice problems (see Suter et al., 2016, for a complete list of problems) such that an EV strategy and the minimax heuristic made different predictions in at least 90% of the problems. Finally, in an *affective evaluation task*, participants rated the amount of negative affect that the side effects and monetary losses triggered (see the Supplemental Material for details).

### 6.1.3. Design and procedure

After first completing the monetary evaluation task, participants rendered choices in 44 medical and 44 monetary problems. The order of the two classes of problems was counterbalanced across participants. Because there were individual differences in how the side effects were evaluated monetarily, the outcomes in the monetary problems varied across participants. Finally, in the affective evaluation task, we counterbalanced (across participants) whether monetary amounts or side effects were presented first. Overall, we employed a  $2 \times 2 \times 2$  design, with the class of choice problems (monetary vs. medical) as a within-subjects factor and the order of medical versus monetary choice problems as well as the order of medical or monetary outcomes in the affective evaluation task as between-subjects factors.

## 6.2. Results

### 6.2.1. How do the CPT parameters characterize monetary and medical choices?

We estimated CPT's parameters for each participant's choices, separately for the monetary and the medical problems.<sup>9</sup> The parameters were estimated using a Bayesian hierarchical approach, where individual-level parameters are assumed to be drawn from a group-level distribution (e.g., Scheibehenne & Pachur, 2015; see Appendix D for details). Parameters were estimated with separate group-level distributions for the monetary and the medical problems. Table 4 shows the means of the group-level distributions for each of the estimated parameters, separately for the two conditions. Fig. 4 depicts the parameter estimates on the individual-level (upper row; each point represents the mean of the posterior distribution of each participant) as well as the resulting value and weighting functions (middle and lower row, respectively). As can be seen, there were pronounced differences in the patterns obtained for monetary and medical choices. Specifically, the  $\gamma$  parameter was considerably

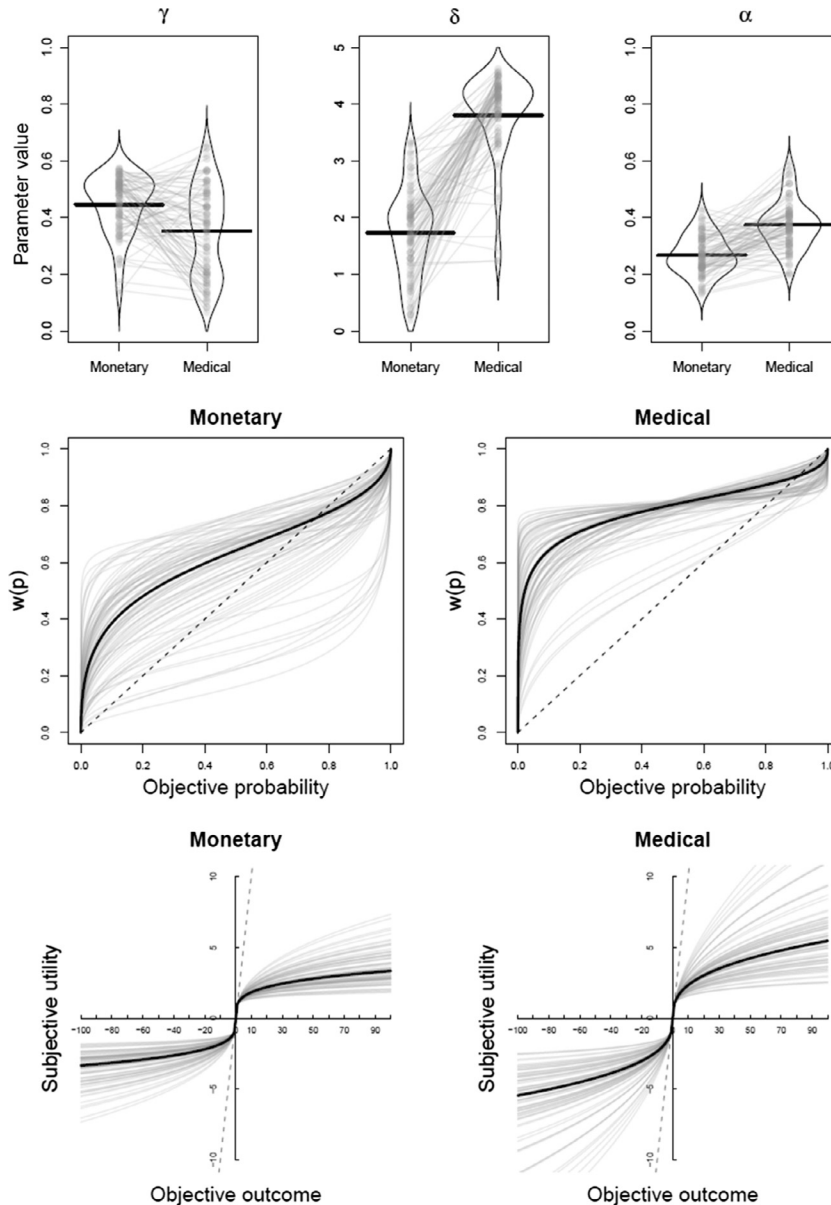
<sup>9</sup> Because all options involved negative (and zero) outcomes and in order to avoid overparameterization, CPT's loss aversion parameter ( $\lambda$ ) was set to be 1. Moreover, we estimated only a single  $\delta$ .

**Table 4**

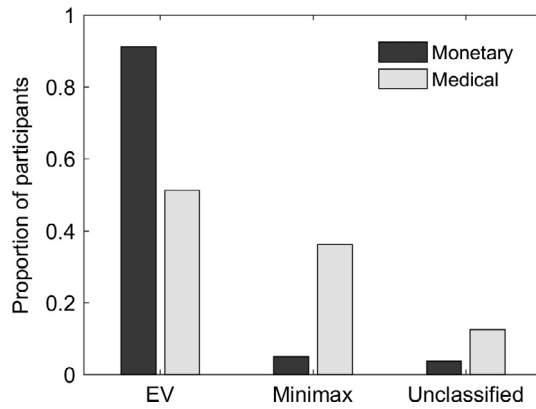
Median group-level estimates for cumulative prospect theory (CPT) parameters obtained for choices in the monetary and medical choice tasks and median difference of the posterior distributions. 95% highest density intervals are in brackets.

Parameter	Choice task		Parameter difference between monetary and medical choice task
	Monetary	Medical	
$\gamma$	0.439 [0.387, 0.501]	0.323 [0.243, 0.416]	0.117 [0.010, 0.217]
$\delta$	1.542 [1.180, 2.008]	4.529 [3.384, 4.978]	-2.954 [-3.612, -1.763]
$\alpha$	0.260 [0.223, 0.300]	0.367 [0.297, 0.447]	-0.107 [-0.195, -0.026]
$\varphi$	4.987 [4.898, 5.000]	1.339 [0.909, 1.967]	3.639 [3.007, 4.073]

Note. The deviance information criterion (DIC) for CPT in the monetary and medical choice tasks was 2397.16 and 3960.42, respectively (with a DIC for guessing of 4879.76).



**Fig. 4.** The upper row shows the distribution of estimated CPT parameters for the weighting function ( $\gamma$  and  $\delta$ ) and value function ( $\alpha$ ), separately for monetary and medical problems. Each point represents the mean of the individuals' posterior distribution on the respective parameter. The panels in the middle row show the weighting functions, separately for the monetary (left) and medical (right) choice task. The gray lines are the functions based on the parameters on the individual level; the black lines are the functions based on the group-level means. The panels in the bottom row show the value functions.



**Fig. 5.** Proportion of participants classified as following the compensatory expected value strategy (EV), the noncompensatory minimax heuristic, or who remained unclassified, separately for the monetary and the medical choice tasks.

lower in the medical than in the monetary choice task, indicating that participants' choices were less sensitive to probabilities in the former than in the latter. The  $\delta$  parameter was considerably higher in the medical choice task—as reflected by a much more elevated weighting function (see middle row, of Fig. 4)—suggesting that participants were more risk averse in the medical than in the monetary problems (Gonzalez & Wu, 1999). Furthermore, as indicated by values of the  $\alpha$  parameter, CPT's value function was less strongly curved in the medical than in the monetary problems, indicating higher outcome sensitivity (see lower row of Fig. 4).

The profile of CPT parameters in the medical problems—low probability sensitivity, high elevation, and high outcome sensitivity—is thus similar to that obtained for the minimax heuristic in the computer simulations (Fig. 1). The CPT analysis of the empirical choices might thus suggest systematic differences in strategy use across monetary and medical problems, with more use of the minimax heuristic in the latter domain. Next, we investigate this possibility directly.

### 6.2.2. Strategy classification

To examine potential differences in strategy use across medical and monetary problems, we modeled people's choices in both domains using two candidate strategies: (a) the EV strategy, according to which each option's outcomes, weighted by their respective probabilities, are integrated and the option with the highest EV is chosen; and (b) the minimax heuristic (Table 1). To derive predicted choices for the medical problems, we used each individual's WTPs (from the monetary evaluation task) as a proxy for how he or she evaluated the side effects. As in the model recovery study, we used a Bayesian latent mixture approach to classify each participant as using the strategy that best described their choices. The probability that person  $i$  chooses option A over option B on item  $j$  using strategy  $k$  is denoted as  $p(A, B)_{ijk}$ . For this analysis, the probability is defined as

$$p(A, B)_{ijk} = \frac{e^{\phi_{ki} V_k(A)}}{e^{\phi_{ki} V_k(A)} + e^{\phi_{ki} V_k(B)}} \times g_{ki} + \frac{g_{ki}}{2}, \quad (5)$$

where  $\phi$  is a choice sensitivity/scaling parameter (see Appendix D), and  $g$  (ranging between 0 and 1) is the probability of random guessing; both parameters are estimated for each person and for each strategy from the data.<sup>10</sup>

For the EV strategy, the subjective valuations of options A and B,  $V(A)$  and  $V(B)$ , were defined as  $V_{EV}(A) = x_A \times p_A$  and  $V_{EV}(B) = x_B \times p_B$  (with  $x$  and  $p$  being the outcome and probability of the nonzero outcomes of the option, respectively); for minimax, the valuations were defined as  $V_{MI}(A) = x_A$  and  $V_{MI}(B) = x_B$ . In addition, we modeled the data with a pure guessing strategy, which assumes that  $p(A, B)_{ij} = 0.5$  across all trials.

