A Prospect Theory Account of the “Hot” Columbia Card Task

Andreas Pedroni[1,2*], Jörg Rieskamp[1], Thorsten Pachur[3], Renato Frey[3,4], Jonathan E. Westfall[5] & Bernd Figner[6*]

[1] Center for Economic Psychology, Department of Psychology, University of Basel, Switzerland

[2] Methods of Plasticity Research, Institute of Psychology, University of Zurich, Switzerland

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

[4] Center for Cognitive and Decision Sciences, University of Basel, Switzerland

[5]Delta State University, Cleveland, United States

[6] Behavioural Science Institute and Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands

* Corresponding authors

Abstract

The investigation of decisions under risk has mainly followed one of two approaches. One relies on observing choices between lotteries in which economic primitives (outcome magnitudes, probabilities, and domains (i.e., gains and losses)) are varied systematically, and this information is described to participants. The systematic variation of the economic primitives allows to formally describe behavior with expectation-based models such as expected utility theory or cumulative prospect theory (CPT), arguably the most prominent descriptive theories of risky choice. One drawback, however, is that lottery tasks can seem artificial, likely reducing the external or ecological validity. A second more naturalistic approach employs dynamic paradigms that mimic features of real-life risky situations and are assumed to have higher ecological validity. Because key information are often not provided to the decision maker, it is impossible to apply the same models as in the first approach. The goal of the present work is to integrate both approaches, by developing models for the "hot" Columbia Card Task (CCT), a task that combines a dynamic decision situation with systematic trial-to-trial variation in economic primitives. In a model comparison on the basis of the data of 191 participants, we identified a best-performing model that describes behavior as a function of CPT's main components, outcome sensitivity, probability weighting, and loss aversion. Our work therefore provides a framework that allows the description of risk-taking behavior in a naturalistic dynamic task based on key psychological constructs (e.g., loss aversion, probability weighting) that are rooted in the factorial variation of economic primitives.

Keywords: risk taking, cognitive modeling, Columbia Card Task, cumulative prospect theory, reference point

Introduction

Who takes risks, and why, are central questions in the behavioral sciences. Research investigating these issues often follows one of two approaches, each with its own strengths and limitations. The first, approach uses highly abstract decision situations with systematically designed stimuli to study people's risk taking, whereas the second, approach uses, dynamically evolving and thus more realistic and ecologically valid tasks. Without making strong claims about connotation, we refer to the first approach as the "systematic-variation" approach to study risk taking and the latter as the "naturalistic" approach. With the present work, we aim to combine the strengths of the systematic-variation and naturalistic approaches, while trying to minimize the limitations of each.

The systematic-variation approach to study risk taking has emerged mainly from research in economics and the decision sciences: Here, researchers aimed to develop models to quantitatively describe individuals' risky decisions as a function of their risk preferences. Risk preferences are expressed relative to risk-neutral behavior and are based on the subjective representation and integration of outcome magnitudes, each weighted by some function of the outcome probabilities (yielding the expectation of an option). Therefore models describing this representation and integration are based on the notion of the maximization of expectation. The most prominent models following this notion are expected utility theory (EUT, Von Neumann & Morgenstern, 1944), prospect theory (Kahneman & Tversky, 1979), and cumulative prospect theory (CPT; Tversky & Kahneman, 1992). To illustrate the idea of expectation-maximization based models we consider CPT, arguably the most influential theoretical framework for describing decisions under risk. According to CPT, people's risky decisions can be captured assuming three core psychological components: sensitivity to nominal differences in outcomes (e.g., how much does the subjective utility of outcomes change if the stakes are doubled?), sensitivity to differences in probabilities (e.g., how much does the subjective decision weights change if the chance of winning is doubled?), and differential attraction to gains versus their aversion to losses (e.g., does losing hurt as much as winning makes you happy?). Thus, once a person's outcome and probability sensitivity and their relative aversion to losses (compared to gains) has been characterized, one can predict how this person will decide in any well-defined situation (i.e., with precisely described outcomes and probabilities).

Quantifying a person's risk preference by means of these three components requires multiple observations of choices between prospects that vary in their outcome probabilities, magnitudes and outcome domains (i.e., gains, losses, mixed lotteries). In pursuit of this goal and in view of these requirements, researchers have used lottery paradigms in which individuals typically make many dozens of choices between monetary lotteries that incorporate such variation in a systematic way. To ensure that people do not make decisions for combinations of lotteries and avoid changes of reference points, the outcomes of these lotteries are often not revealed to the participants and typically only one lottery is randomly drawn and played out at the end of an experiment to assure incentive compatibility. In sum, these paradigms are designed with the goal of ruling out nuisance factors and to precisely assess the relevant key variables in the models that determine individuals' risky decisions.

Arguably owing to the experimental rigor of the systematic-variation approach, it can and has been criticized as lacking important aspects that characterize everyday situations: Risky situations in everyday life are often more complex than choices between well-defined monetary lotteries. They include a variety of different psychological factors, including dynamic changes of the situation, receiving immediate feedback, and thus being provided with the opportunity to learn (likely via more cognitive and via more affective processes) from the outcome of a choice and potentially adjust behavior before making the next choice. From this perspective, the lottery paradigm may seem artificial and abstract, as it does not incorporate relevant processes such as dynamic changes, feedback, and affective and motivational processes. However, these processes have been shown to be important factors in risky decisions (Bell, 1982; Figner, Mackinlay, Wilkening, & Weber, 2009; Frey, Rieskamp, & Hertwig, 2015; Loewenstein, Weber, Hsee, & Welch, 2001; Loomes & Sugden, 1982; Mellers, Schwartz, Ho, & Ritov, 1997; Pachur, Hertwig, & Wolkewitz, 2014). And thus—at least to the extent that the goal is to capture and mimic real-life risky choice situations that *do* involve these processes—these typical lottery methods likely may have limited ecological validity (Schonberg, Fox, & Poldrack, 2011).

The focus of the naturalistic approach, in contrast, is to mimic real-life risky decision making by incorporating into the task paradigms key aspects of such situations. Accordingly, these task paradigms often use more engaging problems in the form of playful, game-like tasks. Often, the decision situation contains a dynamic element and participants receive immediate feedback and thus experience the outcomes of their decisions. All these factors are assumed to contribute to considerable affective engagement (Schonberg et al., 2011). Slovic (1966) was the first to use a game of this sort (often referred to as the Devil’s Task) to investigate children’s risk preferences in an experimental study: Children were asked to pull as many of ten indistinguishable switches as they wished. Nine switches were “safe,” each leading to a gain in the form of sweets. One was a “disaster” switch, leading to a complete loss of all sweets gained so far and ending the experiment. The number of pulled switches was used as a measure of the child's risk preference. Thus, the probability of a loss increased dynamically with each pulled lever, and participants received immediate feedback. Other games have used the same basic idea of dynamic increases in riskiness over time, such as a task used to investigate risk taking in psychopaths (Siegel, 1978), the Balloon Analogue Risk Task (BART; Lejuez et al., 2002), its variant, the Angling Risk Task (Pleskac, 2008), the Sequential Investment Task (SIT; Frey, Rieskamp, & Hertwig, 2015), and the "hot" Columbia Card Task (CCT; Figner et al., 2009).

The shared structure of these games—in which participants make sequential risky choices where the riskiness increases incrementally with each choice and immediate feedback is provided following each choice—is common to many real-world risk-taking situations (Goldberg & Fischhoff, 2000; Leigh, 1999; Moore & Gullone, 1996; Weber & Johnson, 2009) and matches both the economic and lay definition of risk (March & Shapira, 1987): Within a trial, each successive choice of the risky option (i.e., pulling another lever, inflating the balloon by another pump, turning over another card) instead of ending the current trial increases the outcome variability (i.e., risk in the economic sense) and increases the probability of exposure to a negative consequence (i.e., the lay definition of risk).

It has been suggested that these game-like tasks to evoke the affective components that also accompany naturalistic risk taking, such as the feeling of escalating tension and exhilaration when a participant pushes decisions to maximal gain against the probability of a loss (Schonberg et al., 2011). Possibly because they capture the emotional components of risk taking and appear to be more similar to naturalistic risk taking, they have been successful in differentiating healthy controls from different groups of high risk takers—for example substance abusers (Bishara et al., 2009; Bornovalova, Daughters, Hernandez, Richards, & Lejuez, 2005; Coffey, Schumacher, Baschnagel, Hawk, & Holloman, 2011; Crowley, Raymond, Mikulich-Gilbertson, Thompson, & Lejuez, 2006; Hunt, Hopko, Bare, Lejuez, & Robinson, 2005; Ledgerwood, Alessi, Phoenix, & Petry, 2009) or prisoners (Wichary, Pachur, & Li, 2015). Although we are not aware of a systematic review, relative to static lottery-type tasks without feedback, these dynamic task types seem to more often show significant correlations with self-reported “real-life” risk-taking behaviors (e.g., Aklin, Lejuez, Zvolensky, Kahler, & Gwadz, 2005; Bornovalova et al., 2009; Hunt et al., 2005; Lejuez et al., 2003; 2003; MacPherson, Magidson, Reynolds, Kahler, & Lejuez, 2010; Mishra, Lalumière, & Williams, 2010; Skeel, Pilarski, Pytlak, & Neudecker, 2008; Swogger, Walsh, Lejuez, & Kosson, 2010; for a similar argument, see Schonberg, Fox, & Poldrack 2011)).

Typical analyses using dynamic task paradigms involve creating an aggregate score that captures participants’ propensity to choose a risky option, such as the number of pulled levers in Slovic's task, the number of pumps in the BART, or the net score in the IGT (although this measure is difficult to interpret because the net score is computed on the basis of the expected value of the decks, but nevertheless typically interpreted in terms of risk taking). Because the main goal of applying these tasks is often to produce a score that maximally differentiates groups of risk takers or to predict "real-life" risk taking with high accuracy, the task paradigms have been optimized accordingly, although it has typically abstained from systematically varying factors such as outcome magnitudes, probabilities, and domains.

Because these factors are not systematically varied (as they are largely under the participant's control), the observed risk-taking scores are only based on decisions with the possible outcomes of winning against losing within a rather restricted range of possibilities. As a consequence, risk-taking scores obtained in these paradigms may have limited generalizability to other situations with other outcome domains, for example to pure loss or pure gain domains, or other probability ranges. Along with this, the parameters of expectation-based models such as CPT cannot be reliably estimated from the data, which is undesirable as the estimated parameters in CPT would—at least theoretically—hold for any kind of decision situation that can be described by outcome probabilities and outcome magnitudes.

For some dynamic tasks, such as the Iowa Gambling Task (IGT), the BART, or the SIT researchers have developed quantitative models that include specific mechanisms to match task characteristics. For instance, in the IGT, BART, and the SIT participants start with no knowledge about the probabilities and outcome magnitudes of the different choice options. Thus, over the course of the task, participants usually have to learn the probabilities and magnitudes based on experience and feedback.[1] Thus, the learning process becomes of major interest and the quantitative models developed for these types of tasks focus centrally on the learning processes. Because these additional processes, such as learning, are given special attention, other central aspects of risk-taking, such as loss aversion, captured by CPT are given less attention.

Bridging the Gap

In the present article we aim to combine the advantages of the systematic-variation approach of enabling precise quantitative measures of risk-preferences with the advantage of the second naturalistic approach of higher ecological validity by examining a dynamic naturalistic task with feedback. We argue, consistent with the view of Schonberg et al. (2011), that the "hot" Columbia Card Task (CCT; Figner & Voelki, 2004; Figner et al., 2009) represents a task that is suitable to bring both approaches together: The CCT is dynamic and at the same time systematically varies outcome magnitudes and probabilities and thus involves decisions that can be represented as explicitly defined lotteries. This should make it possible to describe behavior at the level of psychological mechanisms with expectation-based models such as CPT. We therefore chose the CCT as a promising candidate to apply expectation-based models to a dynamic task paradigm, thus hoping to combine the advantages of the two discussed approaches: On the one hand, we want to use models to capture and describe risk preferences based on psychological components (such as outcome sensitivity and loss aversion) and on the other hand we hope to combine this with greater real-life similarity and affective involvement due to the dynamic task structure and immediate feedback. To the best of our knowledge, a model for the CCT that allows to identify

[1] Because learning the relevant information takes participants many trials, systematic variation of key variables such as outcome magnitudes and probabilities is simply not feasible, as participants would probably not be able to learn it within a reasonable number of trials.

individual differences by various parameters has so far not been developed and tested.[2] Thus, the main objective of the present study was to establish a model that serves to quantify individual differences in risk taking in the CCT and that provides explanations for behavior via its psychological mechanisms.

The remainder of the paper is structured as follows: We will first describe the CCT and discuss how the sequential nature of this task parallels the structure of other dynamic tasks (such as Slovic's task or the BART), while at the same time providing fully described decision options as in the lottery paradigm (i.e., unlike in the BART, in the CCT participants are provided with the relevant information about gain amounts, loss amounts, and their probabilities and do not have to learn this information based on experience). We then outline a series of models to describe decision making in the CCT and explain our model comparison approach. Next, we describe the specifics of the experimental study in which 191 healthy participants made decisions in a modified version of the hot CCT. The participants also answered a series of self-report risk-taking questionnaires that served to test and compare the convergent validity of observed CCT behavior and estimated model parameters. Thereafter, we report the results of the study and the model comparison analyses.

