Article:

Konstantinidis, E orcid.org/0000-0002-4782-0749, van Ravenzwaaij, D, Güney, Ş et al. (1 more author) (2020) Now for sure or later with a risk? Modeling risky intertemporal choice as accumulated preference. Decision, 7 (2). pp. 91-120. ISSN 2325-9965

https://doi.org/10.1037/dec0000103

Now for sure or later with a risk? Modeling risky inter-temporal choice as accumulated preference

Emmanouil Konstantinidis
University of Leeds and University of New South Wales

Don van Ravenzwaaij
University of Newcastle and University of Groningen

Şule Güney
University of New South Wales and University of Southern California

Ben R. Newell
University of New South Wales

Abstract

Research on risky and inter–temporal decision–making often focuses on descriptive models of choice. This class of models sometimes lack a psychological process account of how different cognitive processes give rise to choice behavior. Here, we attempt to decompose these processes using sequential accumulator modeling (i.e., the Linear Ballistic Accumulator model; LBA). Participants were presented with pairs of gambles that either involve different levels of probability or delay (Experiment 1) or a combination of these dimensions (both probability and delay; Experiment 2). Response times (RTs) were recorded as a measure of preferential strength. We then combined choice data and response times, and utilized variants of the LBA to explore different assumptions about how preferences are formed. Specifically we show that a model which allows for the subjective evaluation of a fixed now/certain option to change as a function of the delayed/risky option with which it is paired provides the best account of the combined choice and RT data. The work highlights the advantages of using cognitive process models in risky and inter–temporal choice, and points towards a common framework for understanding how people evaluate time and probability.

Keywords: Inter–temporal Choice; Risky Choice; Evidence Accumulation; Cognitive Modeling; Linear Ballistic Accumulator

Introduction

You have just won the lottery and the prize is $10,000. Do you use your money now, or do you put it in a bank account, for one year, and then take out $11,500? This choice is an example of an *inter–temporal choice*, it involves tradeoffs between sooner–smaller (SS) and larger–later (LL) options. Consider a second dilemma. You can either choose to keep the prize money in a savings account for a *certain* (probability of 1) return of $1,500 or you can take a trip to the casino and put the $10,000 on your lucky number 17 on the roulette table. This is an example of a *risky choice*, it involves tradeoffs between certain and risky but more valuable options. Most studies of inter–temporal and risky choice have employed context-free monetary choice dyads between SS and LL options on the one hand, for example, a choice between $10 now or $15 in two months (e.g., Chapman & Weber, 2006; Loewenstein & Prelec, 1992), and between certain and risky options on the other hand, for example, a choice between $30 for sure or $40 with 80% chance or nothing otherwise (e.g., Kahneman & Tversky, 1979).

Two hallmarks of traditional research on inter–temporal and risky choice are i) examination of the two types of choice in isolation, and ii) evaluation of preferences in terms of their coherence (or lack thereof) with normative economic principles. This large body of work has revealed key insights into the types of factors that affect risky or inter-temporal choice, but "the interaction between risk and delay is complex and not easily understood" (B. J. Weber & Chapman, 2005, p. 104).

In the domain of inter–temporal choice, the dominant approach is to examine whether choices across time adhere to Discounted Utility Theory (DUT; Samuelson, 1937). DUT implies that decision makers maximize a weighted sum of utilities with exponentially declining discount weights. In the domain of risky choice, research has focused on Expected Utility Theory (EUT). EUT views decision makers as maximizing a weighted sum of utilities with their probabilities of occurrence (e.g., Epper, Fehr-Duda, & Bruhin, 2011; Prelec & Loewenstein, 1991).

DUT and EUT are normative models of choice; they provide principles according to which rational decision makers should behave (Newell, Lagnado, & Shanks, 2015). However, extensive research has documented several violations of these principles (e.g., Allais, 1953; Thaler, 1981). The standard approach to account for these violations is to modify the theories but to retain their core constituents. Thus for inter–temporal choice, hyperbolic functions which allow decreasing discount rates rather than constant (i.e., exponential) rates are used to capture observed choice "anomalies" (e.g., Green & Myerson, 2004). For risky choice, allowing a non–linear probability weighting function provides explanations of commonly observed behavioral effects and preference reversals (e.g., Kahneman & Tversky, 1979; Tversky & Kahneman, 1992).

These modified models are descriptive: they provide a description not of how decision makers should behave but how they appear to behave when confronted with such choices (at least at the

---

Emmanouil Konstantinidis, Centre for Decision Research, University of Leeds, Leeds, England and School of Psychology, University of New South Wales, Sydney, Australia; Don van Ravenzwaaij, Department of Psychology, University of Groningen, Groningen, The Netherlands; Şule Güney, University of Southern California, Los Angeles, CA, USA; Ben R. Newell, School of Psychology, University of New South Wales, Sydney, Australia

This research was supported by an Australian Research Council Discovery Project Grant (DP140101145) awarded to the second and last authors.

Part of the computational modeling work was undertaken on ARC3, part of the High Performance Computing facilities at the University of Leeds, England.

Correspondence concerning this article should be addressed to Emmanouil Konstantinidis, Centre for Decision Research, University of Leeds, Leeds, LS2 9JT, England. E-mail: em.konstantinidis@gmail.com.

aggregate level, cf. Epper et al., 2011). Such models (e.g., Cumulative Prospect Theory for risky choice and Hyperbolic Discounting for inter–temporal choice) are utility–based models: a utility (or subjective value) is calculated for each option, and the option with the highest utility is chosen. However, what these models lack is a psychological process account of why choices are better fit by hyperbolic than exponential functions, or by non–linear than linear weighting functions (cf. Stewart, Chater, & Brown, 2006). In other words, these models do not explain how the utility of each option is estimated and the psychological processes that are involved. To answer this "how" question requires the development of *cognitive process models* which specify the components and relations between the (thought) processes engaged when making such choices (e.g., Appelt, Hardisty, & Weber, 2011; Brandstätter, Gigerenzer, & Hertwig, 2006; Shafir, Simonson, & Tversky, 1993; Vlaev, Chater, Stewart, & Brown, 2011; E. U. Weber et al., 2007).

In the field of speeded multi–alternative forced choice decision–making, such cognitive process models have been in use for almost four decades (e.g., Ratcliff, 1978; Brown & Heathcote, 2008). The cognitive models of choice in the field of response time (RT) research are called sequential accumulator models. Among others, these models have been successfully applied to experiments on perceptual discrimination, letter identification, lexical decision, categorization, recognition memory, and signal detection (e.g., Ratcliff, 1978; Ratcliff, Gomez, & McKoon, 2004; Ratcliff, Thapar, & McKoon, 2006, 2010; van Ravenzwaaij, Dutilh, & Wagenmakers, 2012; van Ravenzwaaij, van der Maas, & Wagenmakers, 2011; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008). Evidence accumulation models such as decision field theory (DFT; Busemeyer & Townsend, 1993) and the leaky competing accumulator (LCA; Usher & McClelland, 2001) have been applied in the domain of risky choice, and the domain of inter–temporal choice (see Dai & Busemeyer, 2014; Rodriguez, Turner, & McClure, 2014).

One of the advantages of such a modeling approach is that it allows researchers to decompose observed RTs and choice proportions into latent psychological processes such as speed of cognitive processing, response caution, and non–decision time. More traditional analyses make no attempt to explain the observed data by means of a psychologically plausible process model.

One key difference between inter–temporal and risky choice on the one hand and traditional RT research on the other is that in the former field, decisions are rarely timed (but see e.g., Dai & Busemeyer, 2014; Rodriguez et al., 2014). Response times could potentially teach us a lot about people's preferences. For instance, consider on the one hand the choice between receiving \$100 now or \$10,000 in two months and on the other hand the choice between receiving \$100 now or \$150 in two months. In both instances, you might prefer the larger–later option. As a result, your choice preference will look the same. However, the first choice was most likely a much easier one to make than the second one. Your *strength of preference* in the first choice in favor of the larger–later option was likely much larger. This strength of preference would likely be reflected in a lower RT compared to the RT associated with your decision for the second choice.