Each participant was classified based on the posterior distribution of the latent mixture variable  $z$ , separately for the monetary and the medical condition. The median (across participants) estimated values of the  $\phi$  and the  $g$  parameters for the EV strategy were 2.96 (2.47) and 0.35 (0.58), respectively, in the monetary (medical) problems; for minimax, the values were 2.50 (2.51) and 0.50 (0.50), respectively. Fig. 5 shows that more participants were classified as following the EV strategy in the monetary than in the medical problems (91.3% vs. 51.3%,  $\chi^2(1) = 25.03$ ,  $p = 0.001$ ; with McNemar's test). For minimax, the opposite pattern emerged (5.0% vs. 36.3%,  $\chi^2(1) = 19.9$ ,  $p = 0.001$ ). The percentage of participants who were best described by guessing, and thus assigned to the category "unclassified", was similar across both domains (3.8% vs. 12.5%,  $\chi^2(1) = 2.77$ ,  $p = 0.096$ ). Consistent with the specific pattern of CPT parameters obtained for the medical choices, the strategy classification suggests that the minimax heuristic was used more frequently in the medical than in the monetary domain. The mapping of parameter profile and minimax is further supported in Table 5. It shows the median individual-level CPT parameters, separately for participants who were best described by the EV strategy and the minimax heuristic, focusing

<sup>10</sup> We also tested a model without the  $g$  parameter. Although the results yielded a similar qualitative classification pattern, the model fit was considerably worse (with DICs of 3801.50 vs. 3512.77 and 4472.66 vs. 4306.89 in the monetary and medical conditions, respectively).

**Table 5**

Average (across individual participants) cumulative prospect theory (CPT) parameter estimates in the medical choice task for participants classified as following the EV strategy or the minimax strategy, respectively.

CPT parameter	Classified strategy in medical choice task	
	EV ( <i>n</i> = 41)	Minimax ( <i>n</i> = 29)
$\gamma$	0.456	0.207
$\delta$	3.653	4.052
$\alpha$	0.356	0.425
$\varphi$	1.729	1.438

(due to the lack of variability in strategy use in the monetary problems) on the medical problems. The CPT parameter profile obtained here is even more similar to the profile of the minimax heuristic (see Fig. 1 and Appendix E).

Overall, the results illustrate that the use of CPT to model variability in choice needs not stop at fitting choices. CPT can also be used to make inferences about differences in information processing—from compensatory integration to noncompensatory probability neglect (as implemented by minimax)—and to indicate the operation of a heuristic. CPT and heuristics can thus complement each other in describing choice behavior. Consistent with the conclusion that medical choices involve less attention to probability information, Pachur et al. (2014; Study 3) found in a process-tracing study that people paid less attention to probability information in medical than in monetary problems (see also Suter et al., 2015). We acknowledge, however, that a portion of participants remained unclassified, meaning that they probably employed other strategies.

## 7. Model recovery: how well can the heuristics be recovered based on their CPT parameter profiles?

The previous analysis showed that the CPT parameter profile of heuristics can guide the search for the operation of a specific heuristic in an empirical data set. Next, we conducted a model recovery study to address this issue more generally. Specifically, we asked how well the minimax heuristic, the maximax heuristic, the priority heuristic, the least-likely heuristic, and the most-likely heuristic can be recovered based on their specific profiles on CPT's parameters (as identified in our main analysis; see Table E1 in Appendix E). To that end, we modeled sets of choices produced by the five heuristics with five "variants" of CPT, each representing CPT with the parameter profile of a specific heuristic. We again varied the amount of noise in the heuristics' choices to see how this would affect recovery performance. The question was how well CPT with the parameter profile that best captured a specific heuristic would also be the one that best models the heuristic's choice pattern. This is relevant because CPT, even when fitted to the choices of a heuristic, is only able to represent the heuristic's choices to some extent, but not perfectly (see Table E1 in Appendix E).

### 7.1. Data generation

We used the same 180 two-outcome problems (60 gain, 60 loss, and 60 mixed problems) as in the other model recovery study. In a first step, choices were generated for the five heuristics on each of the problems. Sets with different error rates were again implemented by varying the percentage of (randomly selected) problems in which the prediction of a given model was reversed. Preliminary analyses showed that the largest changes in recovery performance occurred at error rates between 5% (9 of the 180 choices reversed) and 27.8% (50 of the 180 choices reversed). We therefore report the recovery rates mainly within this interval.

### 7.2. Model recovery

We modeled the heuristics' choices using the choice probabilities (for each of the 180 problems) predicted by CPT based on the five heuristics' parameter profiles (see Appendix D and E), as well as random choice (i.e., constant choice probability of 0.5). With a Bayesian latent mixture approach, we classified each heuristic's set of choices to one of the five CPT profiles or as random choice, based on the posterior distribution of the latent mixture variable  $z$ . Ideally, the recovered CPT profile model was the one that corresponded to the heuristic that actually generated the choices. For each error rate, the procedure was repeated 1000 times for each generating heuristic.

### 7.3. Results

As shown in Table 6, with an error rate of 0% for all heuristics the recovery performance was perfect—that is, the operation of a heuristic was always accurately detected based on their respective CPT profiles. Recovery was highly accurate for all heuristics up to an error rate of 5% (9 out of 180 problems); beyond that, recovery performance became worse. This was the case, in particular, for the maximax heuristic, which was often recovered as having been produced by the CPT profile of the most-likely heuristic. For the other heuristics, recovery performance became gradually worse with increasing error rates (with increasingly more assignments to the guessing mechanism), first for minimax, then for the priority heuristic,

**Table 6**

Proportion of classifications as a function of generating heuristic and error rate in the model recovery study. Correct recoveries are in bold.

Error rate	Generating heuristic	Recovered CPT profile					Recovered as guessing
		MI	MA	PH	LL	ML	
0%	MI	<b>1</b>	0	0	0	0	0
	MA	0	<b>1</b>	0	0	0	0
	PH	0	0	<b>1</b>	0	0	0
	LL	0	0	0	<b>1</b>	0	0
	ML	0	0	0	0	<b>1</b>	0
5%	MI	<b>0.898</b>	0	0	0	0	0.102
	MA	0	<b>0.871</b>	0	0	0.183	0
	PH	0	0	<b>0.993</b>	0	0	0.007
	LL	0	0	0	<b>1</b>	0	0
	ML	0	0	0	0	<b>0.990</b>	0.010
5.6%	MI	<b>0.816</b>	0	0	0	0	0.184
	MA	0	<b>0.527</b>	0	0	0.473	0
	PH	0	0	<b>0.990</b>	0	0	0.010
	LL	0	0	0	<b>1</b>	0	0
	ML	0	0	0	0	<b>0.980</b>	0.020
6.7%	MI	<b>0.656</b>	0	0	0	0	0.344
	MA	0	<b>0.114</b>	0	0	0.886	0
	PH	0	0	<b>0.953</b>	0	0	0.047
	LL	0	0	0	<b>1</b>	0	0
	ML	0	0	0	0	<b>0.946</b>	0.054
8.3%	MI	<b>0.359</b>	0	0	0	0	0.641
	MA	0	<b>0.004</b>	0	0	0.996	0
	PH	0	0	<b>0.889</b>	0	0	0.111
	LL	0	0	0	<b>0.999</b>	0	0.001
	ML	0	0	0	0	<b>0.887</b>	0.113
13.9%	MI	<b>0.012</b>	0	0	0	0	0.988
	MA	0	<b>0</b>	0	0	0.998	0.020
	PH	0	0	<b>0.328</b>	0	0	0.672
	LL	0	0	0	<b>0.960</b>	0	0.040
	ML	0	0	0	0	<b>0.571</b>	0.429
27.8%	MI	<b>0</b>	0	0	0	0	1
	MA	0	<b>0</b>	0	0.001	0.242	0.757
	PH	0	0	<b>0.001</b>	0	0	0.999
	LL	0	0	0	<b>0.437</b>	0	0.563
	ML	0	0	0	0.002	<b>0.076</b>	0.922
50%	MI	<b>0</b>	0	0	0.023	0	0.977
	MA	0	<b>0</b>	0	0.020	0	0.980
	PH	0	0	<b>0</b>	0.020	0.001	0.979
	LL	0	0	0	<b>0.015</b>	0	0.985
	ML	0	0	0	0.021	<b>0</b>	0.979

Note. MI = minimax heuristic, MA = maximax heuristic, PH = priority heuristic, LL = least-likely heuristic, ML = most-likely heuristic.

and the most-likely heuristic. For the least-likely heuristic, about half the cases were correctly recovered even with an error rate of 27.8%. Overall, these results show that even though CPT cannot mimic the heuristic choices perfectly, the CPT profiles of the heuristics are diagnostic for detecting the operation of the underlying heuristic, at least when the amount of noise is not overly pronounced.

The one qualification is that maximax was incorrectly attributed to the CPT profile of the most-likely heuristic already at quite low error rates. There are two reasons for this. First, as can be seen in Table 7, showing the overlap in choice between the heuristics and CPT assuming the heuristics' parameter profile, there was a high overlap (73%) between the choices predicted by maximax and those predicted by CPT with the most-likely heuristic's parameter profile. Second, the choice probabilities predicted by CPT with the parameter profile of the most-likely heuristic were less variable across the 180 problems than those predicted by CPT with the parameter profile of maximax (with standard deviations of 0.23 vs. 0.48, respectively). This renders the former less complex, thus giving it an edge in the Bayesian latent mixture analysis.

## 8. General discussion

We showed how two disjoint approaches to modeling risky choice can be related to each other, despite their fundamentally different algorithmic structures. Such analyses can help to integrate psychological theory and to build bridges between



**Table 7**

Overlap in predicted choices between the heuristics and CPT, using the parameter profile produced by the respective heuristic (see Table E1 in Appendix E). The overlap of the choices of the heuristic with those of its corresponding variant of CPT are in bold.

		CPT based on the respective heuristic's parameter profile				
		MI	MA	PH	LL	ML
Heuristic	MI	<b>0.744</b>	0.433	0.628	0.461	0.456
	MA	0.578	<b>0.956</b>	0.661	0.517	0.733
	PH	0.594	0.606	<b>0.689</b>	0.483	0.511
	LL	0.283	0.328	0.311	<b>0.694</b>	0.522
	ML	0.517	0.539	0.489	0.556	<b>0.661</b>

Note. MI = minimax heuristic, MA = maximax heuristic, PH = priority heuristic, LL = least-likely heuristic, ML = most-likely heuristic.

economics, which relies almost exclusively on algebraic models to describe risk choice, and psychology, which studies both algebraic models and process-oriented heuristic models. Both modeling traditions stand to benefit from such integration. Next we discuss some of the potential benefits. Then we discuss the empirical evidence for heuristics in risky choice, CPT's limits in representing heuristic choice, and how CPT represents heuristics when assuming an extended parameter range.