The Columbia Card Task

Figner and Voelki (2004) introduced the CCT to investigate information integration in risky choice. Figner et al. (2009) used a "hot" affective and a "cold" deliberative CCT version to investigate the influence of affective and deliberative processes in risk taking. In the "hot" (sequential) CCT version, participants play multiple rounds of a computerized card game, in each of which they are shown 32 cards: At the beginning of each round, all cards are shown

[2] Van Duijvenvoorde and colleagues (2015) have applied a risk return model without adjustable parameters to decompose behavior in an fMRI-adapted version of the CCT. Consistent with their goal of identifying neuronal responses to changing objective risks (outcome variability) and changing objective returns (expected value), their computational model did not account for individual differences in subjective representations of those, except for modeling individual differences in sensitivity to risks and returns.

face down. Participants can turn over cards as long as gain cards are encountered. Each gain card adds a specified gain amount to the payoff of the current round, and the player can voluntarily stop the round and claim the obtained payoff (i.e., the round's current score). As soon as a loss card is encountered, the trial terminates; that is, no more cards can be turned over and a specified loss amount is subtracted from the round's current score. Crucially, the loss amount, gain amount, as well as the number of loss cards is transparently displayed.

The sequential nature of the hot CCT with increasingly likely negative consequences has been shown to evoke strong affective responses. In the hot CCT, whether a participant encounters a specific choice situation depends on the preceding decisions. For example, a player could face the decision of whether or not to turn over a 10th card in a game round with 1 loss card, a gain amount of 30 cents and a loss amount of 250 cents. In this example, the participant will only encounter that specific choice situation when deciding to turn over the previous 9 cards, and when all these 9 previous cards were gain cards. Accordingly, participants in the hot CCT receive immediate feedback on every single decision. Deciding to turn over another card after just having successfully turned over a gain card is assumed to reflect a highly arousing experience that involves anticipatory and anticipating emotions. Figner et al. (2009) consistently found higher emotional arousal in the sequential "hot" compared to the more deliberative "cold" CCT, in which participants indicate at the beginning of each round how many cards they wish to turn over and receive feedback only at the end of the game. The same pattern was observed for both self-reported emotional arousal and for emotional arousal operationalized via skin conductance responses (SCR), a widely used physiological measure of emotional arousal (Boucsein, 2012; Critchley, Elliott, Mathias, & Dolan, 2000; Figner & Murphy, 2011). Recent work further supports the role of emotional processes in risky choice in the CCT (e.g., Baumann & DeSteno, 2012; Panno, Lauriola, & Figner, 2013; Panno, Lauriola, & Pierro, 2015). In sum, the hot version of the CCT appears to evoke substantial affective involvement and has been shown to be sensitive in capturing individual differences such as age (e.g., Figner et al., 2009; Huang, Wood, Berger, & Hanoch, 2013; van Duijvenvoorde et al., 2016), personality (e.g., Buelow, 2015), and situational factors such as stress (e.g., Jamieson & Mendes, 2016).

Although the hot CCT is similar to other dynamic tasks that evoke affective involvement, it is distinct from these tasks by provided information about the potential outcomes and the probabilities with which these outcomes occur. Thus in contrast to the BART or IGT no probabilities need to be learned. Unlike the Devil's Task, the BART, and the Angling Risk Task, the CCT systematically varies probability, gain magnitudes, and loss magnitudes and unconfounds the probability and magnitude of potential losses. The decisions in the CCT essentially represent choices between lotteries: The "stopping" option represents a sure-payoff lottery and the "turning over another card" option represents a lottery with two possible outcomes. Thus, although the CCT is a dynamic task, it has the advantage of involving well-defined options, allowing the application of computational models, such as EUT and CPT.

Models for the Hot CCT

In each of the game rounds in the CCT, participants repeatedly choose whether to turn over a card or to stop and cash the accrued payoff. To make their choices, participants need to evaluate whether turning over a card leads to a higher expected benefit compared to stopping. Prospect theory assumes that the evaluation of the two options can differ substantially depending on what people perceive as the reference point that defines gains and losses, and depending on how the probabilities of the potential outcomes determine their expectations. We first discuss two variants of reference points and then outline models that account for different ways of constructing expectations. We describe the basic idea of the models and provide a formal description in Appendix A.

Reference point. In one model variant the reference point consists of the accumulated payoff in each round and is therefore updated after every decision. We refer to this type as the decision-updated (DU) reference point (see also van Duijvenvoorde et al., 2015, Wallsten et al., 2005). Under this model variant, the decision maker chooses between receiving a payoff of zero (i.e., stopping) and gaining or losing some payoff with some probability (i.e., turning over another card).

According to the second model variant, the reference point is updated after the completion of each game round. We refer to this type as the round-updated (RU) reference point. Correspondingly, the option to stop is perceived as receiving the payoff accumulated in the current round. Likewise, the option to turn over another card is perceived as receiving with some probability the accumulated payoff plus the potential gain and with some probability the accumulated payoff minus the potential loss. The decision to turn over a card is thus a function of the difference between the money accrued so far (i.e., the sure outcome of stopping) and the expected outcome of the lottery of turning over one more card plus the accrued money so far.

Therefore, in the first model variant (DU-reference point) participants ignore the history of outcomes in the round and focus entirely on the outcomes and probabilities relevant for the current binary choice between the take-card and the stop option. Accordingly, participants only consider whether the prospect of taking an additional card would yield an outcome greater than zero (i.e., a positive expected utility). In contrast, in the second model variant (RU-reference point), participants take into account the history of the previous gains in the current round. Because both variants appear to be plausible potential decision mechanisms to capture individuals' decisions in the CCT, we implemented both types of reference points in all of the following expectation-based model variants.

Expectation-based models. Expectation-based models describe individuals' behavior in terms of expectation maximization, which is generally expressed as the average of an option's outcomes weighted by its probabilities. Expected-value maximization represents the most basic form of this model. The expected value (EV) model assumes risk-neutral behavior.

To take into account the stochastic nature of people's preferences, we use an exponential choice rule, in which the probability of turning over a card is modeled as a monotonic function of the differences between the expected values of both options. One free parameter (θ) scales the degree of determination in the choice function, with lower values denoting more random choice and higher values denoting more deterministic choice. We use the exponential choice rule for all following expectation-based models.

The expected utility (EU) model assumes that people maximize the expected utility of their decision, where the utility (defined by a utility function) represents the subjective value of the risky outcomes. We formalize the utility function by a power function with one free parameter (α), describing its curvature, and with this quantify individual differences in risk preferences.

However, people have been shown to violate the principles of EU theory (for reviews see Camerer, 1992; Starmer, 2000). CPT accommodates these violations by allowing additional mechanisms to influence risky choices. These mechanisms are embedded in CPT's three basic features: (a) A value function is defined over subjective gains and losses, which accounts for the observation that people are sensitive to changes in their status quo (rather than their overall wealth level). As in in the EU model, the curvature of this function is defined by the parameter (α) that quantifies the decision maker's sensitivity to differences in the magnitude of outcomes. (b) Loss aversion reflects that people typically overweight losses relative to gains. The loss aversion parameter λ, quantifies how participants weight losses relative to gains, with values larger than 1 indicating that losses are weighted more strongly than gains. (c) Probability weighting addresses the fact that people's subjective representation of probability often deviates from the objective probabilities by overweighting small and underweighting large probabilities. The functional form is defined by two parameters (η, δ) quantifying the decision maker's sensitivity to differences in probabilities and under- or overweighting of probabilities, respectively. The details about the implementation of the models are described in Appendix A.

Which features an expectation-based model needs to adequately describe behavior in the hot CCT is an empirical question that is of central interest in this paper: To this end, we compare models that are differently flexible and thus differently complex in explaining behavior in the CCT. Specifically, we test the EV model, and the EU model, as well as 12 variants of CPT (differing in model complexity) against each other. For CPT we estimate the parameters for the value and weighting functions either jointly for gains and losses or separately for gains and losses (separate parameters for gains and losses can capture whether participants' marginal utility and/or perception of probabilities differs for gains versus losses). Further, we consider both a one-parameter and a two-parameter probability weighting function (the two-parameter function additionally accounts for general under-/ overestimation of probabilities, respectively). Therefore, individual differences in making decisions in the CCT are captured by four parameters in the simplest version of CPT (value function: α; probability weighting: η; loss aversion: λ; choice stochasticity: θ), and eight parameters in the most complex version of CPT (value function: $\alpha+$, $\alpha-$[3]; probability weighting: $\eta+$, $\eta-$; probability under-overestimation: $\delta+$, $\delta-$; loss aversion λ; choice stochasticity: θ). A complete overview of all the tested model variants is shown in Table B1, Appendix B.

Constant probability (CP) model. According to the expectation-based models, participants take the information of a game round (i.e., gain amount, loss amount, number of loss cards) into account when making decisions. However, it may be that some participants ignore this information and rather follow the simple strategy of turning over cards with a constant probability (e.g., at each binary decision to turn over a card or to stop, a participant might have a 90% probability that they will turn over a card and a 10% probability that they

[3] The + and – after the Greek parameter letter indicate that the parameter applies only to gains or losses, respectively.

will stop voluntarily). With the sequential structure of the CCT, turning over cards with a constant probability results in a negatively accelerating decrease in compound probability. In other words, this model can account for behavior of participants who ignore gain and loss amounts, but frequently turn over the first few cards and then stop turning over cards with rapidly increasing probability (for example, the overall probability of turning over the first card in our example would be 90%, the probability of turning over a second card would be 0.9 × 0.9 = 81%, the probability of turning over a third card would 0.9 × 0.9 × 0.9 = 73%, etc.). Thus, the CP model should sensitively identify participants who simplify the situation and give little attention to the different information when making choices.

Model Estimation

Each model was estimated separately for each participant using a maximum likelihood method. Specifically, we used the Nelder-Mead Simplex method to estimate the parameter combination that maximizes the likelihood of the observed choices. The search space was restricted to ($0< (\alpha+, \alpha-) < 3$), ($0 < (\eta+, \eta) < 3$), ($0 < (\delta+, \delta-) < 3$), ($0 < \lambda < 10$), and ($0 < \theta < 10$). These ranges include the commonly reported values for CPT (e.g., Glöckner & Pachur, 2012; Scheibehenne & Pachur, 2015) but also allow for values beyond this range (which seems desirable given that CPT has not previously been applied to the CCT). To reduce the risk of local minima, each search was repeated 20 times, with new random starting points in the range of the search field of the parameters. We found for differently complex models that 20 repetitions are sufficient to yield stable parameter estimates with the subsequent optimization. Accordingly, the best-fitting parameter combination was chosen among these 20 repetitions.

Model Selection Approach

The main objective of the present study is to establish a formal model that allows for the decomposition of choice behavior in the hot CCT and may thus be used as a measurement model to quantify individual differences in the CCT. Accordingly, an ideal model should accurately and parsimoniously describe behavior in the CCT for a majority of individuals. At the same time it should be possible to estimate the parameters that account for the individual differences in risk taking in an unbiased and precise fashion. With these goals in mind, we base the model comparisons on overall fit and parsimony in terms of the averaged (across participants) Bayesian Information Criterion (BIC, Schwarz, 1978), which penalizes the models for their numbers of parameters. To further account for the adequacy of the models at the individual level, we additionally calculate the rank of the BIC for each model for each participant and compare the aggregate (across participant) ranks (cf. Buschena & Zilberman, 2000; Hey, 1995; Hey & Orme, 1994; Stott, 2006).

After delineating a set of adequate models we submit these best-ranked models to parameter recovery analyses, to ensure that the best-performing model not only captures the empirical data well but also allows for an unbiased and precise parameter estimation. For that purpose, we use the median parameter estimates across participants of the best-fitting models to simulate choice behavior of a prototypical participant for the CCT. We then estimate the model again using these choices. This process is repeated 1,000 times and thus provides a distribution of estimates. The results are used to evaluate whether the data-generating parameters can be recovered accurately, preferably with unimodal distributions of low variance.

Methods

Participants and Procedure

We used the data of 191 healthy participants (120 woman and 71 men) with an average age of 24.8 years (*SD* = 2.9) taking part in a large-scale study on risk preferences (Pedroni et al., 2017, Frey et al., 2017, Dutilh et al., 2017; https://osf.io/rce7g). We fixed the sample before the start of the study to the first 200 participants of the study. With this sample size we aimed to balance exceeding computational time to estimate each of the 28 models for each participant with a sample large enough to cover large parts of the variability in decision making between the participants. 9 of the 200 first datasets were incompletely recorded due to technical difficulties, resulting in the effective sample of 191 participants. The completion of the CCT and the self-report measures (described below) took about 40 minutes. The experiment took place at the University of Basel, Switzerland (*n* = 73) and the Max Planck Institute for Human Development in Berlin, Germany (*n* = 118). The respective local ethics committees approved the study, which was conducted in accordance with the Declaration of Helsinki. Participants received a detailed explanation of the study, and written informed consent was obtained.