Another difference between inter–temporal and risky choice and traditional RT research is that the response options are about preference and as such, there are no correct answers. This presents a difference on the conceptual level, but on the model level, there are no obstacles as evident by some recent applications of sequential accumulator models to preference data (see e.g., Dai & Busemeyer, 2014; Hawkins et al., 2014; Rodriguez et al., 2014; Trueblood, Brown, & Heathcote, 2014).

In this paper, we present data from two experiments that contained both inter–temporal and risky choices. In Experiment 1 participants had to make a choice between a small-sooner ("now") and a larger-later option for the inter–temporal choice trials, and a choice between a certain-smaller and a risky-larger option for the risky choice trials. In Experiment 2 the inter–temporal and risky components were combined within trials: participants had to make a choice between a certain-now-smaller and a risky-later-larger option. In both experiments the now/certain options remained constant across choice trials: $100 with certainty, right now. We pressed our participants to make their responses as quickly as possible while still being able to show their true preferences. We then applied a cognitive process model to the data.

One of the main objectives of the present analysis is to uncover the potential for a shared component that drives decision–making in both inter–temporal and risky choices. Previous research has identified several parallels and similarities between probability and delay. For example, Chapman and Weber (2006) examined whether two well-documented biases in risky and inter–temporal choice (the *common ratio* effect, and the *common difference* effect, respectively) can be accounted for by the same underlying mechanism. Other studies have found evidence for psychological equivalence between probability and delay, suggesting that probability can be translated or treated as delay (e.g., Rachlin, Raineri, & Cross, 1991; Yi, de la Piedad, & Bickel, 2006), and vice-versa, that delay can be treated as uncertainty (e.g., Baucells & Heukamp, 2010; Keren & Roelofsma, 1995). In addition, recent theoretical and modeling attempts have assumed similar functional forms for delay discounting (i.e., decrease of a reward value with increases in delay) and probability discounting (i.e., decrease of a reward value with decreases in probability). For instance, Vanderveldt, Green, and Myerson (2015) observed that a hyperboloid function of delay and probability discounting can describe discounting patterns in both domains. While their hyperboloid model provides an excellent fit to the choice data, it is a descriptive "as–if" model, in the sense that it does not account for the underlying thought processes that drive preferential choice. With our cognitive process model and the simultaneous examination of choices and RTs (i.e., strength of preference), we take the idea of parallelism in delay and probability discounting one step further by exploring the potential for a unifying psychological process that governs preferences in both inter–temporal and risky choices.

The model also allowed us to test to what extent the absolute attractiveness of the now/certain options (which was always an immediate $100 with certainty in our experiments) change with variations in the delayed/risky choice options with which it was paired. Such a test is difficult with behavioral data or descriptive models of choice that typically only provide insight into the relative attractiveness of two presented choice options.[1] Following this, our model constitutes

[1] Descriptive models also provide measures of absolute attractiveness for each option, i.e., the subjective utility, but it is the relative attractiveness that defines preferences and choice.

a natural test of integrative expected and subjective expected utility-based models which assume an overall fixed utility for each option irrespective of the context or the alternative to which it is compared (see also Brandstätter et al., 2006).

The remainder of this paper is organized as follows: In the next section, we will introduce our sequential accumulator model of choice: the Linear Ballistic Accumulator model (LBA; Brown & Heathcote, 2008). We will then describe our experiments in detail. Next, we discuss the behavioral results and then the modeling results. We conclude with a discussion of the gains of our cognitive modeling approach and the lessons learned about the shared nature of delay and probability.

The Linear Ballistic Accumulator model

In the LBA for multi–alternative RT tasks (Brown & Heathcote, 2008), the decision–making process is conceptualized as the accumulation of information over time. A response is initiated when the accumulated evidence reaches a predefined threshold. An illustration for an inter–temporal choice with two response options is given in Figure 1.

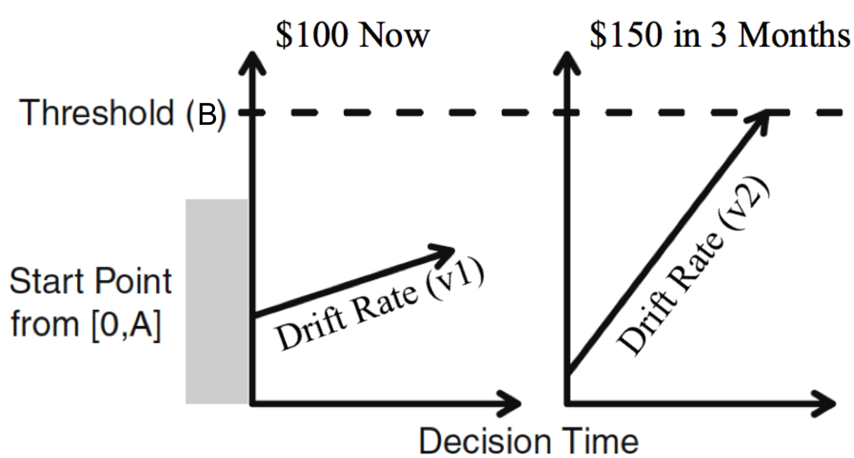

Figure 1. The LBA and its parameters for an inter–temporal choice with two response options. Evidence accumulation begins at start point k, drawn randomly from a uniform distribution with interval $[0, A]$. Evidence accumulation is governed by drift rate d, drawn across trials from a normal distribution with mean ν and standard deviation s. A response is given as soon as one accumulator reaches threshold B. Observed RT is an additive combination of the time during which evidence is accumulated and non–decision time $t0$.

The LBA assumes that the decision process starts from a random point between 0 and A, after which information is accumulated linearly for each response option. The rate of this evidence accumulation is determined by drift rates d_1 and d_2, normally distributed over trials with means ν_1 and ν_2, and standard deviation s, which we assume here to be equal for both accumulators. For this application, drift rates are truncated at zero to prevent negative accumulation rates. Threshold b determines the speed–accuracy tradeoff; lowering b leads to faster RTs at the cost of a higher error rate (but see Rae, Heathcote, Donkin, Averell, & Brown, 2014). The distance between threshold b

and the maximum start point A is quantified by B, such that $b = A + B$.

Together, these parameters generate a distribution of decision times DT. The observed RT, however, also consists of stimulus–nonspecific components such as stimulus encoding, response preparation and motor execution, which together make up non–decision time $t0$. The model assumes that $t0$ simply shifts the distribution of DT, such that $RT = DT + t_0$ (Luce, 1986). Hence, the three key components of the LBA are (1) the speed of information processing, quantified by mean drift rate ν; (2) response caution, quantified by distance from start point to threshold that averages at $b - A/2$; and (3) non–decision time, quantified by $t0$.

The LBA has been applied to a number of perceptual discrimination paradigms (e.g., Cassey, Gaut, Steyvers, & Brown, 2015; Cassey, Heathcote, & Brown, 2014; Forstmann et al., 2008, 2010; Ho, Brown, & Serences, 2009; van Ravenzwaaij, Provost, & Brown, 2017). Recently, the LBA has also been applied to preference data. For instance, Hawkins et al. (2014) applied the LBA to consumer preference data toward mobile phones. In an adaptation of the LBA developed by Trueblood et al. (2014), the model was fit to preference data of three kinds of context effects: similarity, compromise, and attraction. Rodriguez et al. (2014) applied the LBA to inter–temporal choice data and concluded that "perceptual and value–based decision–making may depend on similar comparison and selection processes" (p. 7).