### 8.1. Benefits for algebraic models

#### 8.1.1. From as-if models to psychometric tools

Algebraic models, with their focus on describing preference patterns, are mute about the cognitive processes underlying choice. Milton Friedman (1953) famously argued that in modeling human behavior, “complete ‘realism’ is clearly unattainable, and the question whether a theory is realistic ‘enough’ can be settled only by seeing whether it yields predictions that are good enough for the purpose at hand” (p. 41). For Friedman, the purpose at hand was to capture the behavior of economic institutions, not individuals. Consequently, models such as expected utility theory can be tested only in terms of their behavioral predictions (see, e.g., Gul & Pesendorfer's, 2005, defense of “mindless economics”).

CPT analyses usually do not (at least not explicitly) assume that people are able to or *actually* do calculate subjective values and decision weights according to Eqs. (1), (2), and (4); the CPT framework thus stands in the tradition of algebraic models. Nevertheless, CPT postulates psychological constructs such as loss aversion and risk aversion, whose cognitive and neural roots have been investigated (e.g., Tom, Fox, Trepel, & Poldrack, 2007; Willemsen, Böckenholt, & Johnson, 2011). Our analyses demonstrate still another way to connect CPT (and, by extension, other algebraic models invoking psychological constructs) to underlying cognitive processes. To wit, algebraic models can be employed as measurement tools to characterize, for instance, heuristic processes (e.g., lexicographic search) and the individuals enlisting them in terms of the constructs built into the models' algebraic functions (e.g., risk aversion, loss aversion, and probability sensitivity). Such a “cognitive psychometrics” perspective on formal models is popular in memory research (Batchelder, 1998; Riefer, Knapp, Batchelder, Bamber, & Manifold, 2002), but has barely been invoked in work modeling the cognitive underpinnings of risky choice (and variability therein).

#### 8.1.2. Cognitive foundation of CPT

Prospect theory has psychophysical roots that Kahneman and Tversky (1979) highlighted, for instance, in the context of diminishing sensitivity: “Many sensory and perceptual dimensions share the property that the psychological response is a concave function of the magnitude of physical change” (p. 278). Without challenging either this reference to psychophysics, or Stewart et al.'s (2006) elegant elaboration of possible ecological roots of CPT's functions, we propose that linking CPT with heuristics offers still another foundation of the empirically observed value and weighting functions. From this perspective, individual differences in the shapes of these functions may stem, at least in part, from distinct heuristic principles of information processing.

Connecting CPT with models of heuristics also permits an information-processing interpretation of differences in CPT's parameters due to factors such as gender (Fehr-Duda, De Gennaro, & Schubert, 2006), age (Harbaugh, Krause, & Vesterlund, 2002; Pachur, Mata, & Hertwig, in press) and criminal tendencies (Pachur, Hanoch, & Gummerum, 2010). Fehr-Duda et al., for instance, concluded that women tend to be less sensitive than men to differences in probabilities (see also Booij & van de Kuilen, 2009). One reason could be that they employ different choice heuristics.

We should, however, also point out a potentially important drawback of invoking heuristic processes as possible cognitive underpinnings of characteristic shapes of CPT. For many algebraic models, it has been shown that if the decision maker satisfies particular axioms (e.g., independence, transitivity), it is possible to derive the properties of the underlying utility scale (e.g., von Neumann & Morgenstern, 1947). Some heuristic processes have been shown to violate some of these fundamental axioms (e.g., Tversky, 1969; but see Katsikopoulos & Gigerenzer, 2008). If one accepts heuristics as underlying processes of choice, then such inferences about the utility measurement may not be warranted.

## 8.2. Benefits for models of heuristics

### 8.2.1. Characterizing and comparing heuristics in terms of CPT's constructs

Choice heuristics are usually compared with each other in terms of their building blocks, specifically, their search, stopping, and decision rules (see Brandstätter et al., 2006; Gigerenzer et al., 2011; Payne et al., 1993; see also Shah & Oppenheimer, 2008). Our analysis indicates that they can also be compared in terms of the constructs of algebraic models, such as risk aversion, loss aversion, outcome sensitivity, and probability sensitivity. How the policy of a heuristic translates into these constructs need by no means be obvious, especially when the properties of a heuristic's choices interact with the choice environment. For instance, we found the most-likely heuristic to be risk neutral in randomly constructed gambles (i.e., Table E1 in Appendix E); however, in environments in which outcomes and probabilities were negatively correlated, the heuristic was risk averse in the gain domain and risk seeking in the loss domain (Table E3).

### 8.2.2. Multi-method approach to identifying processes

Models of heuristics are meant to be accounts of both the content and the outcome of the choice process. Our findings suggest a multi-method approach to process identification. Alongside algebraic tools to measure properties of the process, it is also possible to harness process tracing. For instance, Pachur, Schulte-Mecklenbeck, Murphy, and Hertwig (submitted for publication) combined Mouselab, a process-tracing methodology, with CPT and found that decision makers with lower probability sensitivity—as measured by the weighting function parameter  $\gamma$ —paid less attention to probabilities (inferred from how long they inspected probability information) than decision makers with higher probability sensitivity. Conversely, decision makers with higher outcome sensitivity—as measured by the parameter  $\alpha$  in the value function—allocated more attention to outcomes. Moreover, the amount of time they inspected negative relative to positive outcomes correlated positively with the loss aversion parameter  $\lambda$ .

Conversely, it should be emphasized that CPT's ability to accommodate choices generated by heuristics disallows the conclusion that a good model fit of CPT means that the choices were necessarily brought about by a compensatory, expectation-based cognitive mechanism. Instead, it is important to further unpack the processes behind the choices (e.g., by using process-tracing data)—they might be noncompensatory after all.

## 8.3. Heuristic information processing in risky choice

The concern of our analysis was conceptual: What footprints would various *possible* (and proposed) heuristic information-processing rules—varying in, for instance, their consideration of probability information and attention to extreme outcomes—leave in CPT's parametric framework? Having addressed this question, let us briefly turn to the empirical question of whether risky choice appears to be shaped by heuristic processes.

Consistent with the heuristics' signature of limited search and skewed attention to available information, it has been concluded (from observed choices patterns) that people place a strong “emphasis on the amount of money associated with the smallest probability” (Lichtenstein, 1965, p. 168), and process risky options based on a lexicographic ordering of relevant reasons. Indicators for limited search have also emerged in process-tracing studies (Mann & Ball, 1994; Payne & Braunstein, 1978; Russo & Doshier, 1983; Su et al., 2013). In line with Tversky's (1972) suggestion that people simplify risky choice by comparing probability and outcome information separately (and across gambles) rather than integrating it, some studies have found evidence for predominantly between-gamble transitions (Mann & Ball, 1994; Payne & Braunstein, 1978; Rosen & Rosenkoetter, 1976; Russo & Doshier, 1983; Su et al., 2013). In addition, the amount of information examined often varies substantially across choice problems—a finding that has been interpreted as indicative of lexicographic heuristics (e.g., Mann & Ball, 1994; Payne & Braunstein, 1978; Russo & Doshier, 1983; Slovic & Lichtenstein, 1968).

Several investigations have concluded that, at least in some choice environments, models of heuristics offer a promising approach (e.g., Pachur et al., 2013). For instance, Venkatraman et al. (2009), studying complex multi-outcome gambles, found that some participants nearly always employed a heuristic that maximizes the probability of winning and ignores reward magnitude (see also Venkatraman, Payne, & Huettel, 2014). Furthermore, enlisting process-tracing methodologies such as neuroimaging and verbal protocols, Rao, Li, Jiang, and Zhou (2012) and Cokely and Kelley (2009) found evidence consistent with the use of noncompensatory heuristics in risky choice.

Other studies, however, have obtained little supportive evidence for heuristics. For instance, empirical tests of the priority heuristic found that people do not examine risky options as predicted by the heuristic (e.g., Birnbaum & LaCroix, 2008; Fiedler, 2010; Glöckner & Betsch, 2008; Glöckner & Herbold, 2011). For instance, in a process-tracing analysis, Johnson et al. (2008) observed that people transition between outcomes and probabilities more frequently than predicted by the heuristic. In addition, CPT outperformed the priority heuristic—and 12 other heuristics—in predicting individual choices (Glöckner & Pachur, 2012).

At this point, the following conclusion seems fair: Whereas indicators of cognitive processes in risky choice often, but not always, signal the operation of heuristic principles of information processing, no existing model of a heuristic can explain risky choice across a large range of choice ecologies and task structures. One possible explanation is that people switch between different processes (some perhaps more heuristic than others) across trials in ways that are not yet understood (but see Venkatraman et al., 2009). However, the contribution of specific heuristic processes seems to be systematically higher under certain conditions, such as time pressure or affective content (Pachur et al., 2014; Payne et al., 1993; Young,

Goodie, Hall, & Wu, 2012; see also our analysis of medical and monetary choice reported above). Our analyses suggest that CPT can be used as a tool to track such differences in processing.

#### 8.4. Why does CPT fit some heuristics better than others?

Although CPT is able to accommodate choices generated by various heuristics, our analyses also clearly showed limits in CPT's ability to represent the choice patterns of some of the heuristics. First, for some heuristics, the posterior distributions for several parameters had a high mass at the parameter boundaries, indicating inadequacies in model specification (Appendix F). Second, model fit differed across the heuristics (Table E1 in Appendix E). Third, CPT's fit to the priority heuristic's choices differed as a function of the frequency with which the top-ranked reason discriminated. Fourth, CPT's model fit varied as a function of whether outcomes and probabilities were negatively correlated or uncorrelated. Why does CPT capture some heuristics better than others, and why does model fit depend on the environment?

One factor that affects CPT's ability to fit a heuristic's choices is whether the rank of the outcomes (among all possible outcomes, ordered by magnitude) that determines the choice is the same across all problems, or whether it varies. For the minimax and the maximax heuristic, the rank is invariant: these heuristics always base their choices on the best or the worst outcome, respectively. CPT can emulate such policies (resulting in a good model fit) by assuming relatively extreme parameter values in the weighting function, thus giving all the weight to the first or last ranked outcome. The most-likely heuristic, by contrast, bases its choice on the outcome with the highest probability, and this outcome can be the best, worst, or a medium outcome (at least in environments in which outcomes and probabilities are uncorrelated). Consequently, CPT's ability to emulate the choices of this heuristic is more limited.

The variability of the rank of the outcome on which a heuristic bases a choice also explains why CPT captures the choices of the most-likely heuristic better with negatively correlated than with uncorrelated outcomes and probabilities. When outcomes and probabilities are negatively correlated, the most-likely heuristic always chooses based on the smallest gains or the smallest losses, as these are the most probable outcomes. In uncorrelated environments, the rank of the decisive outcome will be variable, as the most probable outcome can be the highest, lowest, or a medium outcome. Therefore, CPT fits the choices of the most-likely heuristic better in the former than in the latter environment.