In addition to the CCT, participants completed seven other incentivized behavioral tasks. At the end of the experiment, one of the incentivized behavioral tasks was randomly selected to be played out. If the CCT was selected to be played out, three game rounds were chosen randomly and summed to a total payoff. One point was worth CHF 0.02 in Switzerland and € 0.015 in Germany. The payoff (which could be negative or positive) was added to a flat-fee payment for participating in the study.

Material

Modifications to the CCT. We used the "hot" version of the CCT, with presentation and structure of the task identical to the original set-up (Figner et al., 2009). However, to obtain precise parameter estimates of the value and probability weighting functions it is advantageous to sample decision behavior for a wide range of outcome magnitudes and probability levels. Thus, we extended the current task to include game rounds with medium to high probabilities of losses; that is, 10, 16, 20, or 28 loss cards (out of the total 32 cards in each round), compared to the maximum of 3 loss cards in the original CCT. This provided additional data points to estimate the probability weighting function of CPT for medium to high loss probabilities. Second, to improve precise estimation of the value functions, we also extended the range of gain amounts to amounts of 10, 20, 50, 75, 100, 150, 200, 300, 400, 600 points per gain card (original version: 10, 20, and 30 points per gain card). Third, the loss amounts were similarly extended to 25, 50, 75, 100, 150, 250, 500, 750 points per loss card (original version: 250, 500 and 750 points per loss cards). Note that the placement of cards of all types was random. Contrary to the original CCT, the combination of the game parameters in the 84 game rounds in this study (see Appendix C) did not represent a fully factorial orthogonal design. Participants were interactively instructed about the game, played two practice rounds, and answered four questions to check whether they fully understood the task. After correctly answering all questions, participants started the game.

Self-report measures of risk taking. We also asked participants to complete the domain-specific risk-taking scale (DOSPERT; Weber, Blais, & Betz, 2002) in its German version (Johnson, Wilke, & Weber, 2004). The DOSPERT asks about the likelihood of engaging in various risky activities in six domains of life ("investment," "gambling," "health/safety," "recreational," "ethical," and "social"), as well as the perceived benefit and the perceived risk of these activities. The DOSPERT assesses risk taking in different everyday domains and also allows measurement of the influence of perceived risks and expected benefits on the risk-taking tendencies (Harrison, Young, Butow, Salkeld, & Solomon, 2005). For each of the 30 items participants rated first the *likelihood* of engaging in the specific risky behavior, second the *perceived risk*, and third the expected *benefits*. The DOSPERT was scored as in Johnson et al. (2004), computing the mean of each of the six domains for the rating scales "likelihood," "perception," "benefit" as well as the total scores across all different domains.

Furthermore, we assessed participants' sensation seeking and impulsivity (Lauriola et al., 2014), both of which have often been linked to risk taking. Zuckerman (2007, p. 27) defines the sensation-seeking trait as individual differences "in the seeking of varied, novel, complex, and intense sensations and experiences, and the willingness to take physical, social, legal, and financial risks for the sake of such experience." Conceptually, sensation-seeking is connected to risk taking by means of the search for excitement and arousal that often comes with high degrees of risk (Zuckerman, 2007). We used the German version (Beauducel, Strobel, & Brocke, 2003) of the *Sensation Seeking Scale (SSS), version V* (Zuckerman, Eysenck, & Eysenck, 1978). The SSS contains four subscales (each with 10 items), representing the factors of "thrill and adventure seeking" (TAS), "experience seeking" (ES), "disinhibition" (DIS), and "boredom susceptibility" (BS). Figner et al. (2009) found that their scale "need-for-arousal"—which is arguably similar to sensation seeking—was significantly correlated with risk taking in the hot CCT.

Impulsivity has been often associated with risky behaviors such as drug abuse, risky driving, unprotected sex, and problem gambling (e.g., Chambers & Potenza, 2003; Dahlen, Martin, Ragan, & Kuhlman, 2005; de Wit, 2009; Hoyle, Fejfar, & Miller, 2000). Impulsivity was measured with the German version (Preuss et al., 2008) of the *Barratt-Impulsiveness-Scale* (BIS-11, Patton et al, 1995). Based on the 30 items, we computed the three second-order factors "attentional," "motor," and "non-planning" impulsiveness.

We used Spearman rank correlations to examine the association of the total number of cards turned over by each participant, the number of game rounds in which the participant encountered a loss card, and the model parameters of the most adequate model (the CPT-1 model, see below) with (a) the DOSPERT measures on the three scales *likelihood* of risky behavior, *perceived risk* and expected *benefit*, (b) trait impulsivity, and (c) trait sensation seeking. Given that the CCT shares similar key features with tasks such as the BART, ART, and Devil's Task with high external validity, we expected that CCT measures would be correlated with these self-report scales.

Results

Card Selection

In this modified version of the hot CCT participants showed risk-seeking behavior, turning over, on average, 414.2 (*SD* = 124.9) cards in 84 trials, relative to the risk-neutral number of 336 cards following from expected-value maximization. In total, participants made

33.9 (*SD* = 13.5) stopping decisions. Thus, 8.2% of all decisions were stopping decisions. For a single round of the game, participants turned over an average of 4.9 cards, exceeding the number of cards (4.0) predicted if participants maximize expected value[4] (which would indicate risk-neutral behavior) (see Figure 1). Due to this risk-seeking behavior, an average of 50 (±13.5) rounds ended with a loss card. Only considering those (on average 34) rounds that were ended by the participants voluntarily, participants took, on average, 6.7 (±3.8) cards per round before stopping.

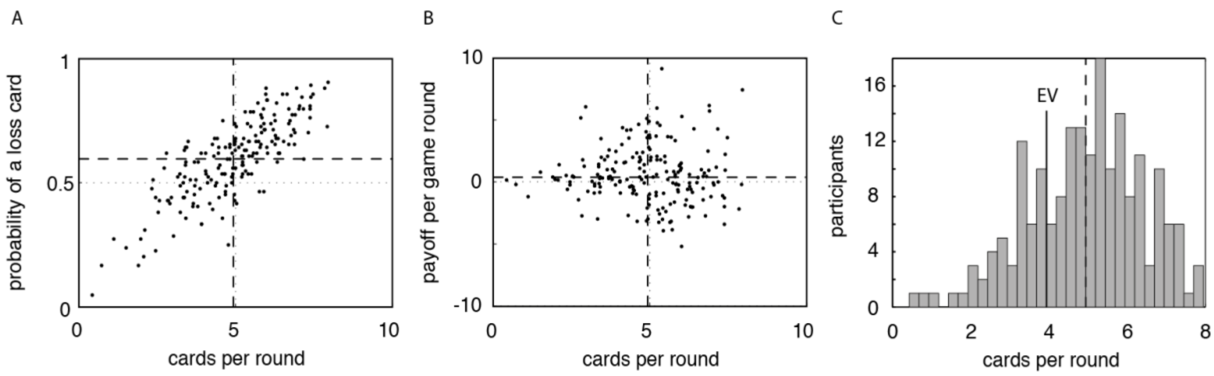

Figure 1. Descriptive Results. The dashed lines in all plots indicate the average behavior observed in the modified CCT. A) illustrates a linear increase in the probability of encountering a loss card with increasing number of cards (consistent with the lay definition of risk: turning over more cards is associated with a higher probability of encountering a negative outcome). Consequently, as evident in B) higher numbers of cards turned over yield more widely spread average payoffs (consistent with the economic definition of risk: turning over more cards is associated with greater outcome variability). C) The histogram of the number of cards turned over illustrates that a large majority of participants turns over more cards than a risk-neutral agent maximizing expected value (indicated by the solid line).

Computational Modeling

In a first step, we compared the expectation-based models with respect to the different reference point specifications and found that the models assuming that the reference point is

[4] Such risk-seeking behavior is commonly observed in the CCT.

updated after each decision (DU-reference point) performed substantially better than the models assuming reference-point updating after each round (RU-reference point). The BICs of the RU models are, on average, approximately twice as large as the BICs of the models with DU reference points (481.4 vs. 248.7) (see Appendix B, Table B1).

In a second step, we therefore focused only on the 14 expectation-based models with DU-reference point and first compared the models with respect to the average BIC across participants. The average BICs of the EV (604.0) and EU model (256.6) were substantially larger than the BICs of the 12 CPT models (all ≤ 221.1), indicating that CPT models more adequately describe behavior in the CCT. Within the CPT models we found comparable average BICs, ranging from 215.3 for the CPT-2 model to 221.1 for the CPT-12 model (see Appendix B, Table B1).

Considering model adequacy at the individual level in terms of the average rank of the BICs (i.e. the model with the lowest, best BIC value was given the lowest, best rank of 1 for the specific individual), the CPT-1, CPT-2, and CPT-7 models (ranks: CPT-1 = 4.3, CPT-2 = 3.6, CPT-7 = 3.9, with lower rank indicating better fit) outperformed the other CPT-models (with average ranks in the range of 5.4 to 11.1). Similarly, the CPT-1, CPT-2, and CPT-7 models had the lowest BIC for a majority of participants. This was the case for 149 out of 192 participants, with 73 participants for whom the CPT-1 model yielded the lowest BIC, and with 45 and 31 participants for whom the CPT-2 and the CPT-7 model, respectively, yielded the lowest BICs (see Appendix B, Table B1).

In a third step, we submitted these three best-performing models to a model recovery analysis, gauging their ability to precisely measure the parameters that generated behavior in the CCT. The results of the parameter recovery are illustrated in Appendix B, Figure B1. For the CPT-7 model, parameter recovery was particularly poor: First, the choice sensitivity parameter θ was severely biased and showed tendencies for a bimodal distribution, whereas the outcome sensitivity parameter for losses $\alpha-$ was overestimated. Second, the variability of the recovered parameters was very high for this model, indicating that parameter estimates may be imprecise; these problems likely stem from the increased complexity of the CPT-7 model. The parameters for the CPT-1 and the CPT-2 model, by contrast, showed little bias. However, the probability weighting parameter η in the CPT-2 model was recovered with large variance and followed a bimodal distribution.

In sum, these results suggest that from the expectation-based models, CPT-1 is to be preferred to describe people's choice behavior in the CCT. CPT-1 includes three parameters: The α parameter adjusts the functional form of the value function, the loss aversion parameter λ , and the η parameter accounts for distortions in probabilities via the probability weighting function. This model is the best-fitting model for the largest group of participants according to its BIC values. In addition, it has the smallest number of free parameters, making it the preferred model by parsimony. At last, it is the only model for which the data-generating parameters can be recovered accurately in an unbiased manner and with reasonably low variance.

CPT-1 yields relatively high estimates for the choice sensitivity parameter θ, meaning that participants are described as making choices in a very deterministic non-random manner. One possible explanation for this may be the sequential structure of the CCT: High values in θ implies that cards will be turned over (or that a participant stops) with high probability even when the subjective values of taking a card versus stopping differ only slightly. In game rounds in which participants turn over more than two cards, it can be advantageous for a model to predict a high probability of taking cards, because the goodness of fit in such game rounds is more influenced by predicting high likelihoods of turning over cards, compared to predicting the stopping of the last decision in a game round (in the case where a participant has stopped). This effect becomes more pronounced in game rounds that end with loss cards: A model achieves a better fit when it predicts a very high probability for card turns. Because our observations include 91.8% of card turn decisions compared to 8.2% of stopping decisions, it is reasonable to assume that these effects led to the high values in θ so that the estimates often fell at the boundaries of the search space (see Table 1). When not restricting the search space for θ, we observed for a proportion of participants that the parameter estimates of θ approached very large values (e.g., 50 and larger); however, with only marginally increasing model fits. We therefore decided to fix θ to a value of 10 and re-estimated the model only for the three remaining parameters and found no qualitative change in the indices of model adequacy as well as the median parameter estimates (see Table 1). Thus, this step reduced model complexity without loss of information.

Table 1. *Maximum likelihood parameter estimates for the CPT-1 model and the CPT-1 model with fixed choice sensitivity (θ = 10)*

CPT-1					CPT-1 (θ = 10)			
	α	η	λ	θ		α	η	λ
Mean	0.13	0.43	0.46	9.48	Mean	0.13	0.31	0.5
First quartile	0.05	0.18	0.19	10	First quartile	0.05	0.17	0.25
Median	0.12	0.27	0.46	10	Median	0.11	0.26	0.49
Third quartile	0.19	0.43	0.67	10	Third quartile	0.19	0.41	0.68
IQR	0.13	0.25	0.49	0	IQR	0.14	0.24	0.43

In a final step, we compared the CPT-1 model with fixed choice sensitivity to the CP (i.e., constant probability) model, which assumes that participants ignore round-specific information. The CPT-1 model outperformed the CP model with respect to the average BIC (CPT-1: BIC=212.6, CP: BIC=229.0); moreover, for 114 out of 191 (60%) participants, CPT-1 was the model with the best descriptive adequacy. The CP model, however, described the observed behavior of 77 (40%) participants better than the CPT-1 model. This indicates that a notable portion of participants seemingly did not take round-specific information into account.