The interpretation of the drift rate parameter changes when applying sequential accumulator models to preference data without an inherent correct answer. Rather than speed of information processing, drift rate reflects the strength of preference for a choice option. For this application, we define drift rates as representing a weighted sum of an option's attribute values (amount, delay, and probability). In other words, each attribute's contribution to the strength of preference depends on the importance or attention placed on each attribute, quantified by scaling parameters (or weights; see the Model Implementation section for more details). This definition of preferential strength allows us to test three specific accounts of how choice preferences vary with different levels of delay and probability.

Specifically, we examine how preferences for the now/certain options are formed in relation to the values of the choice attributes of the delayed/risky options. The first account ("proportional") assumes that the value of the now/certain option changes proportionally to different alternatives for the delayed/risky option. The choice attributes (amount, delay, and probability) in the now/certain and delayed/risky options have different weights, suggesting that the importance or attention placed on each attribute varies between the two options. On the other hand, the "invariant" account simply assumes that the absolute value (preferential strength) of the now/certain option remains constant across all choice trials, irrespective of the attribute values of the delayed/risky option. Consequently, this model suggests that a single absolute value for the now/certain option is estimated (that is, a single drift rate across all trials – no scaling parameters or weights needed). This "invariant" account resembles expected utility-based models, which assume a single fixed value for an option regardless of the context (i.e., alternative options) in which a decision is made. The last account ("symmetrical") presents a compromise between the two aforementioned "extreme" accounts: while the value of the now/certain option does not remain constant, the choice attribute weights are identical between the two options. For each model account, we examined linear and non-linear (i.e., power transformations of each attribute's values) functional forms for the drift rate ν, and different assumptions relating to the upper starting point A. We formally describe these variants of the LBA in the Model Implementation section below.

We fit the LBA to inter–temporal and risky choice data simultaneously. Thus, our work presents the first attempt to examine the potential for a unifying underlying process that governs preferences in both inter–temporal and risky choices.

Experiment 1

We set out to model people's preferences on inter–temporal and risky choices separately. Participants completed a task with two separate blocks of inter–temporal choice and risky choice trials. We report the behavioral results of this experiment, as well as the modeling results provided by fitting the LBA.

Method

Participants. Forty undergraduate students (26 female; $M_{age} = 19.40$) at the University of New South Wales participated in return for course credit. For each participant, one of their preferences from the risky choice trials was randomly selected. If the participant preferred the risky option in that specific choice trial, the gamble was played for real (e.g., $200 with 50% chance). In case of a win, the participant was paid 2% of the amount (e.g., $4 as 2% of $200) and nothing otherwise. Those who preferred the sure option were paid $2 (i.e., 2% of $100 for sure)[2].

Materials. The experiment consisted of 380 inter–temporal choice and 380 risky choice trials. For the inter–temporal choices, participants had to indicate what they preferred: $100 now or \$$X$ in D months, with \$$X$ varying from $120 to $500 in $20 increments (for a total of 20 amounts) and D varying from 2 months to 38 months in 2 month increments (for a total of 19 delays). Thus, every combination of amount and delay was presented to participants once as an alternative to $100 now ($D = 0$).

For the risky choices, participants had to indicate what they preferred: $100 for sure or \$$X$ with P% chance, with \$$X$ varying from $120 to $500 in $20 increments (for a total of 20 amounts) and P varying from 5% to 95% in 5% increments (for a total of 19 probabilities). Thus, every combination of amount and probability was presented to participants once as an alternative to $100 for sure ($P = 100\%$).

Procedure. Participants completed the experiment in two sessions, of 380 trials each (190 risky choice trials and 190 inter–temporal choice trials). Within a session the order of the trials was blocked (i.e., all risky together, all inter–temporal together) and counterbalanced. The two

[2]At the outset, participants were told that one trial from the experiment would be selected and the gamble would be played for real. To facilitate payment, the (pseudo)randomly selected trial always came from the risky choice trials, but participants did not know this, thus creating the impression of equal incentives for both phases of the experiment.

sessions were separated by a minimum of three hours (i.e., some participants completed the sessions on consecutive days, others in the morning and afternoon of the same day).

Implementation of the Model

We used a hierarchical Bayesian implementation of the LBA (Turner, Sederberg, Brown, & Steyvers, 2013). Advantages of the hierarchical Bayesian framework include the ability to fit the LBA to data with relatively few trials, because the model borrows strength from the hierarchical structure. This advantage is important, as we are working with a task for which we essentially have only a single trial per participant for each type of choice (one single combination of $\$X$ and D for inter–temporal choices and a combination of $\$X$ and $P\%$ for risky choices). The Bayesian set–up allows for using Markov chain Monte-Carlo (MCMC) sampling, which is an efficient way of finding the optimal set of parameters (Gamerman & Lopes, 2006; Gilks, Richardson, & Spiegelhalter, 1996; van Ravenzwaaij, Cassey, & Brown, 2018).

We fit three different model accounts that differed on how they model the choice process: proportional, symmetrical, and invariant. Within each model account we examined four different assumptions (variants) relating to the starting point parameters and the functional form of the drift rate. Specifically, the “2A” model-variant assumed two parameters for the upper range of starting point A for each type of choice (i.e., A_I for inter–temporal and A_R for risky choices), whereas the “4A” variant had four starting point parameters for each choice option in the task (i.e., in the inter–temporal choice task A_{I_N} and A_{I_D} for the now and delayed choice options, respectively, and in the risky choice task A_{R_C} and A_{R_R} for the certain and risky choice options, respectively). The “4A” variants assume that a response bias is associated with every choice option, indicating that for some choice options less evidence might be required to reach a decision. For the drift rates, we examined linear and nonlinear versions. Thus, for each model account we fit 4 model-variants: linear-2A, linear-4A, nonlinear-2A, and nonlinear-4A. In total we fit 12 models. For all models we assumed two parameters for threshold B (B_I for inter–temporal and B_R for risky choices). We first describe the “proportional” model, then describe the two other model accounts by referring to changes to the “proportional” model.

For the inter–temporal choice task, the linear “proportional” model (both the 2A and 4A variants) included drift rates for the “now” and “delayed” options as follows:

$$
\begin{aligned}
\nu_N &= \nu_{N_0} - \alpha_{N_X} \times (\$X/20 - 6) - \alpha_{N_D} \times (19 - D/2) \\
\nu_D &= \nu_{D_0} - \alpha_{D_X} \times (25 - \$X/20) - \alpha_{D_D} \times (D/2 - 1)
\end{aligned}
\tag{1}
$$

where ν_N and ν_D denote drift rates for the now and the delayed choice options, ν_{N_0} and ν_{D_0} denote offset parameters for the now and the delayed choice options, $\$X$ denotes the amount in dollars for the delayed choice option, D denotes delay in months for the delayed choice option, α_{N_X} and α_{D_X} are amount scale parameters for the now and the delayed choice options, and α_{N_D} and α_{D_D} are delay scale parameters for the now and the delayed choice options. $\nu_N = \nu_{N_0}$ if $\$X = 120$ and $D = 38$ (the option that most favors the “now” choice). $\nu_D = \nu_{D_0}$ if $\$X = 500$ and $D = 2$ (the option that most favors the "delayed" choice). For the nonlinear version (both the $2A$ and $4A$ variants), we applied a power transformation (β parameters), on the numerical values of amount and delay of each option, thus adding two extra parameters to be estimated (same β for amount and delay across the two options).