Finally, why did we observe that CPT's model fit with the priority heuristic was best when the heuristic's highest ranked reason (minimum outcomes) discriminated in 75% of the problems, decreased when it discriminated in 50%, and then improved again when it discriminated in only 25% of the problems? As explained earlier, the priority heuristic's first two processing steps embody sequential application of the minimax and the least-likely heuristic. As can be seen from the resulting value and weighting functions (Fig. 1), these two heuristics have distinct "parametric personalities." CPT seems better able to fit choices when all or most are brought about by a single policy (either minimax or least-likely). When choices are based on two distinct policies with approximately equal frequency (the 50% choice set), in contrast, CPT is less able to accommodate them using a single set of parameters. It follows that the CPT framework would have more difficulties capturing an ensemble of different choice processes (e.g., aggregated across heterogeneous problem types or across individuals).

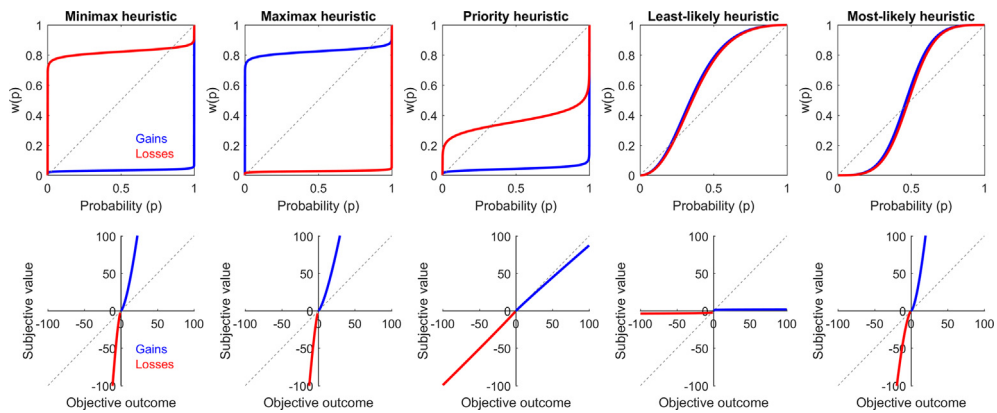
It is important to keep these limitations of CPT in capturing choice patterns conceptually separate from the limitations of CPT identified by Birnbaum and others (e.g., Birnbaum, 2008; Wu & Markle, 2008). The latter concern CPT's *empirical* shortcomings—for example, that decision makers sometimes violate CPT's assumptions of coalescing, stochastic dominance, and gain–loss separability. These violations can be explained by assuming particular ways in which attention is allocated to rank-ordered outcomes (e.g., Birnbaum, 2008). Importantly, these patterns of attention allocation are constant across choice problems. In the heuristics that CPT has difficulty emulating, by contrast, the pattern of attention allocation (i.e., which outcome determines a choice) varies across problems.

#### 8.5. Heuristics reflected when CPT parameter ranges are extended

To what extent do the previously discussed limits in CPT's ability to capture choice patterns compromise efforts to relate CPT and heuristics? On the one hand, evidence for such limits is important in and of itself. It underlines that CPT, notwithstanding its functional complexity and multiple adjustable parameters, is not infinitely flexible. However, even within the parameter bounds (reflecting CPT's characteristic conceptual assumptions), CPT was flexible enough to represent distinct heuristics by assuming different shapes in the weighting and value functions; in addition, these differences were meaningfully associated to differences in the heuristics' information processing architectures. On the other hand, the limits on the range of CPT parameter values are likely to restrict the theory's flexibility as a measurement tool.

To gauge the impact of this limitation, we reran the computer simulations with an extended (but still reasonable) parameter range. Specifically, the parameters were set to  $0 \leq \alpha \leq 3$ ;  $0 \leq \lambda \leq 10$ ;  $0 \leq \gamma \leq 3$ ;  $0 \leq \delta^* \leq 10$ ;  $0 \leq \varphi \leq 5$ , thus also rendering possible, for instance, a convex (concave) value function in gains (losses) (with  $\alpha > 1$ ), an S-shaped weighting function (with  $\gamma > 1$ ), and gain seeking (with  $\lambda < 1$ ). How did CPT accommodate the heuristics' choices with these extended parameter ranges?

It was only in a few cases that the increased model flexibility endowed CPT with a considerably better ability to accommodate the heuristics' choices: For the most-likely heuristic in the choice ecology with problems with similar EVs, and for minimax and maximax in the ecology with problems with dissimilar EVs and the negatively correlated environment, the DIC was considerably lower (indicating a much better ability to capture the data) and the posterior predictives were more accu-



**Fig. 6.** CPT's weighting function (upper row) and value functions (lower row) estimated for the choices of all five heuristics for problems with two-outcome gambles using an extended parameter range (see the [Supplemental Material](#) for the parameter values).

rate (see [Supplemental Material](#)). For all other heuristics, however, model fit was only slightly better than with the more restricted parameter ranges.

What about the estimated parameters? For illustration, [Fig. 6](#) shows the resulting weighting and value functions for problems with two-outcome gambles with similar EVs (the posterior distributions of the parameters as well as the parameter estimates and the results for the other computer simulations are reported in the [Supplemental Material](#)). There are two main differences relative to the previous analyses (e.g., [Fig. 1](#)). First, for the least-likely and most-likely heuristics, the weighting functions were now S-shaped instead of linear. Importantly, however, an S-shaped weighing function still indicates very high sensitivity (in fact, hypersensitivity) to differences in probabilities—consistent with these heuristics' strong focus on probability information. Second, for minimax, maximax, and the most-likely heuristic, the value functions indicated hypersensitivity to differences in outcomes (i.e., a strongly convex value function in gains and a concave value function in losses)—again consistent with these heuristics' focus on outcomes.

In summary, a more permissible parameter space increased—albeit not drastically—CPT's ability to describe choices generated by heuristics. Similarly, some of the heuristics' parametric profiles changed noticeably, but estimates of CPT's weighting and value functions still veridically reflected properties of the heuristics' information processing. Furthermore, in both sets of analyses—with narrower versus wider parameter ranges—CPT was able to represent the distinct heuristics. In our view, there is no 'right' or 'wrong' parameter range to implement CPT. In the case of the narrower range, the ability of CPT as a measurement model is somewhat more limited; yet this parameter range reflects the theoretical commitments of the theory and allows for comparisons with other CPT analyses. In the case of the wider range, CPT becomes a more general measurement tool, covering a wide range of shapes of the value and weighting functions. There are good arguments for either choice of range. The most important point is to consistently use one range within an analysis or to conduct a comparative analysis. We thus conclude that, even with expanded parameter bounds (going beyond the assumptions of CPT), estimates of CPT's weighting and value functions veridically reflect properties of the heuristics' information processing architectures.

## 9. Conclusion

Algebraic models and models of heuristics represent two prominent but disjoint approaches to capturing risky choice. Yet the gulf between them is not as wide as it is often suggested and perceived to be. Connecting the two approaches affords substantial benefits for both. Moreover, the integrative perspective on risky choice adopted here need not be limited to CPT and heuristics. Other models that could likewise be used to measure different heuristic policies on psychological dimensions include the transfer-of-attention-exchange model ([Birnbaum & Chavez, 1997](#)), the perceived relative argument model ([Loomes, 2010](#)), and decision field theory ([Busemeyer & Townsend, 1993](#)), as well as other diffusion models (e.g., [Ratcliff, 1978](#)). Theory integration is only one corrective to the often lamented theoretical fragmentation in psychology but, we submit, a powerful and underused one.

## Author note

Thorsten Pachur, Renata S. Suter, and Ralph Hertwig, Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany.

This research was supported by a grant from the German Research Foundation (DFG) as part of the priority program on *New Frameworks of Rationality* (SPP 1516) to Ralph Hertwig and Thorsten Pachur (HE 2768/7-2). We thank Valerie M. Chase and Susannah Goss for editing the manuscript.

## Appendix A

### A.1. Predictions for CPT's weighting function for the different heuristics

As described in Eq. (3), CPT assumes a rank-dependent transformation of the outcomes' probabilities into decision weights. As a consequence, the decision weight resulting from a probability depends on the rank of its respective outcomes (among the other outcomes). For instance, when a choice problem has three positive outcomes ( $o_1$ ,  $o_2$ , and  $o_3$  in ascending order), the decision weights will be  $w(p_3)$  for the highest outcome  $o_3$ ,  $w(p_2 + p_3) - w(p_3)$  for the medium outcome  $o_2$ , and  $1 - w(p_2 + p_3)$  for the lowest outcome  $o_1$ . How can the decision weights in CPT accommodate heuristic choices?

#### A.1.1. Minimax

For minimax, which considers the lowest outcome only, the decision weights of the three outcomes are  $1 - w(p_2 + p_3)$ , that is,  $1 - 0 = 1$  for the lowest outcome  $o_1$ ;  $w(p_2 + p_3) - w(p_3) = 0 - 0 = 0$  for the medium outcome  $o_2$ ; and  $w(p_3) = 0$  for the highest outcome  $o_3$ . In order to obtain decision weights of 0 for the medium and the highest outcome, the weighting function has to be extremely inverse S-shaped with an extremely low elevation.

#### A.1.2. Maximax

For maximax, which considers the highest outcome only, the decision weights of the three outcomes are  $1 - w(p_2 + p_3)$ , that is,  $1 - 1 = 0$  for the lowest outcome  $o_1$ ;  $w(p_2 + p_3) - w(p_3) = 1 - 1 = 0$  for the medium outcome  $o_2$ ; and  $w(p_3) = 1$  for the highest outcome  $o_3$ . In order to obtain decision weights of 0 for the lowest and the medium outcome, the weighting function has to be extremely inverse S-shaped with an extremely high elevation.

#### A.1.3. Priority heuristic

The priority heuristic first inspects the minimum outcomes and turns to their probabilities only if the minimum outcomes do not discriminate between the options (see Table 1). Its processing scheme is therefore a blend of minimax and the least-likely heuristic, which always decides based on the probabilities of the minimum outcomes (see below), and therefore results in a weighting function that lies between the two. The degree of curvature depends on the proportion of choices in which the heuristic turns to probability information (see our analysis of lexicographic search in the main text).

#### A.1.4. Least-Likely

For the least-likely heuristic, which decides based on the probabilities of the minimum outcomes, choices always take these probabilities into account. This means that the weighting function has to reflect this sensitivity as well as the lack of sensitivity to the higher outcomes' probabilities. With an inverse S-shaped weighting function, the highest outcomes would receive more weight than they should. Therefore, a weighting function with little (if any) curvature and a low elevation is predicted, as it retains the least-likely heuristic's sensitivity to the minimum outcomes' probabilities without giving more weight to higher outcomes.