As shown in Figure 2A, the CP model and the CPT-1 model describe behavior similarly well in rounds with many loss cards, in which only few cards can be turned over. In contrast, the CPT-1 model is additionally capable of capturing behavior in game rounds with few loss cards, in which generally more cards are turned over and which therefore are likely to be diagnostically more valuable (please note also that in the original CCT, there were only few loss cards in all game rounds, namely up to 3 loss cards). Thus, the CPT-1 model provides a relatively unbiased prediction of the average number of turned-over cards across the whole observed range whereas the CP model performs well only in the lower range (which was not included in the original CCT version). Importantly, as shown in Figure 2B the CPT-1 model also captures behavior at the individual-participant level with high accuracy.

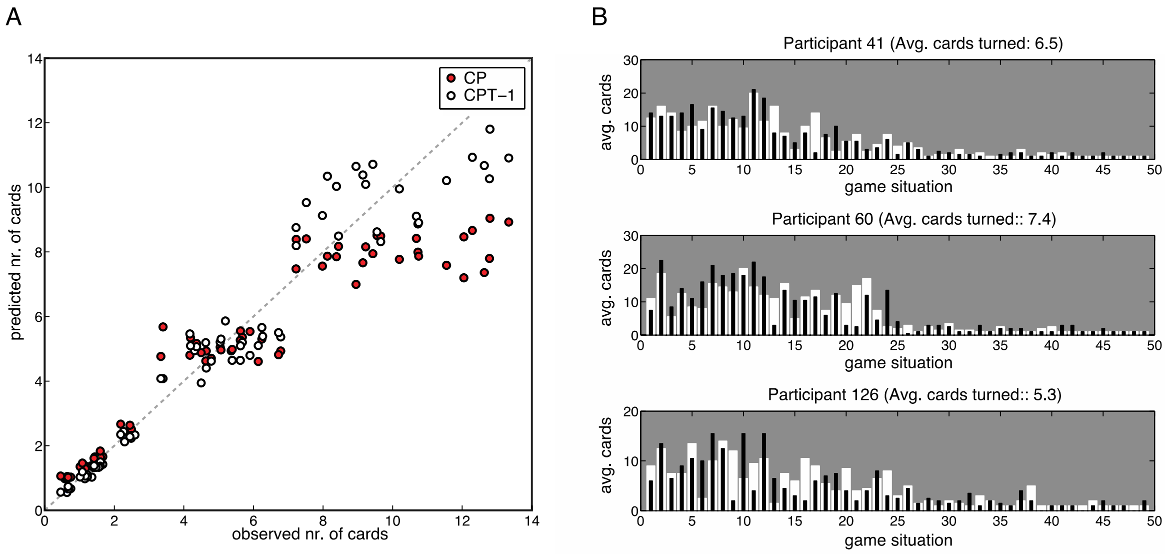

Figure 2. Model predictions. A) Scatterplot of the average number of cards turned over across participants against the predicted number of turned-over cards of the CP model and the CPT-1 model. B) Predictions and observed behavior of three exemplar participants, selected to illustrate participants who differed substantially in the CCT. Depicted is the number of cards that the CPT-1 model predicts (in black) given the estimated parameters in 49 game situations. Note that 35 of the 84 rounds were played twice, resulting in 49 unique rounds. The underlying white bars show the number of turned-over cards of the respective participant.

Parameter estimates. Recall that the majority of participants was risk seeking, as they turned over, on average, 23% more cards than predicted by risk-neutral expected-value maximization. The CPT-1 model accommodated this high risk seeking in three ways: First, by assuming low sensitivity to differences in outcome magnitudes (α: *Md* = 0.11); second, by substantial distortions in probability weighting (η: *Md* = 0.26); and third by assuming a loss aversion parameter lower than one (λ: *Md* = 0.49), thus in fact suggesting gain seeking. This implies that participants put little weight on the nominal differences in the magnitudes of possible outcomes (i.e., gains and losses) when turning over a card, whereas gains are weighted more strongly than losses. The extreme distortion in probability weighting indicates a reduced sensitivity for the differences in gain and loss probabilities (or extreme over-/ underweighting of low and high probabilities, respectively).

To illustrate the impact of the parameters on the choice behavior predicted by the CPT-1 model, consider an agent making choices according to this model (based on the median parameter estimates), facing the decision of whether or not to play a fictitious lottery with two possible outcomes. This lottery yields a gain of $50 with a probability of .6 and a loss of $100 with a probability of .4. The α parameter of 0.11 scales down the nominal differences between the two outcomes to subjective values of 1.54 and -1.66. Because the agent is gain seeking (λ = 0.46) the subjective value for the loss outcome is reduced to -0.76. The distortion in probability weighting also reduces the differences in probabilities, resulting in decision weights of 0.42 for the gain outcome and 0.37 for the loss outcome. Multiplying the decision weights and the subjective values renders a prospect to play the lottery of 0.25 compared to the prospect of not playing of 0. Thus, in contrast to an expected-value maximizer, who would not play the lottery (as the expected value of playing yields -10), the agent would prefer to play the lottery and hence shows pronounced risk seeking, paralleling the observed risk-seeking behavior in our data.

Correlations with Self-Reported Risk Taking

To explore the relationships between CCT behavior and the self-reported risk-related measures, we computed the correlations between participants' behavior in the CCT and the estimated parameters of the CPT-1 model on the one hand and participants' responses to the DOSPERT (i.e., the likelihood, the perceived benefits, and the perceived risk of engaging in various risky activities in six domains of life) and to the subscales of the SSS and BIS-11 scores on the other hand. Table 2 presents the rank correlations between the CCT measures and the risk-related self-report measures. Neither the number of cards turned over nor the number of rounds in which a loss card was encountered correlated with any of the self-report measures (this contrasts with Figner et al., 2009, who found a positive correlation between risk taking in the hot CCT and self-reported need-for-arousal). In contrast, the value function parameter α correlated positively with 4 of 6 subscores of the DOSPERT *likelihood* scale. In addition, the probability weighting parameter η correlated positively with the DOSPERT health subscore on the *likelihood* scale. Interestingly, the parameter α, accounting for the sensitivity in outcome differences was positively related to 5 of the 6 subscales of the DOSPERT *benefit* scale. Thus, the more sensitively a person responded to differences in the magnitude of outcomes (i.e., higher α), the larger s/he perceived the benefits of taking risks in various domains. In contrast, α correlated (negatively) only with the perceived risks in the investment domain, indicating that the more sensitively a person responded to differences in the outcome magnitudes, the less risky she perceived the risk of pursuing financial risky activities. We did not find significant correlations between any dependent variables of the CCT and impulsivity, and only two significant positive correlations with subscales of the sensation seeking scale (α and η with the disinhibition subscale). Because we used a *p* value criterion of $< .01$ that was unadjusted for multiple testing (consistent with the exploratory nature of assessing these potential relationships), one has to be cautious in interpreting the results. But it appears that (a) especially the value parameter α, which accounts for individual differences in outcome sensitivity, might capture an aspect of risky choice that might be relevant for risk taking beyond the context of the CCT itself and it further appears that (b) the CPT modeling (again, mainly in the form of the α parameter) might indeed lead to greater generalizability compared to simply using an aggregate CCT score.

Table 2. *Correlations between variables of the CCT and risk-related self-report measures. An asterisk indicates a significant correlation at p < 0.01 (unadjusted for multiple testing). DOSPERT: E = Ethical, I = Investment, G = Gambling, H = Health, R = Recreational, S = Social, L = Likelihood to engage in risky behavior, PR = Perceived risk if engaging in risky behavior, B = Perceived benefit of risky behavior, BIS-11: Att = Attentional impulsivity, Mot = Motor impulsivity, NP = Non-planning impulsivity, SSS-V: TAS = Thrill and adventure seeking, ES = Experience seeking, D= Disinhibition, BS = Boredom susceptibility, NLC = number of loss cards turned over (i.e., number of rounds in which a loss card was encountered), NC = number of cards turned over.*

Scale		NLC	NC	α	η	λ
DOSPERT E	Likelihood	-0.09	0.01	**0.21***	0.08	-0.01
DOSPERT I	Likelihood	-0.17	-0.04	**0.19***	0.16	-0.00
DOSPERT G	Likelihood	0.00	0.02	0.12	0.04	-0.12
DOSPERT H	Likelihood	-0.13	-0.09	**0.20***	**0.21***	-0.00
DOSPERT R	Likelihood	-0.05	-0.02	**0.21***	0.09	-0.14
DOSPERT S	Likelihood	-0.06	-0.06	0.18	0.08	-0.01
DOSPERT Total	Likelihood	-0.13	-0.04	**0.29***	0.18	-0.09
DOSPERT E	Risk perception	0.00	0.02	-0.07	-0.01	0.01
DOSPERT I	Risk perception	0.13	0.07	**-0.26***	**-0.21***	0.08
DOSPERT G	Risk perception	-0.01	0.07	-0.10	-0.03	0.09
DOSPERT H	Risk perception	-0.04	-0.07	-0.07	-0.10	0.00
DOSPERT R	Risk perception	-0.01	-0.02	-0.16	-0.08	0.11
DOSPERT S	Risk perception	-0.02	0.03	-0.08	-0.02	0.16
DOSPERT Total	Risk perception	-0.00	0.01	**-0.19***	-0.11	0.09
DOSPERT E	Benefit	-0.02	0.02	**0.20***	0.18	0.09
DOSPERT I	Benefit	-0.11	-0.02	**0.20***	**0.21***	0.17
DOSPERT G	Benefit	-0.01	0.05	0.07	0.09	0.05
DOSPERT H	Benefit	-0.01	-0.02	**0.19***	0.12	-0.01
DOSPERT R	Benefit	-0.04	-0.04	**0.23***	0.10	-0.09
DOSPERT S	Benefit	-0.07	-0.02	**0.24***	0.10	-0.07
DOSPERT Total	Benefit	-0.08	-0.02	**0.29***	**0.19***	0.02
BIS-11 Att		-0.04	0.05	0.01	-0.05	-0.10
BIS-11 Mot		0.12	0.06	-0.01	-0.16	-0.18
BIS-11 NP		0.08	0.05	0.06	-0.07	-0.05
SSS-V TAS		-0.03	-0.09	0.11	0.06	-0.16
SSS-V ES		0.09	0.02	0.12	0.05	-0.09
SSS-V D		-0.11	-0.07	**0.22***	**0.19***	0.06
SSS-V BS		-0.00	-0.02	-0.00	0.01	0.05

Discussion

We combined two complementary approaches to investigate decision making under risk. In the systematic-variation approach, individuals' risk taking is assessed using choices in well-defined lotteries. This approach has the advantage that individual behavior can be formally described in terms of psychological constructs (e.g., loss aversion) using expectation-based models and can (at least theoretically) be generalized to other decisions. In the second, naturalistic approach, risk taking is assessed in more playful, game-like tasks that are suggested to reflect more affective and realistic risky situations, thus promising higher ecological relevance and validity. To combine the beneficial aspects of both approaches, we developed and tested theoretically founded expectation-based models for the hot CCT with the goals of (a) better understanding how individuals make decisions in the CCT itself, (b) being able to quantify individual differences in the CCT on the basis of clearly interpretable psychological mechanisms, and (c) exploring whether overt behavior and model parameters of the CCT are related to other risk-related self-report measures.

The model comparisons favored a model according to which, when decision makers choose to stop or take another card in the CCT, they only consider whether taking another card leads to a subjective value that is positive; this suggests that they update their reference point after every binary decision of whether to take a card or stop and ignore the payoffs they have previously received. This is in line with Tversky and Kahneman’s (1991, p. 1046) proposal that “… the reference state usually corresponds to the decision maker's current position.“ Consistent with our results, it has been shown that people readily accommodate gains but not losses (Jervis, 1992; Kahneman, Knetsch, & Thaler, 1991). Potentially, the structure of the CCT further amplifies this tendency, as subsequent decisions (within the same game round) reflect situations after gains (except for the first decision in a new game round), whereas encountering a loss card terminates the current game round. That may also explain differences between the hot and the cold CCT that have been described in earlier work (e.g., Figner et al., 2009): In the cold CCT, participants are forced to make a decision about the total number of cards that they want to turn over in each round, whereas in the hot CCT, participants can avoid such an "overall" decision (which is arguably more complex) and instead rely on a more local, and thus simpler, strategy that considers only whether or not to turn over the very next card, neither considering previous events nor making considerations beyond the very next card.