For the risky choice task, the linear "proportional" model (both the $2A$ and $4A$ variants) included drift rates for the "certain" and "risky" options as follows:

$$\begin{aligned} \nu_C &= \nu_{C_0} - \alpha_{C_X} \times (\$X/20 - 6) - \alpha_{C_P} \times (P/5 - 1) \\ \nu_R &= \nu_{R_0} - \alpha_{R_X} \times (25 - \$X/20) - \alpha_{R_P} \times (19 - P/5) \end{aligned} \tag{2}$$

where ν_C and ν_R denote drift rates for the certain and risky choice options, ν_{C_0} and ν_{R_0} denote offset parameters for the certain and risky choice options, $\$X$ denotes the amount in dollars for the risky choice option, P denotes the payout probability for the risky choice option, α_{C_X} and α_{R_X} are amount scale parameters for the certain and the risky choice options, and α_{C_P} and α_{R_P} are probability scale parameters for the certain and the risky choice options. $\nu_C = \nu_{C_0}$ if $\$X = 120$ and $P\% = 5$ (the option that most favors the certain option). $\nu_R = \nu_{R_0}$ if $\$X = 500$ and $P\% = 95$ (the option that most favors the risky option)[3]. As in the specification for the inter–temporal choice task, the nonlinear drift rates for the risky choice task included power transformations (β parameters) of each option's amount values. Unlike amount and delay, we did not apply a power transformation for probability (i.e., $\beta = 1$).

We fit two other models that test specific assumptions about the underlying choice process. The first model ("invariant"), estimates a single ν_N parameter (i.e., drift rate for the now option) for all inter–temporal choice trials and a single ν_C parameter (i.e., drift rate for the certain option) for all risky choice trials. Conceptually, this simpler model assumes that the absolute value of the now/certain option does not change with different alternatives for the delayed/risky option. Essentially, the objectively invariant option would also be perceived as invariant by the decision makers.

The final model ("symmetrical") presents a compromise between the "proportional" model and the "invariant". The model assumes that $\alpha_{N_X} = \alpha_{D_X}$, $\alpha_{N_D} = \alpha_{D_D}$, $\alpha_{C_X} = \alpha_{R_X}$, and $\alpha_{C_P} = \alpha_{R_P}$. This means that contrary to the "invariant" model, the ν_N parameter and the ν_C parameter are not fixed to a single value. Instead, drift rates for the now/certain option vary symmetrically (though in the opposite direction) with drift rates for the delayed/risky option. Table 1 presents the 12 models that we fit (3 model accounts $\times$ 4 variants) and their associated parameters.

The comparison of the "proportional", the "invariant", and the "symmetrical" models will teach us something about the change in subjective evaluation of the now/certain choice option. Is the subjective evaluation of the now/certain choice option fixed irrespective of the delayed/risky choice option, does the subjective evaluation of the now/certain choice option vary symmetrically with the delayed/risky choice option, or does the now/certain choice option vary non-symmetrically but proportionally with the delayed/risky choice option? In addition, is the linear form of the drift

[3]Note that $t0$ is fixed to be the same for the inter–temporal and the risky choice task.

Table 1

Outline of the 12 models (3 model accounts: Proportional, Invariant, and Symmetrical × 4 variants: Linear–2A, Linear–4A, Nonlinear–2A, and Nonlinear–4A) and their parameters that were fit to the data of Experiment 1. In addition to the parameters listed below, all models include a single $t0$ *(non–decision time) and two threshold parameters* B_I *and* B_R *for the inter–temporal and risky choice tasks, respectively.*

Model	P	Task	Parameters
Proportional			
Linear–2A	17	I	A_I, ν_{N_0}, α_{N_X}, α_{N_D}, ν_{D_0}, α_{D_X}, α_{D_D}
		R	A_R, ν_{C_0}, α_{C_X}, α_{C_P}, ν_{R_0}, α_{R_X}, α_{R_P}
Nonlinear–2A	20	I	A_I, ν_{N_0}, α_{N_X}, α_{N_D}, β_{I_X}, β_{I_D}, ν_{D_0}, α_{D_X}, α_{D_D}
		R	A_R, ν_{C_0}, α_{C_X}, α_{C_P}, β_{R_X}, ν_{R_0}, α_{R_X}, α_{R_P}
Linear–4A	19	I	A_{I_N}, A_{I_D}, ν_{N_0}, α_{N_X}, α_{N_D}, ν_{D_0}, α_{D_X}, α_{D_D}
		R	A_{R_C}, A_{R_R}, ν_{C_0}, α_{C_X}, α_{C_P}, ν_{R_0}, α_{R_X}, α_{R_P}
Nonlinear–4A	22	I	A_{I_N}, A_{I_D}, ν_{N_0}, α_{N_X}, α_{N_D}, β_{I_X}, β_{I_D}, ν_{D_0}, α_{D_X}, α_{D_D}
		R	A_{R_C}, A_{R_R}, ν_{C_0}, α_{C_X}, α_{C_P}, β_{R_X}, ν_{R_0}, α_{R_X}, α_{R_P}
Invariant			
Linear–2A	13	I	A_I, ν_N, ν_{D_0}, α_{D_X}, α_{D_D}
		R	A_R, ν_C, ν_{R_0}, α_{R_X}, α_{R_P}
Nonlinear–2A	16	I	A_I, ν_N, ν_{D_0}, α_{D_X}, α_{D_D}, β_{D_X}, β_{D_D}
		R	A_R, ν_C, ν_{R_0}, α_{R_X}, α_{R_P}, β_{R_X}, β_{R_P}
Linear–4A	15	I	A_{I_N}, A_{I_D}, ν_N, ν_{D_0}, α_{D_X}, α_{D_D}
		R	A_{R_C}, A_{R_R}, ν_C, ν_{R_0}, α_{R_X}, α_{R_P}
Nonlinear–4A	18	I	A_{I_N}, A_{I_D}, ν_N, ν_{D_0}, α_{D_X}, α_{D_D}, β_{D_X}, β_{D_D}
		R	A_{R_C}, A_{R_R}, ν_C, ν_{R_0}, α_{R_X}, α_{R_P}, β_{R_X}
Symmetrical			
Linear–2A	13	I	A_I, ν_{N_0}, ν_{D_0}, α_{I_X}, α_{I_D}
		R	A_R, ν_{C_0}, ν_{R_0}, α_{R_X}, α_{R_P}
Nonlinear–2A	16	I	A_I, ν_{N_0}, ν_{D_0}, α_{I_X}, α_{I_D}, β_{I_X}, β_{I_D}
		R	A_R, ν_{C_0}, ν_{R_0}, α_{R_X}, α_{R_P}, β_{R_X}
Linear–4A	15	I	A_{I_N}, A_{I_D}, ν_{N_0}, ν_{D_0}, α_{I_X}, α_{I_D}
		R	A_{R_C}, A_{R_R}, ν_{C_0}, ν_{R_0}, α_{R_X}, α_{R_P}
Nonlinear–4A	18	I	A_{I_N}, A_{I_D}, ν_{N_0}, ν_{D_0}, α_{I_X}, α_{I_D}, β_{I_X}, β_{I_D}
		R	A_{R_C}, A_{R_R}, ν_{C_0}, ν_{R_0}, α_{R_X}, α_{R_P}, β_{R_X}

Note: P = Number of free parameters per participant; I = Inter–temporal choice task; R = Risky choice task.

rate sufficient to explain how preference is accumulated, or can more complex nonlinear relationships provide a better account for strength of preference? Do we need separate starting points for each choice option in the task ($4A$ models) to account for a priori biases for either the now/delayed or the certain/risky options, or two starting points for the inter–temporal choice and risky choice tasks ($2A$ models)? We will use formal model comparison to find the account best supported by the data. Details on starting values, prior distributions, and number of iterations may be found in the Appendix.