#### A.1.5. Most-likely

The most-likely heuristic chooses the option whose most likely outcome is more attractive. To approximate this policy, the weighting function has to conserve the order of the probabilities, ensuring that the decision weight of the outcome with the highest probability is not smaller than the decision weights of other outcomes. In an environment in which outcomes and probabilities are uncorrelated (i.e., where the highest probability can be associated with every outcome), this can only be achieved by a linear weighting function, as otherwise it is possible that the decision weights given to the other outcomes are too high.

## Appendix B

### B.1. CPT profiles of heuristics in easy choice problems

The analyses in the main text were based on choice problems in which the difference between the options' EVs was relatively small and both options were thus similarly (un)attractive. Can CPT also capture heuristic choice policies in problems with dissimilar EVs, where one option is clearly superior to the other? Given its commitment to an expectation core, CPT might not be flexible enough to represent such choices, resulting in poor model fit and parameter patterns that do not reflect the heuristics' processing properties as meaningfully.

To address this question, we repeated the analyses described in the first computer simulation for problems in which the difference between the options' EVs was relatively large (for a description of the general problem construction procedure, see



**Table B1**

CPT parameter estimates for the choices of each of the five heuristics in problems with two-, three-, and five-outcome gambles, respectively, with dissimilar EVs. Also shown are the deviance information criterion (DIC) of CPT and the results of posterior predictive checks (PP) indicating the proportion of choices for which CPT, based on the posterior distributions of the parameters, generated the same choice as the respective heuristic.

Heuristic	Parameters						DIC	PP
	$\gamma$	$\delta^+$	$\delta^-$	$\alpha$	$\lambda$	$\phi$		
	Two-outcome gambles							
Minimax	0.05	0.06	4.33	0.77	1.44	0.65	56.42	0.94
Maximax	0.02	4.59	0.02	0.89	2.74	0.93	29.37	0.97
Priority	0.05	0.31	1.83	0.56	3.64	2.36	23.51	0.97
Least likely	0.89	2.86	2.18	0.06	3.40	3.81	164.30	0.69
Most likely	0.84	1.27	0.26	0.71	2.25	0.44	52.61	0.94
	Three-outcome gambles							
Minimax	0.04	0.11	4.51	0.89	2.30	0.17	99.82	0.87
Maximax	0.01	4.75	0.01	0.93	2.87	1.03	34.21	0.95
Priority	0.03	0.05	1.37	0.62	1.87	1.19	78.96	0.89
Least likely	0.91	0.39	3.50	0.24	1.45	2.04	170.44	0.70
Most likely	0.91	0.41	1.58	0.66	1.47	0.58	66.49	0.91
	Five-outcome gambles							
Minimax	0.02	0.04	4.47	0.81	1.79	1.11	49.66	0.93
Maximax	0.05	4.59	0.04	0.89	3.14	0.73	49.80	0.93
Priority	0.06	0.30	1.17	0.52	3.16	1.81	90.42	0.86
Least likely	0.87	1.34	3.58	0.42	2.63	0.48	201.59	0.64
Most likely	0.81	1.03	0.66	0.70	1.74	0.50	78.37	0.90

Note. The DIC assuming random choice was 249.53.

Appendix C). Specifically, we constructed problem sets with two-, three-, and five-outcome gambles with the constraint that the ratio of the absolute difference between the options' EVs to the absolute value of the smaller EV was larger than 1—thus ensuring they were dissimilar (in the previous simulation, the ratios was smaller than 1). Each set of problems again contained 60 gain, 60 loss, and 60 mixed problems. The ratios of the absolute difference in EVs to the absolute value of the smaller EV ranged from 1.6 to 145.6, 6.5 to 129.3, and 4.7 to 87.9 for the problems with two-, three-, and five-outcome gambles, respectively. We then determined each heuristic's choices in each of the three problem sets and estimated CPT's parameters for these sets of choices (again using Bayesian parameter estimation).

Table B1 reports CPT's ability to capture the heuristics' choices. As shown, the model fit was good for all five heuristics in all problem sets; in fact, it was better than in the simulation with similar EVs (as indicated by lower DICs and a higher proportion of overlapping choices). In particular, CPT captured the choices generated by the most-likely heuristic much better than before. One possible explanation is that the most-likely heuristic chose the option with the higher EV in a substantially larger percentage of problems than in the context of problems with similar EVs (i.e., in 96.7, 93.3, and 94.4% of the choices in problems with two-, three-, and five-outcome gambles, respectively, relative to 68.9, 67.2, and 67.8% of choices in the context of problems with similar EVs). Most importantly, the resulting parameter patterns were similar to those obtained for the problems with similar EVs (see Table E1 in Appendix E). For instance, the heuristics that prioritize outcome information yielded values close to 1 on the  $\alpha$  parameter, indicating high outcome sensitivity; heuristics that are risk averse for losses (gains) resulted in a strongly (weakly) elevated weighting function, whereas heuristics that are risk seeking resulted in a weak (strong) elevation. We conclude that CPT can veridically capture the processing characteristics of various heuristics, irrespective of whether the choice problems have options with similar or dissimilar EVs.

## Appendix C

### C.1. Construction of the choice problems

For the problems with two-outcome gambles, we used a set constructed by Rieskamp (2008; Study 2). The problem sets with the three-outcome and the five-outcome gambles were designed using the same principles as in Rieskamp's study. Specifically, we generated 10,000 pairs of gambles for the gain, loss, and mixed domains, separately for the three- and the five-outcome problems. The gambles for the mixed problems were constructed such that there were (with equal probability) either one or two (for three-outcome gambles) or two or three (for five-outcome gambles) gain outcomes. The outcomes were drawn with equal probability from integers ranging from 0 to 100 for gains and from -100 to 0 for losses. The probabilities for the three-outcome problems were constructed as follows: Two values from the interval 0 to 1 were selected randomly, dividing the interval into three sections. The smallest value constituted the first probability, the difference between the smallest value and the second value constituted the second probability, and the difference between the second value and 1 constituted the third probability. For illustration, we randomly selected two

values from the interval 0 to 1—for instance, 0.25 and 0.55—to construct probabilities for three-outcome gambles. The smallest value constituted the first probability—in this example, 0.25. The difference between the smallest value (0.25) and the second value (0.55), or 0.3, constituted the second probability. The difference between the second value (0.55) and 1, or 0.45, constituted the third probability. The three resulting probabilities—0.25, 0.3, and 0.45—sum to 1. The probabilities for the five-outcome problems were constructed using the same principle (i.e., based on four values randomly selected from the interval 0 to 1).

Problems in which one gamble stochastically dominated the other were eliminated. Specifically, following [Rieskamp \(2008\)](#), we eliminated those problems for which the ratio of the absolute difference of the gambles' EVs to the absolute value of the gamble with the smaller expected value was larger than 1. From the remaining problems, 60 were selected randomly for each problem set (i.e., gain, loss, and mixed problems), for both the problems with three- and the problems with five-outcome gambles. (As noted in the main text, some problem sets had additional criteria.)

## Appendix D

### D.1. Estimation of CPT parameters

To determine from CPT's predicted valuations of options A and B a probability of choosing option A over option B,  $p(A, B)$ , we applied the softmax choice rule (e.g., [Sutton & Barto, 1998](#)):

$$p(A, B) = \frac{e^{\phi \cdot V(A)}}{e^{\phi \cdot V(A)} + e^{\phi \cdot V(B)}}, \quad (\text{D1})$$

where  $V(A)$  and  $V(B)$  represent the predicted subjective valuations. The parameter  $\phi > 0$  governs how sensitively the model reacts to differences between  $V(A)$  and  $V(B)$ , with higher values indicating higher sensitivity (at  $\phi = 0$ , choices are random). Importantly, note that the value of  $\phi$  is also sensitive to the magnitude of the valuations  $V$ : Because  $\phi$  also serves to map the differences in the options' valuations onto a scale from 0 to 1, when the  $V$ s and differences between them are large (e.g., due to high values of  $\alpha$ ), the choice probability for an option can be high even with rather low values of  $\phi$  (for a discussion, see [Scheibehenne & Pachur, 2015](#)).

We used a Bayesian approach to estimate the CPT parameters for the choices produced by the different heuristics ([Nilsson et al., 2011](#); [Scheibehenne & Pachur, 2015](#)). Bayesian parameter estimation yields, based on prior distributions, posterior distributions of parameter values that are informed by the data; the distributions are characterized by a central tendency and a dispersion. The dispersion of the posterior distribution indicates the uncertainty in the estimate (the larger the dispersion, the higher the uncertainty). The prior distributions for the parameters were set to uniform probability distributions spanning a reasonable range that excluded theoretically impossible values but included parameter values found in previous research. Specifically, in the main analyses the ranges were 0–1 for  $\alpha$  and  $\gamma$ , 1–5 for  $\lambda$ , and 0–5 for  $\delta^+$ ,  $\delta^-$ , and  $\phi$ . The posterior parameter distributions were estimated using Monte Carlo Markov Chain methods implemented in JAGS, a sampler that utilizes a version of the BUGS programming language (version 3.3.0) called from Matlab. We ran four chains, each with 20,000 recorded samples, which were drawn from the posterior distributions after a burn-in period of 1000 samples. To reduce autocorrelations during the sampling process, we recorded only every 20th sample (thinning). The sampling procedures were efficient, as indicated by low autocorrelations of the sample chains, Gelman–Rubin statistics, and visual inspections of the chain plots.

In the hierarchical estimation approach used in the behavioral experiment (Section 6), the individual-level parameters were linked to group-level parameters through probit transformations ([Scheibehenne & Pachur, 2015](#)), yielding a distribution ranging from 0 to 1 on the individual level. In order to extend the range of these distributions from 0 to 5 for  $\delta$  and  $\phi$ , we interposed an additional linear linkage function. All hierarchical group-level means were assumed to be normally distributed, with a mean of 0 and a variance of 1. The priors on the group-level standard deviations were uniformly distributed, ranging from 0 to 10 (thus avoiding extreme bimodal distributions on the individual level). The number of chains, number of recorded samples, and amount of thinning were the same as in the nonhierarchical estimation.

## Appendix E

### E.1. Parameter estimates obtained when applying CPT to the choices of the heuristics in different environments

See [Tables E1–E3](#).

## Appendix F

### F.1. Posterior distributions of the parameter estimation

See [Fig. F1](#).