The finding that models with continuous updating of reference points captured the data best furthermore reveals how participants may perceive the decision situation in the CCT. Specifically, participants seem to perceive the gain amount as constant throughout a game round (and not increasing after every card), whereas the probability of turning over a loss card versus a gain card increases. Thus, risk taking in the CCT may be experienced as exposure to increasingly more likely losses and increasingly less likely gains. Accordingly, the CCT differs in how risk taking is experienced compared to the BART, a closely related task. Similar as in our study, Wallsten, Pleskac, and Lejuez (2005) compared a number of computational models to describe behavior in the BART. The best-performing model in their analysis suggested that in the BART, participants consider in each decision the overall value of the potential gains relative to the starting point instead of the updated marginal value of a potential gain (but see Schonberg et al., 2012 for an alternative interpretation). Our results suggest that the choices in the CCT may be perceived by participants similarly to the decisions in risk tasks of the systematic-variation approach, particularly ones such as the Holt and Laury (2002, 2005) procedure, in which the probabilities of monetary outcomes are varied systematically, while the monetary outcomes are held constant.

The observed risk-seeking behavior in the CCT, however, clearly diverges from the pattern of risk aversion usually found in tasks following the lottery paradigm (e.g., Holt &

Laury, 2002, 2005; but see, e.g., Berg et al., 2005). Risk aversion is also regularly observed in studies using game-like paradigms (e.g., Lejuez, 2002). In our view, this risk-seeking behavior in the present study (note that other studies using the CCT have also found risk-seeking behavior, e.g., Figner et al., 2009) may be due to how participants evaluate the valence and magnitudes of outcomes and probabilities, as reflected in the parameter estimates of our favored computational model.

Using Cumulative Prospect Theory for modeling behavior in the hot CCT

The most adequate model (the CPT-1 model) incorporates a set of three clearly identifiable parameters (tested with a parameter recovery approach), which parameterize the key components of CPT: The α parameter adjusts the functional form of the value function, reflecting participants' sensitivity to differences in outcomes. The loss aversion parameter λ reflects differential weighting of gains relative to losses, and the η parameter accounts for distortions in probabilities via the probability weighting function.

Our results show rather low average values for the three parameter estimates compared to parameter estimates that are typically found when participants make risky choices in the commonly used lottery paradigms (e.g., Tversky & Kahneman 1992). The low average estimates for the value function parameter that we observed imply that in our study a typical participant rather insensitively evaluates the nominal differences in outcomes. In addition, the estimates of the probability weighting function suggest pronounced overweighting of small and underweighting of large probabilities, with insensitivity to changes in medium range probabilities. Finally, the estimates of the loss aversion parameter are on average below 1, indicating that a typical participant may be more inclined to seek gains than to avoid losses. Thus, the combination of these components offers an explanation for the overall risk-seeking behavior, suggesting that participants in the CCT were generally more attracted by gains compared to losses, while only marginally taking into account nominal differences between gains and losses and the underlying probabilities. Therefore, seemingly paying relatively little attention to the outcomes and underlying probabilities, participants may feel attracted to turn over cards and do so quite excessively.

Risk-seeking Behavior in the CCT

We believe that several factors might be contributing to the characteristic pattern of parameter estimates and the associated high level of risk seeking in the CCT. Moreover, these factors are likely to contribute to increased dynamic risk-taking in real-world situations as well. First, and perhaps most obviously, the "hot" version of the CCT presents the decision maker with up to 32 different cards that can be turned over, and one single button for stopping. Thus, the hot CCT may incorporate demand characteristics that make turning over a card the default action, presumably increasing the probability of choosing a card rather than not drawing another card. In some real-life risk-taking situations, taking the risk may also be the default option (e.g., ordering another drink in a bar; drive home in the car you came with, although you have been drinking; not interrupting to put on a condom).

Second, participants may strive not only to maximize their expectation of potential monetary gains, but also may strive to experience rewarding anticipatory emotions such as feeling thrilled and exhilarated while drawing cards, again a factor likely contributing to risk-taking in the real-world. The study of Figner et al. (2009) supports that this factor contributes to risk-taking in the CCT, as the participants reported more "excitement" and "gut" decisions in the "hot" than the "cold" (i.e., non-sequential) version of the CCT.

Third, as discussed by Figner and Weber (2011), in the "hot" CCT each new game round starts with a rather low loss probability, meaning that early decisions in a game round to turn over a card are typically rewarded with a gain, likely reinforcing a tendency to turn over more cards, which then requires possibly some form of instrumental intervention or inhibitory control to stop a game round voluntarily (see also Gladwin & Figner, 2014; Gladwin, Figner, Crone, & Wiers, 2011). In real-life situations (such as smoking or unprotected sex), the negative consequences often happen only after repeated risk-taking, either because they have a generally low probability of occurring or because they are delayed for other reasons (e.g., unwanted pregnancy is unlikely to occur after a single instance of unprotected sex, and the consequences are noticed only after a while).

Fourth, the hot version of the CCT was designed to trigger affective involvement (as this is an important factor in many real-life risk-taking situations), and the observed low probability sensitivity in the CCT is thus consistent with earlier work showing that involvement of affective processes reduces probability sensitivity (Pachur et al., 2014; Rottenstreich & Hsee, 2001; Suter, Pachur, Hertwig, Endestad & Biele, 2015; Suter, Pachur, & Hertwig, 2015). In sum, it is conceivable that the CCT may promote behavior that is rather common in some situations of dynamic risk taking in real life, namely that people may pay more attention to the possible gains than the losses, are relatively insensitive to the probabilities and the outcome magnitudes, and pay more attention to the gains than the losses. These patterns are reminiscent of what can be observed in smoking, unprotected sex, and drunk driving, and thus direct empirical investigations of these relationships could be a crucial next step in better understanding the psychological mechanisms underlying excessive dynamic risk taking.

Individual Differences

In addition to the overall pattern of risk-seeking behavior in the CCT, we observe large heterogeneity in observed behavior: Risk taking quantified by the average number of turned-over cards per game round differed substantially between participants, ranging from 3.9 cards per game round in the first quartile to 6 cards per game round in the third quartile. At first sight it is not obvious how differences in parameter estimates relate to these differences. Following our conclusion that participants seem to update their reference point after each decision, the choice between turning over a card or not reflects a decision between a certain outcome (i.e., stopping) and a mixed two-outcome lottery. Because the parameter for the value function is jointly estimated for gains and losses, decreasing concavity of the shape of the value function (if considered in isolation) implies both risk aversion with respect to the evaluation of the gain-part of the lottery of turning over a card and risk seeking with respect to the evaluation of the loss-part of turning over a card[5]. The same applies for probability weighting: Underweighting of large probabilities to turn over a loss (gain) card parallels the overweighting of small probabilities to turn over a gain (loss) card. Depending on the specific game round (i.e., what gain amount, what loss amount, and how many loss cards), the functional forms of the value and probability weighting function as well as loss aversion contribute to risk taking in a complex interplay. To gain insight into this interplay, we ran simulations in which we determined the number of cards an agent would turn over given combinations of the three parameter values.

The simulations for four exemplary game rounds are illustrated in Figure 3 and show that increasing values of parameter α (in the value function) lead to less risk taking in rounds in which the loss amount is larger than the gain amount (typically in rounds with few loss cards, Figure 3AB). On the other hand, increasing values of α lead to riskier behavior in rounds in which the gain amount is larger than the loss amount (typically in rounds with many loss cards, Figure 3CD). Similarly, increasing values in parameter η lead to more risk-taking behavior in rounds with few loss cards and less risk-taking behavior in rounds with many loss cards. In contrast, increases in the loss aversion parameter λ always lead to less risk-taking behavior. Therefore, α and η can be interpreted as indicators of how sensitively participants evaluate differences in outcome magnitudes and probabilities, respectively, and can, depending on the type of outcome and probability level, lead to risk aversion or risk seeking, paralleling the fourfold pattern of risk attitudes (Kahneman & Tversky, 1992).

[5] Note, that the interpretation of the curvature of the value function in the present case of the CPT-1 model differs from its classical EUT interpretation as a direct index of a person's risk attitude.

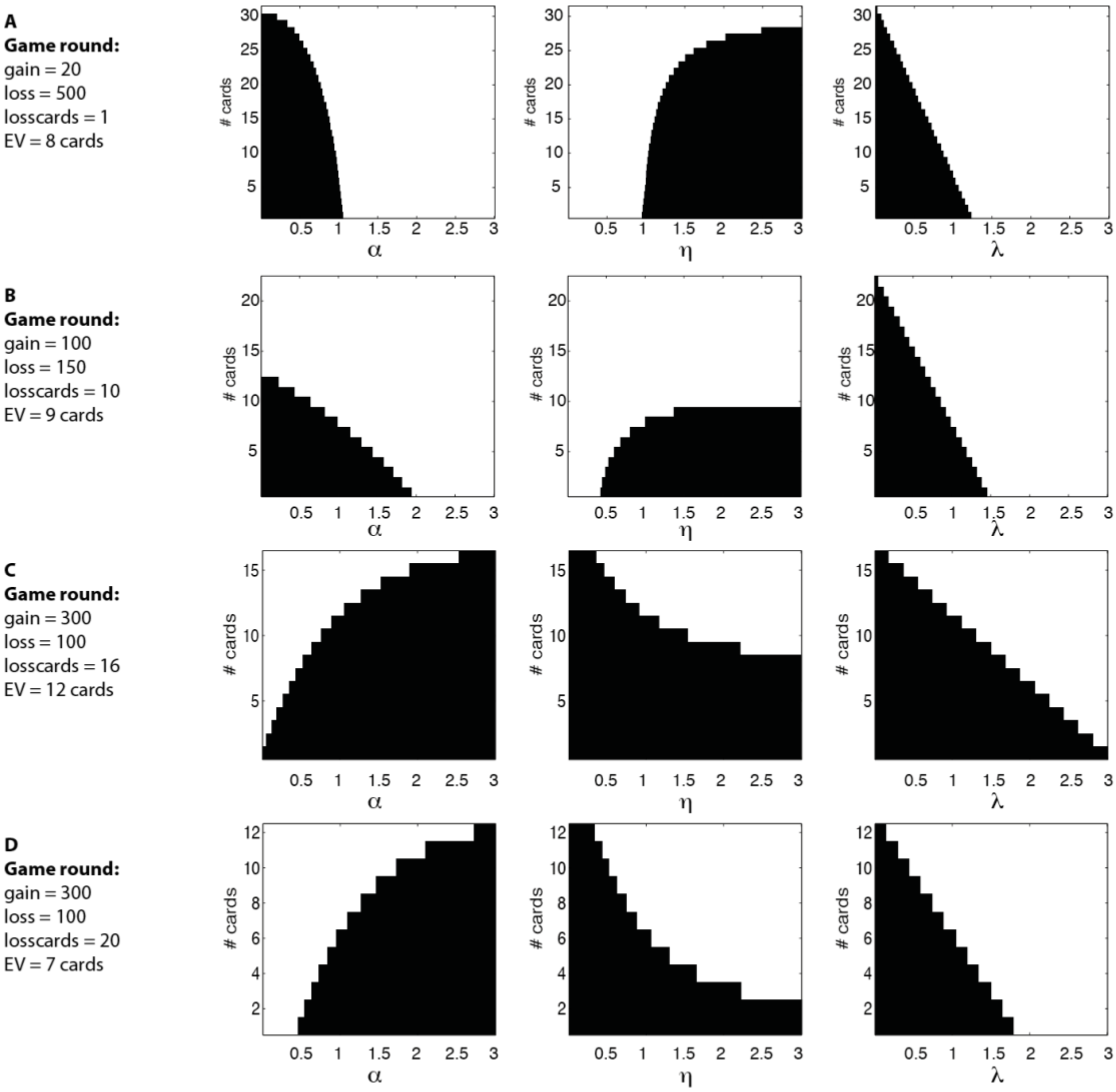

Figure 3. Interacting effects of parameters and game rounds on risk taking. A-D) In each panel, the black bars indicate up to which level (how many cards turned over) risk taking has a positive utility. We calculated the number of cards turned over by assuming that an agent makes decisions according to CPT-1 for different parameter values that cover the search space of the estimation procedure. In each column one parameter is varied (left column: α; middle column: η; right column: λ) while holding the other two constant at 1.

Convergent Validity

The parameter α, which accounts for differences in the sensitivity to outcomes, showed the most consistent pattern of correlations to self-report measures of risk taking and associated trait measures. Specifically, α was positively (although weakly) correlated with the reported likelihood of engaging in risky behavior in 4 of the 6 risk domains of the DOSPERT, and it was also positively correlated with the benefits participants would expect in 5 of the 6 risk domains. In contrast, α correlated with the risk perception scale of the DOSPERT in only one domain. This suggests that individuals who are more sensitive to differences in outcomes may also perceive more benefits in real-life risky behavior, rather than the potential risks and for this reason may report a higher likelihood of taking these risks. Previous research has similarly emphasized the role of outcome sensitivity in risk taking. For example, Wallsten et al. (2005) found that outcome sensitivity to gains (in that study referred to as γ) in the BART was predictive of the number of drugs a person had tried, whether he or she had ever stolen something, and how often he or she had engaged in unprotected sex in the last year. Similarly, Wichary et al. (2015) found that outcome sensitivity in the BART was positively correlated with the propensity to engage in various real-world risky activities (e.g., sky diving, smoking). In addition, a number of studies have shown that trait sensitivity to reward and/or loss as assessed by self-report measures (Carver & White, 1994) is related to risk taking in laboratory tasks (Brunborg, Johnsen, Mentzoni, Molde, & Pallesen, 2011; Demaree, DeDonno, Burns, & Erik Everhart, 2008; Penolazzi, Gremigni, & Russo, 2012) and real-life risk taking such as horse race gambling (Balodis, Thomas, & Moore, 2014).