Subjective value and discounted utility models. To compare the performance of the LBA model against standard approaches in risky and inter–temporal choice, we fit a Prospect Theory model to the risky choice data (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992), and a hyperbolic discounting model to the inter–temporal choice data (Myerson & Green, 1995). For the risky choice data, the subjective value V of a risky prospect is given by

$$V = \sum w(p_i)u(X_i) \tag{3}$$

where $w(p_i)$ is the decision weight (transformed value of objective probability p, as produced by a probability weighting function; see Equation 5) and u is the utility of receiving reward X. For utility u, we assume a power utility function

$$u(X) = X^{\alpha} \tag{4}$$

For the probability weighting function, we used the two-parameter version proposed by Gonzalez and Wu (1999):

$$w(p) = \frac{\delta p^{\gamma}}{\delta p^{\gamma} + (1-p)^{\gamma}} \tag{5}$$

where γ represents the curvature and δ represents the elevation of the weighting function.

Similarly, for the inter–temporal choice data, the discounted value V of a delayed option is

$$V = \sum D(t_i)u(X_i) \tag{6}$$

where D represents the discount function. We used a two-parameter variant of hyperbolic discounting as below:

$$D(t) = \frac{1}{(1+kt)^{s}} \tag{7}$$

where k is the discounting rate and s governs the curvature of the hyperbola (Myerson & Green, 1995). For utility, we used the same power function as for the risky choice task (Equation 4).

Both models assume deterministic choice, that is, the choice option with the highest expected or discounted value is selected. Here we assumed a probabilistic choice rule, where the probability of choosing the safe option over the risky option ($p(S)$), or the probability of choosing the now option over the delayed ($p(N)$) is given by the softmax rule:

$$p(S) = \frac{\exp\{\theta V(S)\}}{\exp\{\theta V(S)\} + \exp\{\theta V(R)\}} \tag{8}$$

where θ denotes the sensitivity (inverse-temperature) parameter, indicating the degree to which choice probabilities adhere to numerical differences between $V(S)$ and $V(R)$ in risky choice, and $V(N)$ and $V(D)$ for inter–temporal choice. To find the best fitting parameters, we used maximum likelihood estimation techniques. We fit both models to the risky and inter–temporal choice data simultaneously for each participant (6 parameters in total: $\alpha, \gamma, \delta, k, s, \theta$). The procedure was a combination of grid search (300 different starting points for each set of parameters) and Nelder-Mead simplex methods (Nelder & Mead, 1965).

Behavioral Results

Choice. All participants completed the experiment[4]. Two participants were excluded from analysis: one participant had chosen the "delayed" option for every single choice, and another participant had seemingly responded randomly, producing responses that seemed largely inconsistent when compared against one another. Also, we excluded extreme RTs that were slower than 7s and faster than 250ms (2.8% of all trials). Preference data for the inter-temporal choice trials can be found in the top-left panel of Figure 2. The figure shows group average proportion data. Proportions close to 0, displayed in blue, indicate a uniform preference for the $100 now option. Proportions close to 1, displayed in yellow, indicate a uniform preference for the delayed option. The results show that participants prefer the now option when the delayed option does not pay very well (i.e., amounts not much higher than $120) or when the delay is long (i.e., close to 38 months). In contrast, participants prefer the delayed option when the delayed option pays well (i.e., amounts close to $500) or when the delay is short (i.e., close to now).

To examine the factors affecting choice of the smaller-sooner (SS; coded 0) or larger-later options (LL; coded 1) in the inter–temporal choice trials, we performed a mixed–effects logistic regression with amount and delay of the LL option as fixed effects (centered and scaled) and participant-specific random intercepts and slopes (for amount and delay). As expected, there was a significant positive effect of amount ($b = 2.13, z = 8.940, p < .001$), indicating that as amount offered by the LL option increased, so did the likelihood of selecting the delayed option. In line with the observations from Figure 2, as delay increased participants were more likely to select the SS ("now") option ($b = -3.05, z = -3.06, p < .001$).

Preference data for the risky choice trials can be found in the bottom-left panel of Figure 2. The figure shows group average proportion data. Proportions close to 0, displayed in blue, indicate a uniform preference for the risk-free option of $100. Proportions close to 1, displayed in yellow, indicate a uniform preference for the risky option. The results show that participants prefer the

[4]All data, analyses, and modeling scripts from Experiments 1 and 2 are available on Open Science Framework: https://osf.io/4dchn/. The analyses reported in this article contain all variables of interest and experimental conditions that we tested.

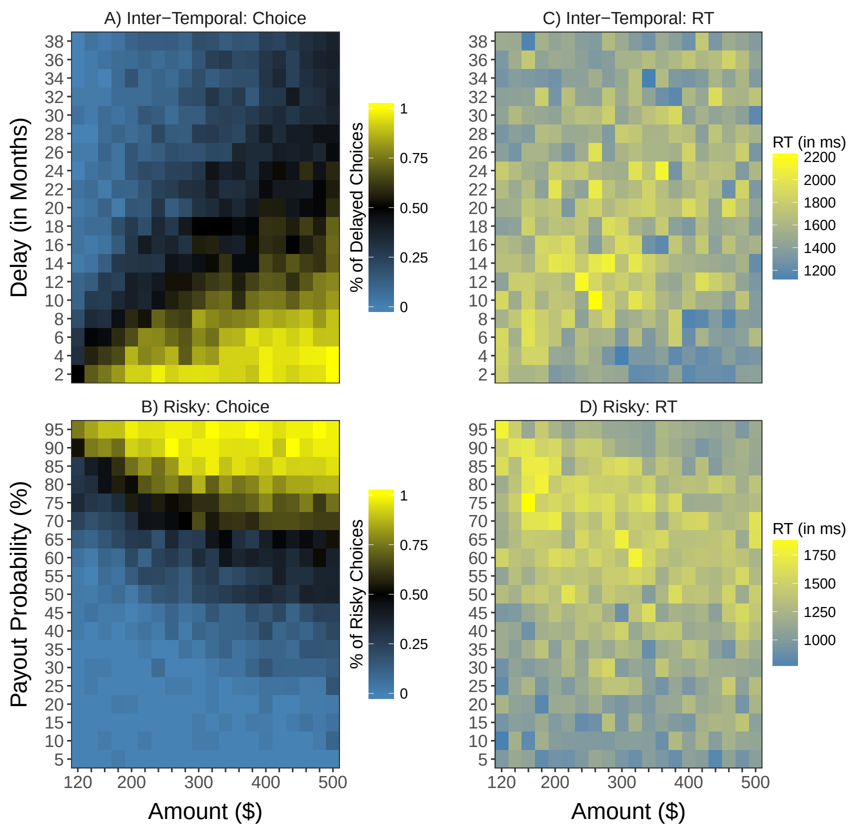

Figure 2. Behavioral data of the inter–temporal and risky choice trials in Experiment 1 averaged over participants. A) Proportion of preference data for inter–temporal choices. A proportion of 0 (blue) indicates exclusive preference for the \$100 now option, a proportion of 1 (yellow) indicates exclusive preference for the delayed option. Black boxes represent proportions around 0.50. B) Proportion of preference data for risky choices. A proportion of 0 (blue) indicates exclusive preference for the \$100 certain option, a proportion of 1 (yellow) indicates exclusive preference for the risky option. Black boxes represent proportions around 0.50. C) RT data for inter–temporal choices. D) RT data for risky choices. Low RTs are closer to blue on the color spectrum.

certain option when the risky option does not pay very well (i.e., amounts not much higher than \$120) or when the payout probability is low (i.e., close to 5%). In contrast, participants prefer the risky option when the risky option pays well (i.e., amounts close to \$500) or when the payout probability is high (i.e., close to 95%). We performed the same analysis for the risky choices, with amount and payout probability of the risky option as fixed effects and participant-specific random intercepts and slopes, which showed that both payout probability ($b = 5.67, z = 12.03, p < .001$) and amount ($b = 1.32, z = 8.78, p < .001$) are significant predictors of risky choice rates. The positive sign of both regression coefficients indicates that participants were more likely to select the risky option when amount and payout probability increased.