**Table E1**

CPT parameter estimates for the choices of each of the five heuristics in problems with two-, three-, and five-outcome gambles, respectively, with similar EVs. Also shown are the deviance information criterion (DIC) of CPT and the results of posterior predictive checks (PP) indicating the proportion of choices for which CPT, based on the posterior distributions of the parameters, generated the same choice as the respective heuristic.

Heuristic	Parameters						DIC	PP
	$\gamma$	$\delta^+$	$\delta^-$	$\alpha$	$\lambda$	$\phi$		
Two-outcome gambles								
Minimax	0.11	0.04	4.77	0.93	2.48	0.37	65.01	0.91
Maximax	0.10	4.76	0.04	0.90	3.19	0.99	38.68	0.95
Priority	0.17	0.05	0.62	0.82	1.37	0.35	160.58	0.73
Least likely	0.90	2.81	2.89	0.04	3.53	3.95	196.51	0.63
Most likely	0.96	1.72	1.23	0.95	1.29	0.08	226.51	0.59
Three-outcome gambles								
Minimax	0.20	0.16	4.12	0.95	1.59	0.15	140.93	0.81
Maximax	0.13	4.68	0.04	0.85	1.24	0.67	89.43	0.87
Priority	0.14	0.18	0.30	0.79	1.90	0.29	181.2	0.70
Least likely	0.92	0.69	3.67	0.22	1.97	1.89	228.27	0.57
Most likely	0.95	0.90	1.81	0.93	1.56	0.09	207.86	0.63
Five-outcome gambles								
Minimax	0.02	0.04	4.47	0.81	1.79	1.11	49.66	0.93
Maximax	0.05	4.59	0.04	0.89	3.14	0.73	49.8	0.93
Priority	0.06	0.30	1.17	0.52	3.16	1.81	90.42	0.86
Least likely	0.87	1.34	3.58	0.42	2.63	0.48	201.59	0.64
Most likely	0.81	1.03	0.66	0.70	1.74	0.50	78.37	0.90

Note. The DIC assuming random choice was 249.53.

**Table E2**

CPT parameter estimates for the choices of the priority heuristic in problems with two-outcome gambles, separately for problems where the choice was based on the top-ranked reason (probability of minimum gain/loss) in 25%, 50%, or 75% of cases, respectively (the choice was otherwise based on the second reason: the probability of the minimum gain/loss). Also shown are the deviance information criterion (DIC) of CPT and the results of posterior predictive checks (PP) indicating the proportion of choices in which CPT, based on the posterior distributions of the parameters, generated the same choice as the heuristic.

% decisions based on first reason	Parameters						DIC	PP
	$\gamma$	$\delta^+$	$\delta^-$	$\alpha$	$\lambda$	$\phi$		
75%	0.30	0.04	0.19	0.91	2.28	0.50	97.96	0.85
50%	0.49	0.06	0.64	0.84	1.72	0.28	167.43	0.71
25%	0.94	0.20	0.56	0.85	1.70	0.35	125.18	0.80

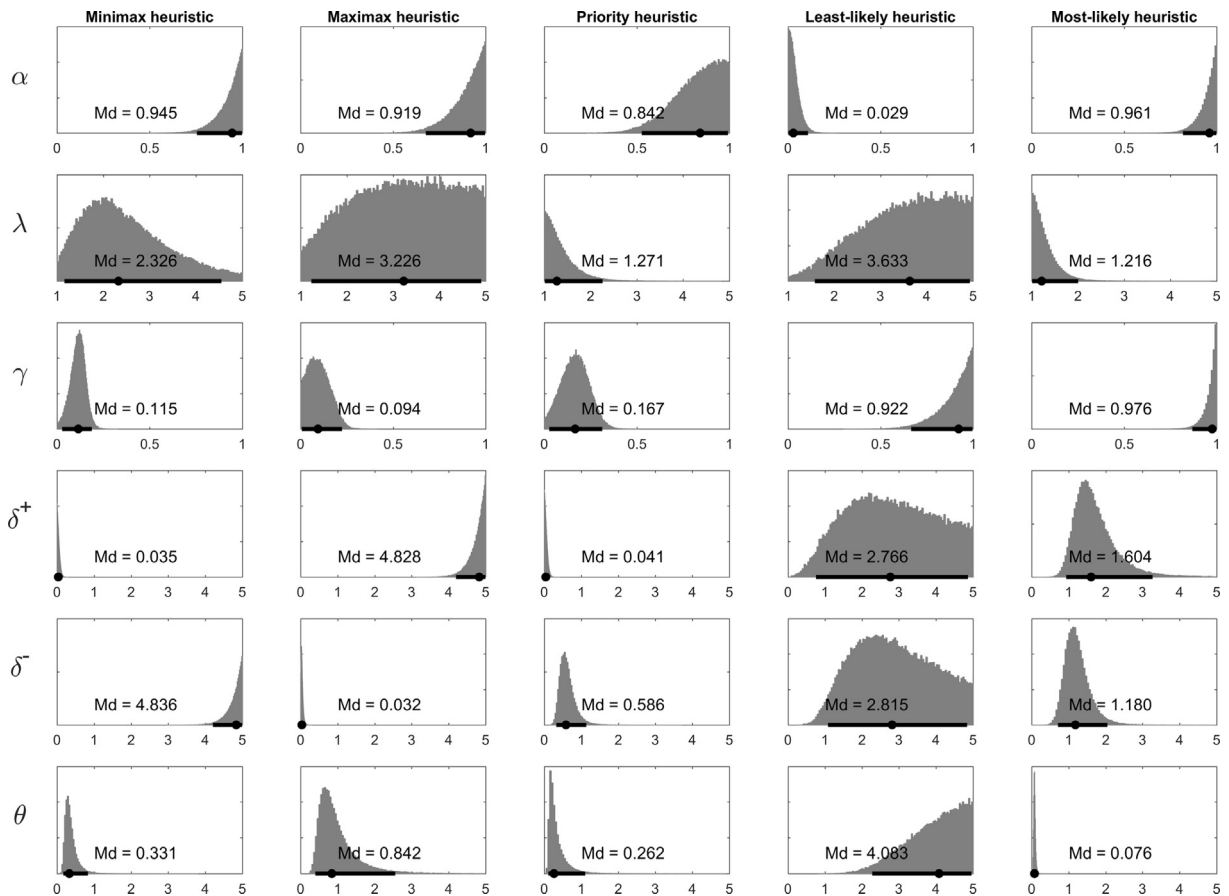
Note. The DIC assuming random choice was 249.53.

**Table E3**

CPT parameter estimates for the choices of each of the five heuristics in problems with two-, three-, and five-outcome gambles, respectively, with a negative relationship between outcomes and probabilities. Also shown are the deviance information criterion (DIC) of CPT and the results of posterior predictive checks (PP), indicating the proportion of choices for which CPT, based on the posterior distributions of the parameters, generated the same choice as the respective heuristic.

Heuristic	Parameters						DIC	PP
	$\gamma$	$\delta^+$	$\delta^-$	$\alpha$	$\lambda$	$\phi$		
Two-outcome gambles								
Minimax	0.02	0.07	4.55	0.77	1.94	1.08	32.60	0.96
Maximax	0.04	4.69	0.02	0.77	2.77	0.93	51.37	0.93
Priority	0.23	0.15	2.63	0.67	3.37	2.11	18.62	0.98
Least likely	0.90	1.94	1.52	0.01	3.77	4.21	182.61	0.65
Most likely	0.87	0.27	0.08	0.48	2.72	2.87	37.38	0.95
Three-outcome gambles								
Minimax	0.09	0.03	4.72	0.83	1.29	1.12	58.66	0.92
Maximax	0.04	4.68	0.02	0.86	2.89	1.74	32.92	0.96
Priority	0.22	0.20	0.63	0.78	1.68	0.63	126.08	0.79
Least likely	0.87	1.73	4.03	0.05	3.12	2.38	221.60	0.58
Most likely	0.56	0.03	0.01	0.69	2.73	3.19	95.54	0.83
Five-outcome gambles								
Minimax	0.10	0.07	4.12	0.72	1.14	1.16	117.38	0.83
Maximax	0.06	4.57	0.01	0.87	3.51	1.21	52.96	0.93
Priority	0.25	0.23	0.56	0.78	1.25	0.67	163.75	0.72
Least likely	0.70	2.66	3.27	0.08	2.35	1.45	249.45	0.51
Most likely	0.77	0.07	0.02	0.72	3.27	3.92	73.55	0.87

Note. The DIC assuming random choice was 249.53.



**Fig. F1.** Posterior distribution of the CPT parameter estimates for the choices of each of the five heuristics in problems with two-outcome gambles with similar EVs. The dots indicate the median of the distribution, the horizontal lines indicate the 95% highest density interval.

## Appendix G. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.cogpsych.2017.01.001>.

## References

- Abdellaoui, M., L'Haridon, O., & Zank, H. (2010). Separating curvature and elevation: A parametric probability weighting function. *Journal of Risk and Uncertainty*, 41, 39–65. <http://dx.doi.org/10.1007/s11166-010-9097-6>.
- Batchelder, W. H. (1998). Multinomial processing tree models and psychological assessment. *Psychological Assessment*, 10, 331–344.
- Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, 22, 23–36. <http://dx.doi.org/10.2307/1909829> (Original work published 1738).
- Bhatia, S. (2014). Sequential sampling and paradoxes of risky choice. *Psychonomic Bulletin and Review*, 21, 1095–1111. <http://dx.doi.org/10.3758/s13423-014-0650-1>.
- Bhatia, S. (2015). Choice rules and accumulator networks. *Decision*. <http://dx.doi.org/10.1037/dec0000038>. Advance online publication.
- Birnbaum, M. H. (2004). Tests of rank-dependent utility and cumulative prospect theory in gambles represented by natural frequencies: Effects of format, event framing, and branch splitting. *Organizational Behavior and Human Decision Processes*, 95, 40–65. <http://dx.doi.org/10.1016/j.obhdp.2004.05.004>.
- Birnbaum, M. H. (2008). New paradoxes of risky decision making. *Psychological Review*, 115, 463–501. <http://dx.doi.org/10.1037/0033-295X.115.2.463>.
- Birnbaum, M. H., & Chavez, A. (1997). Tests of theories of decision making: Violations of branch independence and distribution independence. *Organizational Behavior and Human Decision Processes*, 71, 161–194. <http://dx.doi.org/10.1006/obhd.1997.2721>.
- Birnbaum, M. H., & LaCroix, A. R. (2008). Dimension integration: Testing models without trade-offs. *Organizational Behavior and Human Decision Processes*, 105, 122–133. <http://dx.doi.org/10.1016/j.obhdp.2007.07.002>.
- Booij, A. S., & van de Kuilen, G. (2009). A parameter-free analysis of the utility of money for the general population under prospect theory. *Journal of Economic Psychology*, 30, 651–666. <http://dx.doi.org/10.1016/j.joep.2009.05.004>.
- Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, 113, 409–432. <http://dx.doi.org/10.1037/0033-295X.113.2.409>.
- Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, 100, 432–459. <http://dx.doi.org/10.1037/0033-295X.100.3.432>.
- Cavagnaro, D. R., Gonzalez, R., Myung, I. J., & Pitt, M. A. (2013). Optimal decision stimuli for risky choice experiments: An adaptive approach. *Management Science*, 59, 358–375. <http://dx.doi.org/10.1287/mnsc.1120.1558>.

- Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making*, 4, 20–33.
- Cooksey, R. W. (1996). *Judgment analysis: Theory, methods, and applications*. San Diego, CA: Academic Press.
- Coombs, C. H., Dawes, R. M., & Tversky, A. (1970). *Mathematical psychology: An elementary introduction*. Englewood Cliffs, NJ: Prentice-Hall.
- Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review*, 106, 180–209. <http://dx.doi.org/10.1037/0033-295X.106.1.180>.
- Edwards, W. (1953). Probability-preferences in gambling. *American Journal of Psychology*, 66, 349–364.
- Einhorn, H. J. (1970). The use of nonlinear noncompensatory models in decision making. *Psychological Bulletin*, 73, 221–230.
- Erev, I., Roth, A. E., Slonim, R. L., & Barron, G. (2002). Combining a theoretical prediction with experimental evidence Retrieved from: <<http://ssrn.com/abstract=1111712>>.
- Fehr-Duda, H., De Gennaro, M., & Schubert, R. (2006). Gender, financial risk, and probability weights. *Theory and Decision*, 60, 283–313. <http://dx.doi.org/10.1007/s11238-005-4590-0>.
- Fiedler, K. (2010). How to study cognitive decision algorithms: The case of the priority heuristic. *Judgment and Decision Making*, 5, 21–32.
- Fishburn, P. C. (1974). Lexicographic orders, utilities and decision rules: A survey. *Management Science*, 20, 1442–1471. <http://dx.doi.org/10.1287/mnsc.20.11.1442>.
- Fox, C. R., & Poldrack, R. A. (2014). Prospect theory and the brain. In E. Fehr & P. W. Glimcher (Eds.), *Neuroeconomics* (2nd ed., pp. 533–567). San Diego: Academic Press.
- Friedman, M. (1953). *Essays in positive economics*. Chicago, IL: University of Chicago Press.
- Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). *Heuristics: The foundations of adaptive behavior*. Oxford, United Kingdom: Oxford University Press.
- Gigerenzer, G., & Kurz, E. (2001). Vicarious functioning reconsidered: A fast and frugal lens model. In K. R. Hammond & T. R. Stewart (Eds.), *The essential brunswick: Beginnings, explications, applications* (pp. 342–347). New York, NY: Oxford University Press.
- Gilovich, T., & Griffin, D. W. (2002). Introduction – Heuristics and biases: Then and now. In T. Gilovich, D. W. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 1–18). Cambridge, United Kingdom: Cambridge University Press.
- Glöckner, A., & Betsch, T. (2008). Do people make decisions under risk based on ignorance? An empirical test of the priority heuristic against cumulative prospect theory. *Organizational Behavior and Human Decision Processes*, 107, 75–95. <http://dx.doi.org/10.1016/j.obhdp.2008.02.003>.
- Glöckner, A., & Herbold, A.-K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making*, 24, 71–98.
- Glöckner, A., & Pachur, T. (2012). Cognitive models of risky choice: Parameter stability and predictive accuracy of prospect theory. *Cognition*, 123, 21–32. <http://dx.doi.org/10.1016/j.cognition.2011.12.002>.
- Goldstein, W. M., & Einhorn, H. J. (1987). Expression theory and the preference reversal phenomena. *Psychological Review*, 94, 236–254. <http://dx.doi.org/10.1037/0033-295X.94.2.236>.
- Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, 38, 129–166. <http://dx.doi.org/10.1006/cogp.1998.0710>.
- Gul, F., & Pesendorfer, W. (2005). *The case for mindless economics*. NJ: Princeton University. Retrieved from: -pesendor/mindless.pdf" xlink:type="simple" id="ir015"><<http://www.princeton.edu/~pesendor/mindless.pdf>>.
- Hammond, K. R., Hursch, C. J., & Todd, F. S. (1964). Analyzing the components of clinical inference. *Psychological Review*, 71, 438–456. <http://dx.doi.org/10.1037/h0040736>.
- Harbaugh, W. T., Krause, K., & Vesterlund, L. (2002). Risk attitudes of children and adults: Choices over small and large probability gains and losses. *Experimental Economics*, 5, 53–84. <http://dx.doi.org/10.1023/A:1016316725855>.
- Hoffman, P. J. (1960). Paramorphic representation of clinical judgment. *Psychological Bulletin*, 47, 116–131.
- Jarvstad, A., Hahn, U., Rushton, S. K., & Warren, P. A. (2013). Perceptuo-motor, cognitive, and description-based decision-making seem equally good. *Proceedings of the National Academy of Sciences*, 110, 16271–16276. <http://dx.doi.org/10.1073/pnas.1300239110>.
- Johnson, E. J., & Meyer, R. J. (1984). Compensatory choice models of noncompensatory processes: The effect of varying context. *Journal of Consumer Research*, 11, 528–541. <http://dx.doi.org/10.1086/208989>.
- Johnson, E. J., & Ratcliff, R. (2014). Computational and process models of decision-making in psychology and behavioral economics. In P. W. Glimcher & E. Fehr (Eds.), *Neuroeconomics: Decision making and the brain* (2nd ed., New York, NY: Academic Press.
- Johnson, E. J., Schulte-Mecklenbeck, M., & Willemsen, M. (2008). Process models deserve process data: Comment on Brandstätter, Gigerenzer, & Hertwig (2006). *Psychological Review*, 115, 263–272. <http://dx.doi.org/10.1037/0033-295X.115.1.263>.
- Juslin, P., Olsson, H., & Olsson, A.-C. (2003). Exemplar effects in categorization and multiple-cue judgment. *Journal of Experimental Psychology: General*, 132, 133–156. <http://dx.doi.org/10.1037/0096-3445.132.1.133>.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291. <http://dx.doi.org/10.2307/1914185>.
- Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values, and frames*. Cambridge, United Kingdom: Cambridge University Press.
- Katsikopoulos, K. V., & Gigerenzer, G. (2008). One-reason decision-making: Modeling violations of expected utility theory. *Journal of Risk and Uncertainty*, 37, 35–56.
- Lattimore, P. K., Baker, J. R., & Witte, A. D. (1992). The influence of probability on risky choice: A parametric examination. *Journal of Economic Behavior and Organization*, 17, 337–400. [http://dx.doi.org/10.1016/S0167-2681\(95\)90015-2](http://dx.doi.org/10.1016/S0167-2681(95)90015-2).
- Lee, M. D. (2016). Bayesian outcome-based strategy classification. *Behavior Research Methods*, 48, 29–41. <http://dx.doi.org/10.3758/s13428-014-0557-9>.
- Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge, United Kingdom: Cambridge University Press.
- Lejarraga, T., Pachur, T., Frey, R., & Hertwig, R. (2016). Decisions from experience: From monetary to medical gambles. *Journal of Behavioral Decision Making*, 29, 67–77. <http://dx.doi.org/10.1002/bdm.1877>.
- Leland, J. W. (1998). Similarity judgments in choice under uncertainty: A reinterpretation of the predictions of regret theory. *Management Science*, 44, 659–672. <http://dx.doi.org/10.1287/mnsc.44.5.659>.
- Lichtenstein, S. (1965). Bases for preferences among three-outcome bets. *Journal of Experimental Psychology*, 69, 162–169. <http://dx.doi.org/10.1037/h0021570>.
- Loomes, G. (2010). Modeling choice and valuation in decision experiments. *Psychological Review*, 117, 902–924. <http://dx.doi.org/10.1037/a0019807>.
- Lopes, L. L. (1987). Between hope and fear: The psychology of risk. *Advances in Experimental Psychology*, 20, 255–295. [http://dx.doi.org/10.1016/S0065-2601\(08\)60416-5](http://dx.doi.org/10.1016/S0065-2601(08)60416-5).
- Lopes, L. L. (1995). Algebra and process in the modeling of risky choice. *Psychology of Learning and Motivation*, 32, 177–220. [http://dx.doi.org/10.1016/S0079-7421\(08\)60310-2](http://dx.doi.org/10.1016/S0079-7421(08)60310-2).
- Lopes, L. L., & Oden, G. C. (1999). The role of aspiration level in risky choice: A comparison of cumulative prospect theory and SP/A theory. *Journal of Mathematical Psychology*, 43, 286–313. <http://dx.doi.org/10.1006/jmps.1999.12s59>.
- Lopes, L. L. (1990). Remodeling risk aversion. In G. M. von Furstenberg (Ed.), *Acting under uncertainty: Multidisciplinary conceptions* (pp. 267–299). Boston, MA: Kluwer.
- Luan, S., Schooler, L. J., & Gigerenzer, G. (2011). A signal detection analysis of fast-and-frugal trees. *Psychological Review*, 118, 316–338. <http://dx.doi.org/10.1037/a0022684>.
- Luce, R. D. (1956). Semiordeers and a theory of utility discrimination. *Econometrica*, 24, 178–191. <http://dx.doi.org/10.2307/1905751>.
- Mann, L., & Ball, C. (1994). The relationship between search strategy and risky choice. *Australian Journal of Psychology*, 46, 131–136. <http://dx.doi.org/10.1080/00049539408259487>.