Besides the rather consistent pattern of correlations between α and the DOSPERT risk taking and benefit scales, we find isolated correlations between η and single DOSPERT scales, as well as correlations of α and η with the "disinhibition" sensation-seeking subscale. Figner et al. (2009) found that the average number of cards turned over was positively related to Need-for-Arousal, a construct related to sensation seeking. We observe no correlation of any of the self-report measures either with the number of cards turned over or with the number of encountered loss cards. Given this somewhat surprising null-finding, we also correlated the overt measures only considering game rounds with 1 or 3 loss cards (the range of loss cards described in the current short version of the CCT outlined in Figner & Weber, 2011); however, again there was no significant relationship (all $r < .09$) with any of the self-report measures.

In sum, these results suggest that outcome sensitivity, as captured in the parameter α, is an important driver of observed behavior in the CCT and seems to correspond with other risk-related measures, whereas overt behavior may be less related to risk taking outside the CCT (although other explanations are also possible, for example that the newly added game rounds or the whole study context with several other similar risk-taking tasks played a role; but even if so, our results are consistent with the idea that the CPT decomposition leads to more reliable estimates of relevant processes involved in the CCT, compared to using simply the overt behavior).

Future Directions and Applications

The results of our study should be considered in the light of the finding that for 40% of the participants the CP (i.e., constant probability) model was more adequate than the CPT model. This shows not only that there is a considerable heterogeneity in directly observable risk taking in the CCT, but that participants also differ in their choice strategies, with a substantial proportion of the participants only marginally taking into account the information presented in the game rounds (see also Figner et al.'s 2009, Pedroni et al., 2017). However, this does not necessarily imply that these participants behave randomly. Turning over cards at a constant probability yields an accelerating decrease in compound probability, that is they have a high probability of turning over the first few cards but then quickly stop turning over more cards. Thus, these participants might also have in mind a target for their willingness to take risks. Interestingly, post-hoc analyses indicate that following this simple strategy is not a particularly bad one, as these participants do not perform worse in terms of their payoff (if the outcomes of all game rounds had been paid out in real money) than participants who are better described by the CPT-1 model ($t(189)= 1.26$, $p = .21$). One possible reason for this may be that although they ignored the gain and loss amounts in the respective game rounds, overall they turned over fewer cards than participants who were better described by the CPT model (on average 0.52 less cards per round, $t(189)= -2.27$, $p = 0.02$). In addition, the model comparison, which is based on the individual BICs, does not take into account in *which* game rounds either model better describes their behavior. Our results suggest that the CP-model only fits well in game rounds in which just a *few* cards are turned over (see Figure 2A). In contrast, the CPT model is also able to successfully account for behavior in trials in which *many* cards are turned over, which suggests that it is a more encompassing model and should be applicable to a wider range of choice situations.

We further emphasize that the modeling results are by nature influenced by the selection of models and thus our conclusions should be considered within this selection. Our modeling approach focused on the family of expectation-based models primarily because they are the most influential models in economic research. Because the CCT conforms to the structure and appeal of commonly applied game-like risk-taking tasks in psychological research, modeling the CCT using an expectation-based model makes the psychological mechanisms underlying risk taking directly comparable, facilitating the integration of research between the two streams of research.

To conclude, our study had the goal of bridging the gap between two approaches to studying risk taking, by uncovering clearly interpretable mechanisms that underlie decision making in a dynamic risk-taking task. We derived a parsimonious expectation-based model, reflecting the simplest possible specification of CPT, which made it possible to reveal the psychological underpinnings of the characteristic risk-seeking behavior in the CCT. The model provides indications for why participants behave in this way and how individual differences between participants may be explained. Therefore, our modeling approach may offer new possibilities to investigate individual differences in naturalistic risk taking at the level of psychological mechanisms. For instance, future studies could explore how sensitivity to differences in outcomes and probabilities as well as loss aversion develop across the life span. These findings could be integrated with current knowledge on the localization of the neural processes of these aspects and developmental changes in the brain.

In addition, in combination with the modeling approach, the CCT might be a useful tool for studying differences among neuropsychologically and/or clinically distinct populations, in a similar spirit to studies that used a model-based approach to decompose behavior in the Iowa Gambling Task (IGT) to better understand the specific characteristics in decision making in these populations (see, e.g., Yechiam, Busemeyer, Stout, & Bechara, 2005). In fact, there exist datasets of the CCT that may offer themselves for such purposes, such as a recent study of Kluwe-Schiavon, Viola, Sanvicente-Vieira, Pezzi, and Grassi-Oliveira (2016) that has shown differences in “raw” scores of the CCT between female crack cocaine users and adolescents. It would thus be interesting to investigate whether estimated model parameters may shed additional insights into how the differences in overt behavior come about. Furthermore, we believe that the CCT may be especially suitable for investigating decision making in populations that exhibit pronounced risk-seeking behavior, such as pathological gamblers. The characteristics of the CCT stimulate risk-seeking behavior, because turning over a card is likely experienced as a “pleasurable” action per se and may reflect more strongly a default option compared to stopping. Thus, the CCT possibly has characteristics that are common to casino situations in which pathological gamblers typically take excessive risks when exposed to a multitude of risky options and fail to choose the single “safe” option of leaving the casino.

References

Aklin, W. M., Lejuez, C., Zvolensky, M. J., Kahler, C. W., & Gwadz, M. (2005). Evaluation of behavioral measures of risk taking propensity with inner city adolescents. *Behaviour Research and Therapy, 43*(2), 215-228.

Balodis, S., Thomas, A., & Moore, S. (2014). Sensitivity to reward and punishment: Horse race and EGM gamblers compared. *Personality and Individual Differences, 56*, 29-33.

Baumann, J., & DeSteno, D. (2012). Context explains divergent effects of anger on risk taking. *Emotion*, *12*(6), 1196.

Beauducel, A., Strobel, A., & Brocke, B. (2003). Psychometrische Eigenschaften und Normen einer deutschsprachigen Fassung der Sensation Seeking-Skalen, Form V. *Diagnostica*, *49*(2), 61-72.

Berg J., Dickhaut J., & McCabe K. (2005) Risk preference instability across institutions: A dilemma. *Proceedings of the National Academy of Sciences of the United States of America* 102(11):4209-4214.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations Research, 30*(5), 961-981.

Bishara, A. J., Pleskac, T. J., Fridberg, D. J., Yechiam, E., Lucas, J., Busemeyer, J. R., . . . Stout, J. C. (2009). Similar processes despite divergent behavior in two commonly used measures of risky decision making. *Journal of Behavioral Decision Making, 22*(4), 435-454.

Bornovalova, M. A., Cashman-Rolls, A., O'Donnell, J. M., Ettinger, K., Richards, J. B., Dewit, H., & Lejuez, C. (2009). Risk taking differences on a behavioral task as a function of potential reward/loss magnitude and individual differences in impulsivity and sensation seeking. *Pharmacology Biochemistry and Behavior, 93*(3), 258-262.

Bornovalova, M. A., Daughters, S. B., Hernandez, G. D., Richards, J. B., & Lejuez, C. (2005). Differences in impulsivity and risk-taking propensity between primary users of crack cocaine and primary users of heroin in a residential substance-use program. *Experimental and Clinical Psychopharmacology, 13*(4), 311.

Boucsein, W. (2012). *Electrodermal activity:* Springer Science & Business Media.

Brunborg, G. S., Johnsen, B. H., Mentzoni, R. A., Molde, H., & Pallesen, S. (2011). Individual differences in evaluative conditioning and reinforcement sensitivity affect bet-sizes during gambling. *Personality and Individual Differences, 50*(5), 729-734.

Buelow, M. T. (2015). Predicting performance on the Columbia Card Task effects of personality characteristics, mood, and executive functions. *Assessment,* 22(2), 178-187.

Buschena, D., & Zilberman, D. (2000). Generalized expected utility, heteroscedastic error, and path dependence in risky choice. *Journal of Risk and Uncertainty, 20*(1), 67-88.

Busemeyer, J. R. (1980). Importance of measurement theory, error theory, and experimental design for testing the significance of interactions. *Psychological Bulletin, 88*(1), 237.

Camerer, C. F. (1992). Recent tests of generalizations of expected utility theory *Utility theories: Measurements and applications* (pp. 207-251): Springer.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology, 67*(2), 319-333.

Cavanagh, J. F., Neville, D., Cohen, M. X., Van de Vijver, I., Harsay, H., Watson, P., . . . Ridderinkhof, K. R. (2014). Individual differences in risky decision-making among seniors reflect increased reward sensitivity. Frontiers in Neuroscience, 6, 1-7.

Chambers, R. A., & Potenza, M. N. (2003). Neurodevelopment, impulsivity, and adolescent gambling. *Journal of Gambling Studies, 19*(1), 53-84.

Coffey, S. F., Schumacher, J. A., Baschnagel, J. S., Hawk, L. W., & Holloman, G. (2011). Impulsivity and risk-taking in borderline personality disorder with and without substance use disorders. *Personality Disorders: Theory, Research, and Treatment, 2*(2), 128.

Critchley, H. D., Elliott, R., Mathias, C. J., & Dolan, R. J. (2000). Neural activity relating to generation and representation of galvanic skin conductance responses: a functional magnetic resonance imaging study. *Journal of Neuroscience, 20*(8), 3033-3040.

Crowley, T. J., Raymond, K. M., Mikulich-Gilbertson, S. K., Thompson, L. L., & Lejuez, C. W. (2006). A risk-taking “set” in a novel task among adolescents with serious conduct and substance problems. *Journal of the American Academy of Child & Adolescent Psychiatry, 45*(2), 175-183.

Dahlen, E. R., Martin, R. C., Ragan, K., & Kuhlman, M. M. (2005). Driving anger, sensation seeking, impulsiveness, and boredom proneness in the prediction of unsafe driving. *Accident Analysis & Prevention, 37*(2), 341-348.

de Wit, H. (2009). Impulsivity as a determinant and consequence of drug use: a review of underlying processes. *Addiction Biology, 14*(1), 22-31.

Demaree, H. A., DeDonno, M. A., Burns, K. J., & Erik Everhart, D. (2008). You bet: How personality differences affect risk-taking preferences. *Personality and Individual Differences, 44*(7), 1484-1494.

Dutilh, G., Vandekerckhove, J., Ly, A., Matzke, D., Pedroni, A., Frey, R., ... & Wagenmakers, E. J. (2017). A test of the diffusion model explanation for the worst performance rule using preregistration and blinding. Attention, Perception, & Psychophysics, 79(3), 713-725.

Figner, B., Mackinlay, R. J., Wilkening, F., & Weber, E. U. (2009). Affective and deliberative processes in risky choice: age differences in risk taking in the Columbia Card Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(3), 709.

Figner, B., & Murphy, R. O. (2011). Using skin conductance in judgment and decision making research. In M. Schulte-Mecklenbeck, A. Kuehberger, & R. Ranyard (Eds.), *A handbook of process tracing methods for decision research* (pp. 163-184). New York, NY: Psychology Press

Figner, B., & Voelki, N. (2004). Risky decision making in a computer card game: An information integration experiment. *Polish Psychological Bulletin, 35*(3), 135-139.

Figner, B., & Weber, E. U. (2011). Who takes risks when and why? Determinants of risk taking. *Current Directions in Psychological Science*, *20*(4), 211-216.

Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. Science advances, 3(10), e1701381.

Frey, R., Rieskamp, J., & Hertwig, R. (2015). Sell in May and go away? Learning and risk taking in nonmonotonic decision problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(1), 193–208.

Gladwin, T. E., & Figner, B. (2014). ‘Hot’ cognition and dual systems: introduction, criticisms, and ways forward. *Frontiers of Cognitive Psychology Series: Neuroeconomics, Judgment and Decision Making*.

Gladwin, T. E., Figner, B., Crone, E. A., & Wiers, R. W. (2011). Addiction, adolescence, and the integration of control and motivation. *Developmental Cognitive Neuroscience*, 1(4), 364-376.

Glöckner, A., & Pachur, T. (2012). Cognitive models of risky choice: Parameter stability and predictive accuracy of prospect theory. *Cognition,* 123(1), 21-32.

Goldberg, J., & Fischhoff, B. (2000). The long-term risks in the short-term benefits: Perceptions of potentially addictive activities. *Health Psychology, 19*(3), 299.

Harrison, J. D., Young, J. M., Butow, P., Salkeld, G., & Solomon, M. J. (2005). Is it worth the risk? A systematic review of instruments that measure risk propensity for use in the health setting. *Social Science & Medicine*, 60(6), 1385-1396.