Response Times. The choice data indicates, perhaps unsurprisingly, that people prefer high payout, short delays, and high probability. What can we learn from the RT data? Overall, higher RTs were associated with inter-temporal choices ($M = 1,862$ ms) compared to risky choices

($M = 1,420$ ms). Also, certain or now options were chosen faster ($M = 1,422$ ms) than risky or delayed options ($M = 2,031$ ms). Figure 2C shows group-average RT data for the inter-temporal choice trials. Low RTs are closer to blue on the color spectrum and high RTs are closer to yellow. The results show that the more extreme preferences in terms of proportion are accompanied by lower RTs (see also Dai & Busemeyer, 2014). The choices for which preferences varied among participants (~50%; the black diagonal in Figure 2A) are accompanied by higher RTs (the yellow diagonal in Figure 2C), indicating a lower absolute strength of preference.

Figure 2D shows group-average RT data for the risky choice trials. Similar to the inter-temporal choice RTs, the results show that the more extreme preferences in terms of proportion are accompanied by lower RTs. The choices for which preferences varied among participants (the darker axis from the top-left to the mid-right in Figure 2B) are accompanied by higher RTs (the yellow axis in Figure 2D), indicating a lower absolute strength of preference.

In sum, people prefer high payout, short delays, and high probability. As for response times, people take less time making risky choices than inter-temporal choices, and take a relatively short time to make choices for which response options are extreme. In the next section, we turn to the modeling results. We are looking for two things: 1) Are strengths of preference observed in the behavioral results reflected in the pattern of drift rates in the models we consider? 2) How do people weigh delay and probability in their choices?

Modeling Results

Parameter convergence was satisfactory as indicated by the individual chains mixing properly.[5] Numerically, we compare the "proportional", the "invariant", and the "symmetrical" models (and their variants) by calculating the Deviance Information Criterion (DIC; Spiegelhalter, Best, Carlin, & van der Linde, 2002), a measure which balances goodness of fit against model complexity.[6] DIC values for the 12 models can be found in Table 2. The table shows that in terms of the DIC criterion, the best–fitting model is the linear–4A "proportional" model and the worst model belongs to the "invariant" account (nonlinear–4A model). This suggests that decision makers' absolute valuation of the now and certain options in the inter–temporal and risky domains, respectively, change with different alternatives (different values of delay and probability) for the delayed and risky options. In addition, the best–fitting model shows that the linear functional form of the drift rate and separate starting points for the drift process of each choice option provide better fits than its competitors.

We examined posterior predictive data for the best fitting linear–4A "proportional" model, and compared these to the behavioral data (posterior predictive data for the other two models – invariant and symmetrical – can be seen in the Supplemental Materials available online). Figure 3

[5]The focus of the modeling results is on the LBA and we will refer to the expected and discounted utility models (Prospect Theory and Hyperbolic Discounting) wherever necessary.

[6]DIC is similar to the well–known BIC and AIC measures. However, in hierarchical models, the number of free parameters is not well–defined. As such, DIC quantifies model complexity as across–sample variability in model fit instead. Lower values of DIC indicate better support for a model from the data.

Table 2
DIC values summed over participants for all 12 models fit to the Experiment 1 data set. The best model for these data is the linear–4A proportional model (in bold).

Model	P	Deviance	pD	DIC
Proportional				
Linear–2A	17	57,459.82	549.43	58,558.68
Nonlinear–2A	20	57,881.67	914.51	59,710.69
Linear–4A	**19**	**57,135.14**	**651.98**	**58,439.09**
Nonlinear–4A	22	57,885.15	920.59	59,726.34
Invariant				
Linear–2A	13	65,653.85	382.53	66,418.90
Nonlinear–2A	16	65,652.03	490.78	66,633.61
Linear–4A	15	64,988.77	388.96	65,766.65
Nonlinear–4A	18	65,053.51	537.21	66,127.92
Symmetrical				
Linear–2A	13	60,102.31	407.51	60,917.34
Nonlinear–2A	16	60,074.14	495.70	61,065.56
Linear–4A	15	59,576.31	471.22	60,518.74
Nonlinear–4A	18	59,758.51	555.81	60,870.14

Note: P = Number of free parameters per participant; Deviance = -2 times the likelihood of the mean parameter estimate; pD = -2 times the mean likelihood of the overall model + 2 times the likelihood of the mean parameter estimate; DIC = Deviance+2pD.

indicates that the model fits the RT and choice data well. It provides good fits for response times of all choice options, and accounts for observed choice proportions (see Choice panel).[7] Figure 3 also shows choice simulations based on the best fitting parameters for the Prospect Theory and Hyperbolic Discounting models. It can be seen that these models provide a good fit to the data, almost indistinguishable from the predictions of the LBA model. In other words, accounting for RT data (in addition to choice data) does not affect the way that the LBA accounts for and predicts participants' preferential choices.[8]

What can we learn from the resulting estimated parameters? Table 3 contains median values of the group parameters of the best–fitting model (i.e., the linear–4A "proportional"; see Supple-

[7]It is important to note that every cell in our design contained only a single observation (i.e., any participant contributed only a single choice for each amount/delay or amount/probability combination). As such, our data are relatively noisy, and the model fits reflect that noise.

[8]The Mean Squared Error (*MSE*) on the simulated choice proportions for each model and choice type, $\frac{1}{n}\sum_{i=1}^{n}(C_i - \hat{C}_i)^2$ (where n=380: the number of trials for each type of choice; and C: proportion of choices for the delayed/risky options for each individual choice): $\mathrm{LBA_I} = 0.083$, Hyperbolic Discounting = 0.053, $\mathrm{LBA_R} = 0.093$, Prospect Theory = 0.070. However, caution is advised when attempting direct quantitative comparisons between the two modeling frameworks: 1) the LBA assumes deterministic choice whereas the PT and HD models assume probabilistic choice (as they were implemented in this paper), 2) Different methodologies were used to estimate the parameters, and 3) the two modeling frameworks differ in terms of model complexity (e.g., number of free parameters) and functional form (see Pitt & Myung, 2002).

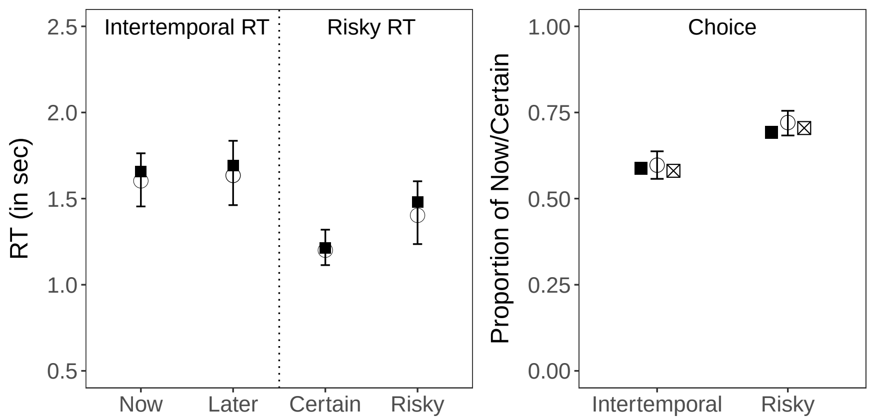

Figure 3. Posterior predictives of the linear–4A proportional model for RTs and choice data in Experiment 1. For all panels, circles represent mean posterior predictive data (error bars indicate 95% credible intervals of the posterior) and squares represent experimental data. Crossed-out squares (only for the Choice panel) indicate simulated predictive data for Prospect Theory (risky choice task) and Hyperbolic Discounting (inter–temporal choice task).

mental Materials for group-level posterior distributions of parameters). Given the particular set of delays, probabilities, and amounts we used, there are three things that the model parameters indicate: first, amount factors more in the decision for inter–temporal choices than for risky choices as evidenced by larger values for α_{N_X} and α_{D_X} than for α_{C_X} and α_{R_X}. Second, risk factors more in the decision than delay as evidenced by larger values for α_{C_P} and α_{R_P} than for α_{N_D} and α_{D_D}. Finally, decisions are made quicker on average for risky than for inter–temporal choices as evidenced by higher drift rates and a lower response threshold for risky (B_R) than for inter–temporal choices (B_I)[9].