- Massaro, D. W., & Friedman, D. (1990). Models of integration given multiple sources of information. *Psychological Review*, 97, 225–252. <http://dx.doi.org/10.1037/0033-295X.97.2.225>.
- Nilsson, H., Rieskamp, J., & Wagenmakers, E. (2011). Hierarchical Bayesian parameter estimation for cumulative prospect theory. *Journal of Mathematical Psychology*, 55, 84–93. <http://dx.doi.org/10.1016/j.jmp.2010.08.006>.
- Pachur, T., Schulte-Mecklenbeck, M., Murphy, R. O., & Hertwig, R. (2016). Prospect theory reflects selective allocation of attention. Submitted for publication.
- Pachur, T., & Galesic, M. (2013). Strategy selection in risky choice: The impact of numeracy, affect, and cross-cultural differences. *Journal of Behavioral Decision Making*, 26, 260–271. <http://dx.doi.org/10.1002/bdm.175>.
- Pachur, T., Hanoch, Y., & Gummerum, M. (2010). Prospects behind bars: Analyzing decisions under risk in a prison population. *Psychonomic Bulletin and Review*, 17, 630–636. <http://dx.doi.org/10.3758/PBR.17.5.630>.
- Pachur, T., Hertwig, R., Gigerenzer, G., & Brandstätter, E. (2013). Testing process predictions of models of risky choice: A quantitative model comparison approach. *Frontiers in Psychology*, 4, 646. <http://dx.doi.org/10.3389/fpsyg.2013.00646>.
- Pachur, T., Hertwig, R., & Wolke, R. (2014). The affect gap in risky choice: Affect-rich outcomes attenuate attention to probability information. *Decision*, 1, 64–78. <http://dx.doi.org/10.1037/dec0000006>.
- Pachur, T., & Kellen, D. (2013). Modeling gain-loss asymmetries in risky choice: The critical role of probability weighting. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual conference of the cognitive science society* (pp. 3205–3210). Austin, TX: Cognitive Science Society.
- Pachur, T., Mata, R., & Hertwig, R. (in press). Who dares, who errs? Disentangling cognitive and motivational roots of age differences in decisions under risk. *Psychological Science*. <http://dx.doi.org/10.1177/0956797616687729>, in press.
- Payne, J. W. (1973). Alternative approaches to decision making under risk: Moments versus risk dimensions. *Psychological Bulletin*, 80, 439–453. <http://dx.doi.org/10.1037/h0035260>.
- Payne, J. W. (2005). It's whether you win or lose: The importance of the overall probabilities of winning or losing in risky choice. *Journal of Risk and Uncertainty*, 30, 5–19. <http://dx.doi.org/10.1007/s11166-005-5831-x>.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge, United Kingdom: Cambridge University Press.
- Payne, J. W., & Braunstein, M. L. (1978). Risky choice: An examination of information acquisition behavior. *Memory and Cognition*, 6, 554–561. <http://dx.doi.org/10.3758/BF03198244>.
- Pitt, M. A., & Myung, I. J. (2002). When a good fit can be bad. *Trends in Cognitive Sciences*, 6, 421–425. [http://dx.doi.org/10.1016/S1364-6613\(02\)01964-2](http://dx.doi.org/10.1016/S1364-6613(02)01964-2).
- Pleskac, T. J. (2007). A signal detection analysis of the recognition heuristic. *Psychonomic Bulletin and Review*, 14, 379–391. <http://dx.doi.org/10.3758/BF03194081>.
- Pleskac, T. J., Diederich, A., & Wallsten, T. S. (2015). Models of decision making under risk and uncertainty. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.), *The oxford handbook of computational and mathematical psychology* (pp. 209–231). New York, NY: Oxford University Press.
- Pleskac, T. J., & Hertwig, R. (2014). Ecologically rational choice and the structure of the environment. *Journal of Experimental Psychology: General*, 143, 2000–2019. <http://dx.doi.org/10.1037/xge0000013>.
- Prelec, D. (1998). The probability weighting function. *Econometrica*, 66, 497–527. <http://dx.doi.org/10.2307/2998573>.
- Rao, L. L., Li, S., Jiang, T., & Zhou, Y. (2012). Is payoff necessarily weighted by probability when making a risky choice? Evidence from functional connectivity analysis. *PLoS ONE*, 7, e41048. <http://dx.doi.org/10.1371/journal.pone.0041048>.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108. <http://dx.doi.org/10.1037/0033-295X.85.2.59>.
- Riefer, D. M., Knapp, B. R., Batchelder, W. H., Bamber, D., & Manifold, V. (2002). Cognitive psychometrics: Assessing storage and retrieval deficits in special populations with multinomial processing tree models. *Psychological Assessment*, 14, 184–201. <http://dx.doi.org/10.1037/1040-3590.14.2.184>.
- Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1446–1465. <http://dx.doi.org/10.1037/a0013646>.
- Rosen, L. D., & Rosenkoetter, P. (1976). An eye fixation analysis of choice of judgment with multiattribute stimuli. *Memory and Cognition*, 4, 747–752. <http://dx.doi.org/10.3758/BF03213243>.
- Rottenstreich, Y., & Hsee, C. K. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science*, 12, 185–190.
- Rubinstein, A. (1988). Similarity and decision-making under risk (Is there a utility theory resolution to the Allais-paradox?). *Journal of Economic Theory*, 46, 145–153. [http://dx.doi.org/10.1016/0022-0531\(88\)90154-8](http://dx.doi.org/10.1016/0022-0531(88)90154-8).
- Russo, J. E., & Doshier, B. A. (1983). Strategies for multiattribute binary choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 676–696. <http://dx.doi.org/10.1037/0278-7393.9.4.676>.
- Savage, L. J. (1954/1972). *The foundations of statistics*. New York, NY: Dover Publications.
- Scheibehenne, B., & Pachur, T. (2015). Using Bayesian hierarchical parameter estimation to assess the generalizability of cognitive models of choice. *Psychonomic Bulletin and Review*, 22, 391–407.
- Scheibehenne, B., Rieskamp, J., & González-Vallejo, C. (2009). Models of preferential choice: Comparing the decision field theory with the proportional difference model. *Cognitive Science*, 33, 911–939. <http://dx.doi.org/10.1111/j.1551-6709.2009.01034.x>.
- Schmidt, U., & Zank, H. (2005). What is loss aversion? *Journal of Risk and Uncertainty*, 30, 157–167. <http://dx.doi.org/10.1007/s11166-005-6564-6>.
- Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, 112, 610–628. <http://dx.doi.org/10.1037/0033-295X.112.3.610>.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461–464. <http://dx.doi.org/10.1214/aos/1176344136>.
- Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin*, 134, 207–222.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69, 99–118. <http://dx.doi.org/10.2307/1884852>.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63, 129–138.
- Slovic, P., & Lichtenstein, S. (1968). Relative importance of probabilities and payoffs in risk taking. *Journal of Experimental Psychology*, 78, 1–18. <http://dx.doi.org/10.1037/h0026468>.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B*, 64, 583–639. <http://dx.doi.org/10.1111/1467-9868.00353>.
- Stewart, N. (2011). Information integration in risky choice: Identification and stability. *Frontiers in Psychology*, 2, 301. <http://dx.doi.org/10.3389/fpsyg.2011.00301>.
- Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, 53, 1–26. <http://dx.doi.org/10.1016/j.cogpsych.2005.10.003>.
- Stewart, N., Hermens, F., & Matthews, W. J. (2016). Eye movements in risky choice. *Journal of Behavioral Decision Making*, 29, 116–136.
- Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty*, 32, 101–130. <http://dx.doi.org/10.1007/s11166-006-8289-6>.
- Su, Y., Rao, L., Sun, H., Du, X., Li, X., & Li, S. (2013). Is making a risky choice based on a weighting and adding process? An eye-tracking investigation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 39, 1765–1780. <http://dx.doi.org/10.1037/a0032861>.
- Suter, R. S., Pachur, T., & Hertwig, R. (2016). How affect shapes risky choice: Distorted probability weighting versus probability neglect. *Journal of Behavioral Decision Making*, 29, 437–449. <http://dx.doi.org/10.1002/bdm.1888>.
- Suter, R. S., Pachur, T., Hertwig, R., Endestad, T., & Biele, G. (2015). The neural basis of risky choice with affective outcomes. *PLoS ONE*, 10(4), e0122475. <http://dx.doi.org/10.1371/journal.pone.0122475>.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.
- Svenson, O. (1979). Process descriptions of decision making. *Organizational Behaviour and Human Performance*, 23, 86–112.
- Thorngate, W. (1980). Efficient decision heuristics. *Behavioral Science*, 25, 219–225. <http://dx.doi.org/10.1002/bs.3830250306>.
- Todd, P. M., & Gigerenzer, G. the ABC Research Group. (2012). *Ecological rationality: Intelligence in the world*. New York, NY: Oxford University Press.
- Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, 315, 515–518.

- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76, 31–48.
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, 79, 281–299. <http://dx.doi.org/10.1037/h0032955>.
- Tversky, A., & Kahneman, D. (1974). Heuristics and biases: Judgement under uncertainty. *Science*, 185, 1124–1130. <http://dx.doi.org/10.1126/science.185.4157.1124>.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297–323. <http://dx.doi.org/10.1007/BF00122574>.
- Venkatraman, V., Payne, J. W., Bettman, J. R., Luce, M. F., & Huettel, S. A. (2009). Separate neural mechanisms underlie choices and strategic preferences in risky decision making. *Neuron*, 62, 593–602. <http://dx.doi.org/10.1016/j.neuron.2009.04.007>.
- Venkatraman, V., Payne, J. W., & Huettel, S. A. (2014). An overall probability of winning heuristic for complex risky decisions: Choice and eye fixation evidence. *Organizational Behavior and Human Decision Processes*, 125, 73–87. <http://dx.doi.org/10.1016/j.obhdp.2014.06.003>.
- von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.
- Wakker, P. P. (2001). Testing and characterizing properties of nonadditive measures through violations of the sure-thing principle. *Econometrica*, 69, 1039–1059. <http://dx.doi.org/10.1111/1468-0262.00229>.
- Wakker, P. P. (2010). *Prospect theory for risk and ambiguity*. Cambridge, United Kingdom: Cambridge University Press.
- Willemsen, M. C., Böckenholt, U., & Johnson, E. J. (2011). Choice by value encoding and value construction: Processes of loss aversion. *Journal of Experimental Psychology: General*, 140, 303–324. <http://dx.doi.org/10.1037/a0023493>.
- Wu, G., & Markle, A. (2008). An empirical test of gain-loss separability in prospect theory. *Management Science*, 54, 1322–1335. <http://dx.doi.org/10.1287/mnsc.1070.0846>.
- Wu, G., Zhang, J., & Gonzalez, R. (2004). Decision under risk. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 399–423). Oxford, United Kingdom: Blackwell.
- Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, 55, 95–115. <http://dx.doi.org/10.2307/1911158>.
- Yechiam, E., & Busemeyer, J. (2005). Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic Bulletin and Review*, 12, 387–402. <http://dx.doi.org/10.3758/BF03193783>.
- Young, D. L., Goodie, A. S., Hall, D. B., & Wu, E. (2012). Decision making under time pressure, modeled in a prospect theory framework. *Organizational Behavior and Human Decision Processes*, 118, 179–188. <http://dx.doi.org/10.1016/j.obhdp.2012.03.005>.
- Zeigenfuse, M. D., & Lee, M. D. (2010). Finding the features that represent stimuli. *Acta Psychologica*, 133, 283–295. <http://dx.doi.org/10.1016/j.actpsy.2009.07.014>.