Hey, J. D. (1995). Experimental investigations of errors in decision making under risk. *European Economic Review, 39*(3), 633-640.

Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica, 62*, 1291-1326.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644-1655.

Holt, C. A., & Laury, S. K. (2005). Risk aversion and incentive effects: New data without order effects. *American Economic Review,* 95(3), 902-904.

Hoyle, R. H., Fejfar, M. C., & Miller, J. D. (2000). Personality and sexual risk taking: A quantitative review. *Journal of Personality*, *68*(6), 1203-1231.

Huang, Y., Wood, S., Berger, D., & Hanoch, Y. (2013). Risky choice in younger versus older adults: Affective context matters. *Judgment and Decision Making*, 8(2), 179.

Hunt, M. K., Hopko, D. R., Bare, R., Lejuez, C., & Robinson, E. (2005). Construct validity of the Balloon Analogue Risk Task (BART). Associations with psychopathy and impulsivity. *Assessment*, *12*(4), 416-428.

Jamieson, J. P., & Mendes, W. B. (2016). Social stress facilitates risk in youths. *Journal of experimental psychology: general*, 145(4), 467.

Jervis, R. (1992). Political implications of loss aversion. *Political Psychology*, 187-204.

Johnson, J., Wilke, A., & Weber, E. U. (2004). Beyond a trait view of risk taking: A domain-specific scale measuring risk perceptions, expected benefits, and perceived-risk attitudes in German-speaking populations. *Polish Psychological Bulletin*, *35*, 153-172.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *The Journal of Economic Perspectives*, 193-206.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263-291.

Kluwe-Schiavon, B., Viola, T. W., Sanvicente-Vieira, B., Pezzi, J. C., & Grassi-Oliveira, R. (2016). Similarities between adult female crack cocaine users and adolescents in risky decision-making scenarios. *Journal of clinical and experimental neuropsychology*, 1-16.

Lauriola, M., Panno, A., Levin, I. P., & Lejuez, C. W. (2014). Individual differences in risky decision making: A meta-analysis of sensation seeking and impulsivity with the balloon analogue risk task. *Journal of Behavioral Decision Making*, 27(1), 20-36.

Ledgerwood, D. M., Alessi, S. M., Phoenix, N., & Petry, N. M. (2009). Behavioral assessment of impulsivity in pathological gamblers with and without substance use disorder histories versus healthy controls. *Drug and Alcohol Dependence*, 105(1), 89-96.

Leigh, B. C. (1999). Peril, chance, adventure: concepts of risk, alcohol use and risky behavior in young adults. *Addiction*, *94*(3), 371-383.

Lejuez, C., Aklin, W. M., Jones, H. A., Richards, J. B., Strong, D. R., Kahler, C. W., & Read, J. P. (2003). The balloon analogue risk task (BART) differentiates smokers and nonsmokers. *Experimental and Clinical Psychopharmacology*, *11*(1), 26.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., ... & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, 8(2), 75.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, 127(2), 267.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, 805-824.

Luce, R. D., & Suppes, P. (1965). Preference, utility, and subjective probability. *In vol. 3 of Handbook of Mathematical Psychology*, edited by R. D. Luce, R. R. Bush and E. Galanter, 249-410. Wiley: New York.

MacPherson, L., Magidson, J. F., Reynolds, E. K., Kahler, C. W., & Lejuez, C. (2010). Changes in sensation seeking and risk-taking propensity predict increases in alcohol use among early adolescents. *Alcoholism: Clinical and Experimental Research, 34*(8), 1400-1408.

March, J. G., & Shapira, Z. (1987). Managerial perspectives on risk and risk taking. *Management Science, 33*(11), 1404-1418.

Marley, A. (2002). Random utility models and their applications: Recent developments. *Mathematical Social Sciences, 43*(3), 289-302.

Mellers, B., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision affect theory: Emotional reactions to the outcomes of risky options. *Psychological Science*, 423-429.

Mishra, S., Lalumière, M. L., & Williams, R. J. (2010). Gambling as a form of risk-taking: Individual differences in personality, risk-accepting attitudes, and behavioral preferences for risk. *Personality and Individual Differences, 49*(6), 616-621.

Moore, S., & Gullone, E. (1996). Predicting adolescent risk behavior using a personalized cost-benefit analysis. *Journal of Youth and Adolescence, 25*(3), 343-359.

Mosteller, F., & Nogee, P. (2006). An experimental measurement of utility. *Selected Papers of Frederick Mosteller* (pp. 179-219): Springer.

Pachur, T., Hertwig, R., & Wolkewitz, R. (2014). The affect gap in risky choice: affect-rich outcomes attenuate attention to probability information. *Decision, 1*(1), 64-78.

Panno, A., Lauriola, M., & Figner, B. (2013). Emotion regulation and risk taking: Predicting risky choice in deliberative decision making. *Cognition & Emotion, 27*(2), 326-334.

Panno A, Lauriola M, Pierro A (2015) Regulatory Mode and Risk-Taking: The Mediating Role of Anticipated Regret. *PLoS ONE* 10(11): e0143147.

Patton, J. H., & Stanford, M. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology*, 51(6), 768-774.

Pedroni, A., Frey, R., Bruhin, A., Dutilh, G., Hertwig, R., & Rieskamp, J. (2017). The risk elicitation puzzle. Nature Human Behaviour, 1(11), 803.

Penolazzi, B., Gremigni, P., & Russo, P. M. (2012). Impulsivity and reward sensitivity differentially influence affective and deliberative risky decision making. *Personality and Individual Differences, 53*(5), 655-659.

Pleskac, T. J. (2008). Decision making and learning while taking sequential risks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 167.

Prelec, D. (1998). The probability weighting function. *Econometrica, 497-527.*

Preuss, U. W., Rujescu, D., Giegling, I., Watzke, S., Koller, G., Zetzsche, T., . . . Möller, H.-J. (2008). Psychometrische Evaluation der deutschsprachigen Version der Barratt-Impulsiveness-Skala. *Nervenarzt, 79*(3), 305-319.

Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1446.

Rottenstreich, Y., & Hsee, C. K. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science*, *12*(3), 185-190.

Scheibehenne, B., & Pachur, T. (2015). Using Bayesian hierarchical parameter estimation to assess the generalizability of cognitive models of choice. *Psychonomic Bulletin & Review, 22*(2), 391-407.

Schonberg, T., Fox, C. R., Mumford, J. A., Congdon, E., Trepel, C., & Poldrack, R. A. (2012). Decreasing ventromedial prefrontal cortex activity during sequential risk-taking: an fMRI investigation of the balloon analog risk task. *Frontiers in Neuroscience, 6*.

Schonberg, T., Fox, C. R., & Poldrack, R. A. (2011). Mind the gap: bridging economic and naturalistic risk-taking with cognitive neuroscience. *Trends in Cognitive Sciences, 15*(1), 11-19.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461-464.

Siegel, R. A. (1978). Probability of punishment and suppression of behavior in psychopathic and nonpsychopathic offenders. *Journal of Abnormal Psychology, 87*(5), 514.

Skeel, R. L., Pilarski, C., Pytlak, K., & Neudecker, J. (2008). Personality and performance-based measures in the prediction of alcohol use. *Psychology of Addictive Behaviors, 22*(3), 402.

Slovic, P. (1966). Risk-taking in children: Age and sex differences. *Child Development, 37,* 169-176.

Savage, L. J. (1972). *The foundations of statistics.* Courier Corporation.

Starmer, C. (2000). Developments of non-expected utility theory: the hunt for a descriptive theory of choice under risk. *Journal of Economic Literature, 38,* 332-382.

Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty*, *32*(2), 101-130.

Swogger, M. T., Walsh, Z., Lejuez, C., & Kosson, D. S. (2010). Psychopathy and risk taking among jailed inmates. *Criminal Justice and Behavior*, *37*(4), 439-452.

Suter, R. S., Pachur, T., & Hertwig, R. (2015). How affect shapes risky choice: Distorted probability weighting versus probability neglect. *Journal of Behavioral Decision Making*.

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).

Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice: A Reference-Dependent Model. *Quarterly Journal of Economics*, *106*(4), 1039-1061.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty, 5*(4), 297-323.

van Duijvenvoorde, A. C., Huizenga, H. M., Somerville, L. H., Delgado, M. R., Powers, A., Weeda, W. D., ... & Figner, B. (2015). Neural correlates of expected risks and returns in risky choice across development. *The Journal of Neuroscience*, *35*(4), 1549-1560.

van Ravenzwaaij, D., Dutilh, G., & Wagenmakers, E.-J. (2011). Cognitive model decomposition of the BART: assessment and application. *Journal of Mathematical Psychology, 55*(1), 94-105.

Neumann, J. V., & von Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press, Princeton, NJ (1944)

Wallsten, T. S., Pleskac, T. J., & Lejuez, C. (2005). Modeling behavior in a clinically diagnostic sequential risk-taking task. *Psychological Review, 112*(4), 862.

Weber, E. U., Blais, A. R., & Betz, N. E. (2002). A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making, 15*(4), 263-290.

Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annual Review of Psychology, 60*, 53-85.

Wichary, S., Pachur, T., & Li, M. (2015). Risk-taking tendencies in prisoners and nonprisoners: Does gender matter? *Journal of Behavioral Decision Making.*

Yechiam, E., Busemeyer, J. R., Stout, J. C., & Bechara, A. (2005). Using cognitive models to map relations between neuropsychological disorders and human decision-making deficits. *Psychological Science, 16*(12), 973-978.

Zuckerman, M. (2007). *Sensation seeking and risky behavior*. Washington, DC: American Psychological Association.

Zuckerman, M., Eysenck, S. B., & Eysenck, H. J. (1978). Sensation seeking in England and America: Cross-cultural, age, and sex comparisons. *Journal of Consulting and Clinical Psychology, 46*(1), 139-149.

Appendix A: Computational Models

Stochastic Choice

We begin the description of the applied models with the formalization of the choice behavior: The basic observed behavior in the CCT in a specific choice situation *i* is whether a participant turns over a card or ends the current game round and thus cashes the money accrued in the current game round. Several studies have shown that choice behavior is stochastic, with decision makers not always making the same choice in identical situations, and the strength of this inconsistency depending on the differences of the prospects' subjective values (Busemeyer, 1980; Luce & Suppes, 1965; Marley, 2002; Mosteller & Nogee, 2006; Rieskamp, 2008). We account for the stochastic nature of choices by applying an exponential choice rule, in which the probability of choosing between the subjective value of stopping $V_i(stop)$ versus the subjective value of taking a card $V_i(take)$ is modeled as a monotonic function of the differences between the prospects' subjective values:

$$p_i(take) = \frac{1}{1+e^{\theta(V_i(stop)-V_i(take))}} \tag{1}$$

Hence, it is assumed that a participant evaluates whether the subjective value of taking a card $V_i(take)$ is larger than the subjective value to stop at the i^{th} card, $V_i(stop)$. The free parameter θ captures the degree of stochasticity in the choices, with lower values for more stochastic choices.

Expectation-Based Models

We implemented two variants of how participants may set the reference points for their decisions. In the first version, the reference point is updated in each binary take-card versus stop decision (DU-reference point). Thus, the subjective value of stopping represents no deviation from the status quo and equals zero (see also van Duijvenvoorde et al., 2015).

DU-reference point: $\mathrm{V}_i(\mathrm{stop}) = 0$ (2)

In the second version, in which the reference point is updated at the beginning of every game round (RU-reference point), the subjective value of stopping equals the accrued amount of gains, each worth g up to card draw i. As such it follows:

RU-reference point: $\mathrm{V}_i(\text{stop}) = (ig)^{\alpha+}$ (3)

whereas the parameter α+ modulates the curvature of the subjective value function.

In the version with the DU-reference point, the subjective value of turning over a card is on the one side the possible gain g of the next card, minus the possible loss l if turning over a loss card.

DU-reference point: $\mathrm{V}_i(\text{take}) = \pi(p_i(win))\ g^{\alpha+} - \pi(1 - p_i(win))\lambda l^{\alpha-}$ (4)

The subjective value of taking a card $\mathrm{V}_i(\text{take})$ in the RU-reference point version reflects on the one side the possible gain of all gains accrued in the current game round i plus the next card's worth g, minus the possible loss l if turning over a loss card. CPT distinguishes between pure gain lotteries and mixed lotteries (Tversky & Kahneman, 1992). Thus, for choice situations in which the expected value of gains minus the expected value of a loss is greater than zero or equal zero, the subjective value of drawing another card is given by equation (5).

RU-reference point, pure gain lotteries:

$\mathrm{V}_i(\text{take}) = \ \pi(p_i(win))\big((i+1)g\ \big)^{\alpha+} +\ (1 - \pi(p_i(win)))\ ((ig)^{a+}\ - l^{\alpha-})$ (5)

If the expected value of gains minus the expected value of a loss is below zero, the lottery represents a mixed lottery and the subjective value is given by equation (6).