In order to delve more deeply into the modeling results, we examine individual drift rates for each choice and each participant by entering the appropriate parameters into Equation 1 and Equation 2. The difference between the resulting drift rates for all participants' inter–temporal choice data, $\nu_N - \nu_D$, can be found in Figure 4. A highly positive difference between the now drift rate and the delayed drift rate, displayed in blue, indicates a strong preference for the now choice option. A highly negative difference between the now drift rate and the delayed drift rate, displayed in yellow, indicates a strong preference for the delayed choice option. The results show there are considerable individual differences in the extent to which participants weigh amount and delay. For example, P5 is mostly driven by the amount in dollars (as indicated by the predominantly vertical transition in colors), whereas P1 tends to be driven by the delay (as indicated by the predominantly horizontal transition in colors). We also see differences in the proportions of choices: participants for whom yellow dominates generally prefer the "now" option, whereas participants for whom blue dominates generally prefer the "delayed" option.

The difference between the resulting drift rates for all participants' risky choice data, $\nu_C -$

[9]Results from a parameter recovery analysis are presented in the Supplemental Materials

Table 3

Estimated parameters of the linear–4A proportional model of the data from Experiment 1. Displayed are the median parameter values of the group parameters, with a 50% credible interval of the posterior presented in parentheses. Rows represent parameters and columns represent the two group parameters.

Description	Parameter	μ	σ
Inter–temporal Choice			
Starting Points (A)			
Now	A_{I_N}	0.90 (0.50, 1.29)	3.27 (2.90, 3.68)
Delayed	A_{I_D}	2.18 (1.66, 2.57)	2.20 (1.75, 2.71)
Threshold	B_I	0.92 (0.71, 1.11)	1.11 (0.98, 1.28)
Inter–temporal drift rates (ν_N, ν_D)			
Offset: Now	ν_{N_0}	2.96 (2.83, 3.09)	0.97 (0.87, 1.08)
Offset: Delayed	ν_{D_0}	3.43 (3.29, 3.55)	0.93 (0.81, 1.08)
Amount Scale: Now	α_{N_X}	0.12 (0.10, 0.13)	0.13 (0.12, 0.15)
Amount Scale: Delayed	α_{D_X}	0.12 (0.11, 0.13)	0.06 (0.06, 0.07)
Delay Scale: Now	α_{N_D}	0.13 (0.13, 0.14)	0.05 (0.05, 0.06)
Delay Scale: Delayed	α_{D_D}	0.29 (0.27, 0.32)	0.22 (0.19, 0.24)
Risky Choice			
Starting Points (A)			
Certain	A_{R_C}	0.43 (0.21, 0.71)	1.90 (1.68, 2.14)
Risky	A_{R_R}	2.40 (2.05, 2.66)	1.75 (1.40, 2.22)
Threshold	B_R	0.91 (0.77, 1.03)	0.72 (0.62, 0.84)
Risky drift rates (ν_C, ν_R)			
Offset: Certain	ν_{C_0}	3.20 (3.11, 3.29)	0.62 (0.53, 0.75)
Offset: Risky	ν_{R_0}	3.80 (3.69, 3.93)	0.62 (0.46, 0.84)
Amount Scale: Certain	α_{C_X}	0.03 (0.03, 0.03)	0.03 (0.03, 0.03)
Amount Scale: Risky	α_{R_X}	0.09 (0.08, 0.09)	0.06 (0.05, 0.07)
Probability Scale: Certain	α_{C_P}	0.17 (0.16, 0.18)	0.07 (0.06, 0.07)
Probability Scale: Risky	α_{R_P}	0.48 (0.45, 0.52)	0.27 (0.23, 0.30)
Non-Decision time	$t0$	0.10 (0.07, 0.13)	0.12 (0.10, 0.13)

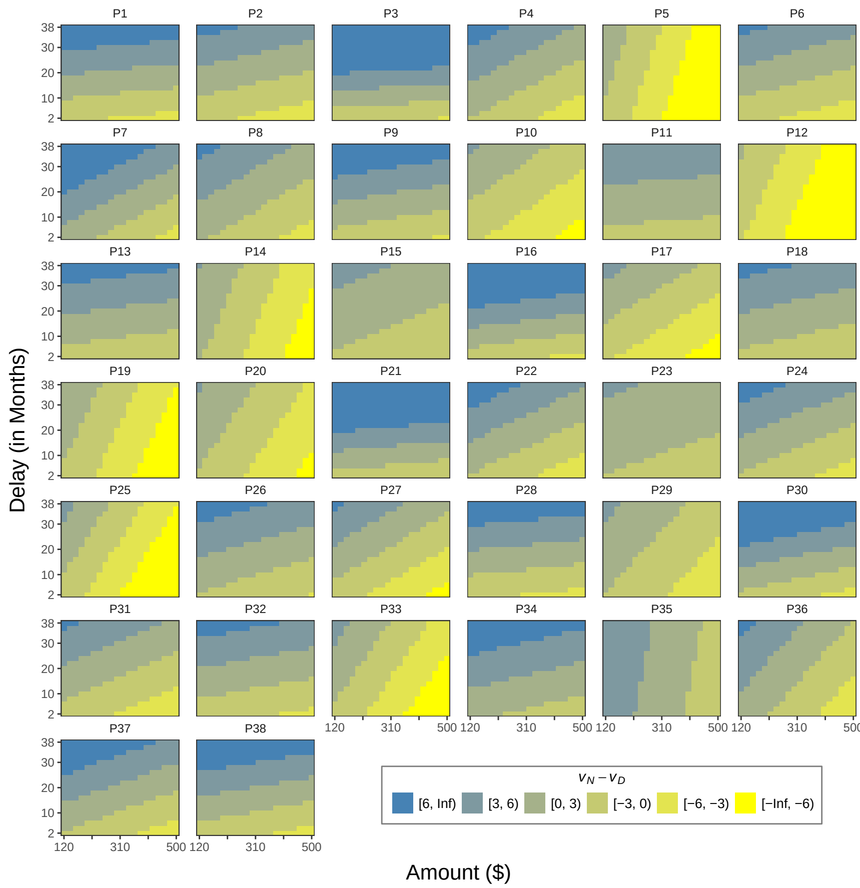

Figure 4. Absolute difference between the drift rates for the now and delayed options ($\nu_N - \nu_D$) across choices and participants (i.e., "P" panels) for the inter–temporal choice trials in Experiment 1. Positive drift rates reflect a preference for the now option and are displayed in colors that are closer to blue on the color spectrum. Negative drift rates reflect a preference for the delayed option and are displayed in colors that are closer to yellow on the color spectrum.

ν_R, can be found in Figure 5. Note that participants are location–matched across the figures. The results show that most participants had their strength of preference almost exclusively be determined by probability, rather than amount (the transition among colors goes predominantly along the vertical axis). There are still a few exceptions to this rule, for instance P12 who seems to weigh amount and probability almost evenly. The other stand–out observation here is that people are very risk–averse: across the board, we see a lot more blue than we see yellow.