RU-reference point, mixed gain lotteries:

$\mathrm{V}_i(\text{take}) = \pi(p_i(win))((i+1)\ g)^{\alpha+} - \pi(1 - p_i(win))\lambda l^{\alpha-}$ (6)

In equations 4 to 6 *p(.)* refers to the probability weighting function, transforming objective probabilities into subjective probabilities. Probabilities are transformed by

$\pi(p_i) = e^{-\delta(-\ln p_i)^{\eta}}$ (Prelec, 1998) (7) with η specifying the inverse (for $\eta < 1$) s-shaped transformation of the weighting function. The free parameter δ accounts for the elevation of the weighting function and reflects a general underestimation (if $\delta > 1$) or overestimation (if $\delta < 1$) of probabilities, which can be seen as an indicator of how optimistically a prospect is valued.

In CPT it is assumed that losses have subjectively more impact on decisions than gains. This is implemented in the model by the loss aversion parameter λ in equation 4 and 6.

Finally, the probability to win $p_i(win)$ is determined by the number of loss cards n^l and the number of cards already turned over *(i-1)* in a game round:

$$p_i(win) = 1 - \frac{n^l}{32-(i-1)} \tag{8}$$

Likelihood Function

To fit the CPT-model to observed data and to estimate the parameter values that are most consistent with the decision makers' behavior, the parameters are connected to the data via the likelihood function. The likelihood function takes an identical structure as in other dynamic risk tasks such as the BART (Wallsten et al., 2005). For the fully specified CPT model, the probability of the data, $p(D|\alpha+, \alpha-, \eta+, \eta-, \delta+, \delta-, \lambda, \theta)$ for all trials n_k, and for all draws within each game round, $n_{i(k)}$, depends on the probability that the decision maker will turn over another card for each game round k for each opportunity i to take a card.

$$p(D|\alpha^+,\alpha^-,\eta^+,\eta^-,\delta^+,\delta^-,\lambda,\theta) = \prod_{k=1}^{n_k} \prod_{i=1}^{n_{i(k)}} p_{ki}^{take} \left(1 - p_{k,n_{i(k)}+1}^{take}\right)^{d_k} \tag{9}$$

where $d_k = 1$ if the decision maker stopped on opportunity i and $d_k = 0$ if the decision maker has turned over a loss card on trial k. This quantity is basically the product of all probabilities that the decision maker will turn over a card times one minus this probability on the occasions where the decision maker stopped.

Constant Probability Model

The constant probability model does not involve psychological processing but assumes that the participants ignore all information but turn over cards at a constant probability in all game rounds. Therefore, this model contains only one free parameter, $p_{ik}(\text{take})$, which is estimated from the data via maximum likelihood estimation for each participant as in equation 9 (cf. Wallsten et al., 2005).

Appendix B. Model Comparison

Table B1. Model Comparisons indicating the number of free parameters (df), the mean BIC and deviance (DEV), the number of participants for whom the respective model has the lowest BIC (nr. b. fit), the average rank of each model with respect to the BIC (rank; lower numbers indicate a better fit for more participants), as well as the explained variance (R^2).

	Model		df	BIC	DEV	nr. b. fit	rank	R^2
Decision-updated (DU) reference point	EV	θ	1	604.0	599.6	0	26.6	0.11
	EUT	α, θ	2	265.6	252.4	5	12.5	0.47
	CPT-1	$\alpha, \eta, \lambda, \theta$	4	216.9	199.2	73	4.3	0.47
	CPT-2	$\alpha, \eta, \delta, \lambda, \theta$	5	215.3	193.1	45	3.6	0.49
	CPT-3	$\alpha, \eta, \delta+, \delta-, \lambda, \theta$	6	217.4	190.9	4	6.0	0.50
	CPT-4	$\alpha, \eta+, \eta-, \lambda, \theta$	5	219.0	196.8	1	6.9	0.48
	CPT-5	$\alpha, \eta+, \eta-, \delta, \lambda, \theta$	6	217.6	191.0	3	6.4	0.49
	CPT-6	$\alpha, \eta+, \eta-, \delta+, \delta-, \lambda, \theta$	7	221.6	190.6	0	10.2	0.48
	CPT-7	$\alpha+, \alpha-, \eta, \lambda, \theta$	5	214.0	191.9	31	3.9	0.50
	CPT-8	$\alpha+, \alpha-, \eta, \delta, \lambda, \theta$	5	215.4	188.9	13	5.4	0.51
	CPT-9	$\alpha+, \alpha-, \eta, \delta+, \delta-, \lambda, \theta$	7	218.1	187.1	2	8.2	0.51
	CPT-10	$\alpha+, \alpha-, \eta+, \eta-, \lambda, \theta$	6	216.9	190.3	11	7.1	0.50
	CPT-11	$\alpha+, \alpha-, \eta+, \eta-, \delta, \lambda, \theta$	7	218.0	187.0	1	8.1	0.53
	CPT-12	$\alpha+, \alpha-, \eta+, \eta-, \delta+, \delta-, \lambda, \theta$	8	222.1	186.6	0	11.1	0.53
Round-updated (RU) reference point	EV	θ	1	611.3	606.9	0	27.6	0.09
	EUT	α, θ	2	580.8	567.5	1	26.0	0.10
	CPT1	$\alpha, \eta, \lambda, \theta$	4	470.3	452.5	0	20.3	0.23
	CPT2	$\alpha, \eta, \delta, \lambda, \theta$	5	455.5	433.4	1	15.2	0.22
	CPT3	$\alpha, \eta, \delta+, \delta-, \lambda, \theta$	6	457.4	430.8	0	16.7	0.23
	CPT4	$\alpha, \eta+, \eta-, \lambda, \theta$	5	474.4	452.2	0	22.6	0.23
	CPT5	$\alpha, \eta+, \eta-, \delta, \lambda, \theta$	6	459.7	433.1	0	18.8	0.24
	CPT6	$\alpha, \eta+, \eta-, \delta+, \delta-, \lambda, \theta$	7	461.5	430.5	0	20.6	0.22
	CPT7	$\alpha+, \alpha-, \eta, \lambda, \theta$	5	461.3	439.2	0	18.5	0.25
	CPT8	$\alpha+, \alpha-, \eta, \delta, \lambda, \theta$	5	457.3	430.8	0	16.5	0.23
	CPT9	$\alpha+, \alpha-, \eta, \delta+, \delta-, \lambda, \theta$	7	459.8	428.8	0	18.8	0.23
	CPT10	$\alpha+, \alpha-, \eta+, \eta-, \lambda, \theta$	6	465.1	438.5	0	21.2	0.26
	CPT11	$\alpha+, \alpha-, \eta+, \eta-, \delta, \lambda, \theta$	7	461.4	430.4	0	20.6	0.22
	CPT12	$\alpha+, \alpha-, \eta+, \eta-, \delta+, \delta-, \lambda, \theta$	8	464.0	428.5	0	22.2	0.22

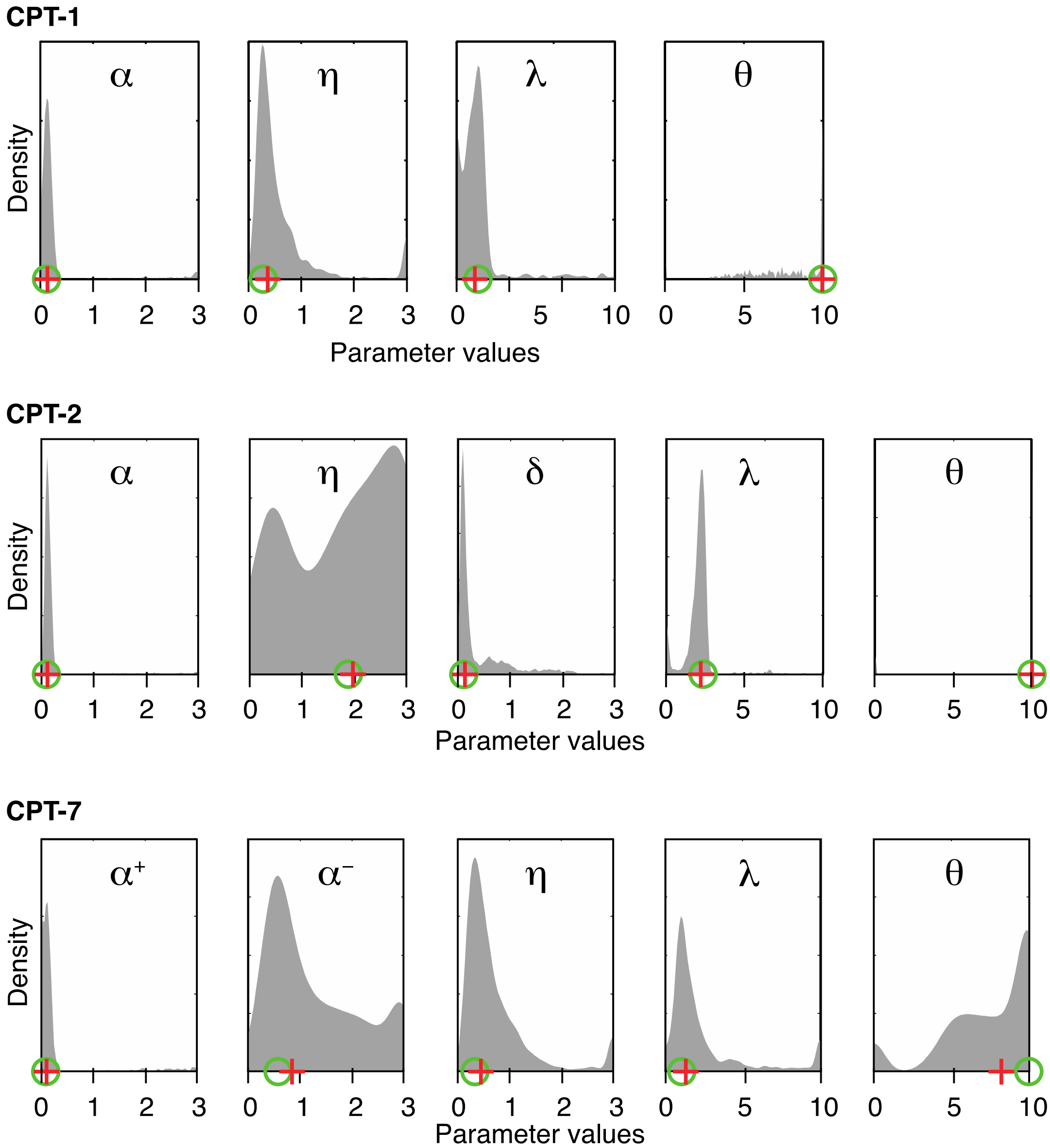

Figure B1. Parameter recovery of the models with the lowest mean BIC rank. The green circle depicts the input (median parameter estimate across participants for the respective model) for the recovery; the red cross indicates the median of the re-estimated parameters. The estimated parameters of the CPT-1-model are reasonably well recovered (i.e., unbiased and with low variance). In the CPT-2 model the parameter estimates are unbiased but η is estimated with a large variance, spanning the whole search space. In the CPT-7 model, α^- and θ are biased and show substantial variance.

Appendix C

Table C1. Experimental design of the game rounds. R = game round, G = gain amount, L = loss amount, NLC = number of loss cards.

R	G	L	NLC	R	G	L	NLC	R	G	L	NLC	R	G	L	NLC
1	150	75	20	22	10	500	1	43	20	750	1	64	400	100	20
2	50	100	1	23	300	150	20	44	300	50	20	65	50	250	1
3	200	100	10	24	20	100	1	45	150	100	10	66	20	500	3
4	200	50	28	25	150	50	16	46	150	100	10	67	600	50	28
5	20	750	1	26	200	50	20	47	300	100	20	68	20	100	3
6	300	100	16	27	50	100	3	48	10	750	3	69	20	500	3
7	10	500	3	28	300	100	20	49	50	750	1	70	150	50	20
8	10	250	1	29	20	500	1	50	20	750	3	71	75	50	10
9	50	750	3	30	10	250	1	51	600	150	28	72	200	100	20
10	50	250	1	31	100	50	10	52	50	100	1	73	150	50	20
11	300	100	16	32	10	500	3	53	20	250	3	74	20	250	3
12	600	300	28	33	100	50	20	54	50	750	1	75	600	50	28
13	10	250	3	34	20	250	1	55	20	100	1	76	75	50	16
14	50	500	3	35	20	250	1	56	10	500	1	77	10	100	1
15	10	250	3	36	150	50	16	57	10	100	3	78	50	250	3
16	10	750	1	37	20	100	3	58	300	50	28	79	50	500	1
17	20	750	3	38	10	750	1	59	150	100	16	80	10	100	1
18	10	750	3	39	100	50	10	60	75	50	16	81	150	100	16
19	75	50	10	40	50	750	3	61	200	100	10	82	50	500	1
20	150	25	20	41	50	500	3	62	600	100	28	83	100	50	28
21	10	100	3	42	50	100	3	63	20	500	1	84	50	250	3