Discussion

Experiment 1 showed that in an inter–temporal choice setting, people prefer high payouts and short delays. In a risky choice setting, they prefer high payouts and high payout probabilities. We have showed how RT data can augment the information provided by choice responses: in conjunction with choice responses they give a measure of strength of preference. In our experiment, decisions were made quicker on average for risky than for inter–temporal choices. In addition, the

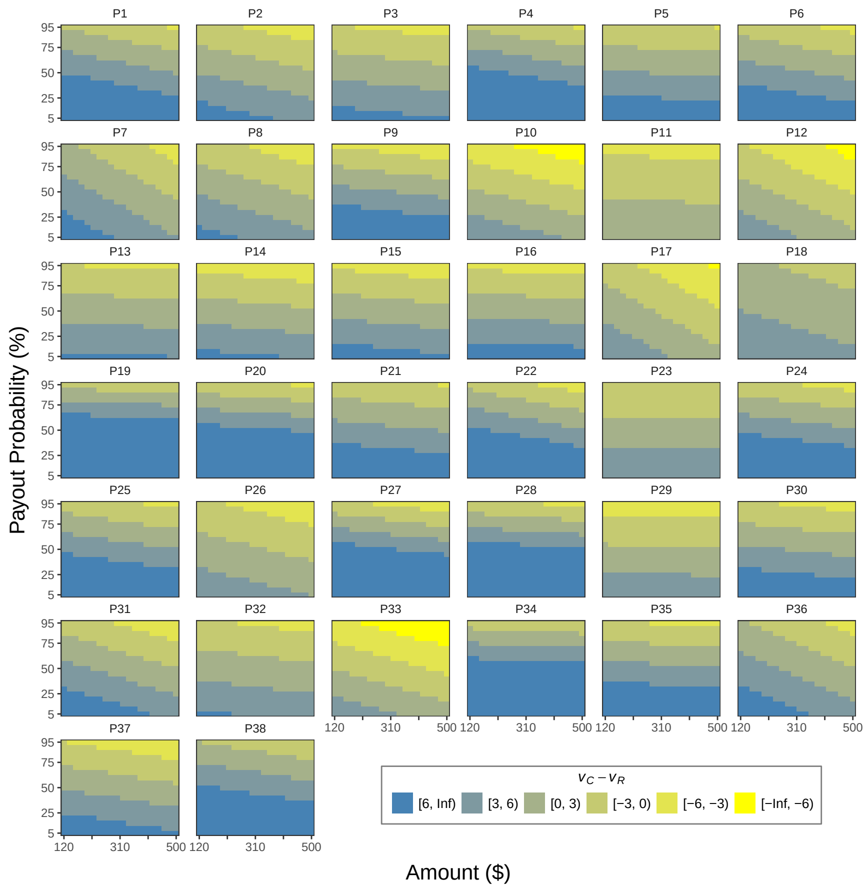

Figure 5. Absolute difference between the drift rates for the certain and risky options ($\nu_C - \nu_R$) across choices and participants (i.e., “P” panels) for the risky choice trials in Experiment 1. Positive drift rates reflect a preference for the certain option and are displayed in colors that are closer to blue on the color spectrum. Negative drift rates reflect a preference for the risky option and are displayed in colors that are closer to yellow on the color spectrum.

formal comparison between the three model accounts of the LBA (“invariant”, “proportional”, and “symmetrical”) revealed that the absolute attractiveness of the now/certain choice option changes with different alternatives for the delayed/risky options as implemented by the “proportional” model. The LBA produced almost identical predictions when compared to standard modeling approaches in risky and inter–temporal choice (such as Prospect Theory and Hyperbolic Discounting), suggesting that accounting for RT data does not invalidate good predictions for the choice data: our cognitive process model can account for both streams of behavioral data (choice and RT) equally well.

Experiment 2

Experiment 1 suggested that our instantiation of LBA can provide a good fit to the choice and RT data from inter–temporal and risky choices based on a simple concept of accumulated preferential strength. In Experiment 2 we examine whether our cognitive process model can provide a good explanation for choice and RT data when probability and delay combine in a single option. We also test the three different accounts of the relationship between now/certain and delayed/risky options ("invariant", "symmetrical", and "proportional") and put LBA to the test by fitting a standard discounted expected utility model to inter–temporal risky choice data. Thus, the main objective of Experiment 2 is twofold: first, to extend the LBA to account for the combined effect of probability and delay, and second, to examine whether the "proportional" model will be the best fitting model.

On each trial of the experiment participants faced a choice between an option that was available now with certainty and one that differed from the fixed option in terms of probability, time of play, and amount of money. The full factorial combination of all the amounts, delays and probabilities for which we wished to elicit preferences resulted in a very large number of trials (i.e., 7,220; see Method). For this reason, in Experiment 2 we decided to collect a large amount of data from a small number (4) of participants.

Method

Participants. Four graduate students (2 female; $M_{age} = 23$) at the University of New South Wales participated in return for a $15 participation fee. In addition, they were paid $2 (i.e., outcome of the sure option) in each of 10 experimental sessions[10].

Material. The experiment consisted of a total of 7,220 inter–temporal and risky choice trials. For all choices, participants had to indicate what they preferred: $100 now for sure or $\$X$ in D months with P% certainty, with $\$X$ varying from $120 to $500 in $20 increments (for a total of 20 amounts), D varying from 2 months to 38 months in 2 month increments (for a total of 19 delays), and P varying from 5% to 95% in 5% increments (for a total of 19 probabilities). Thus, every combination of amount, delay, and probability was presented to the participant once as an alternative to $100 now for sure.

Procedure. Participants completed the experiment in 10 separate experimental sessions, each comprising of 722 choice trials. Experimental sessions were again separated by a minimum of three hours for each participant. Presentation of the sure option, and the inter–temporal and risky option on the screen was counterbalanced across participants.

Implementation of the Model

Fitting the model to the inter–temporal and risky choice data in Experiment 1 revealed that the linear–4A model was the best fitting variant within each model account, indicating that linear drift rates and separate starting points for each choice option improves model fits and predictive accuracy. We assumed the same model variant in Experiment 2, with linear drift rates and as many parameters for the upper starting point A as the number of choice options (i.e., two: A_{NC} for the now/certain option and A_{DR} for the delayed/risky option). Thus, in Experiment 2, we fit the

[10] As in Experiment 1, participants were told that one trial from the experiment would be selected and the gamble would be played for real.

same three model accounts (proportional, invariant, and symmetrical), but focusing on the linear–$2A$ variants of these model accounts. The linear–$2A$ variant is identical to the linear–$4A$ variant in Experiment 1, with the only difference being that there were twice as many choice options in Experiment 1 (four, hence four starting parameters A) compared to only two in Experiment 2 (hence the $2A$). As in the implementation of the model in Experiment 1, we assumed one parameter for threshold B, and one non–decision time parameter $t0$.

In Experiment 1 we observed that the same evidence accumulation (or strength of preference) process provided a good fit to both risky and inter–temporal choices. This allowed us to assume that expanding the drift rates to account for the combination of probability and delay in the same choice option would provide a good account for the risky inter–temporal choice data. The definition of the drift rates of the linear–$2A$ "proportional" model for the risky inter–temporal choice task follows the same principles as for the drift rates when the two dimensions are examined in isolation, that is, a weighted sum of the attribute values of each option. Hence, we extended the model presented in Experiment 1 to account for the joint effect of delay and probability as follows:

$$\begin{aligned} \nu_{NC} &= \nu_{NC_0} - \alpha_{NC_X} \times (X/20 - 6) - \alpha_{NC_D} \times (19 - D/2) - \alpha_{NC_P} \times (P/5 - 1) \\ \nu_{DR} &= \nu_{DR_0} - \alpha_{DR_X} \times (25 - X/20) - \alpha_{DR_D} \times (D/2 - 1) - \alpha_{DR_P} \times (19 - P/5) \end{aligned} \tag{9}$$

where X, D, and P denote the amount in dollars, delay in months, and payout probability, respectively, for the delayed/risky option, ν_{NC} and ν_{DR} denote drift rates for the now/certain and the delayed/risky choice options, ν_{NC_0} and ν_{DR_0} denote offset parameters for the now/certain and the delayed/risky choice options, α_{NC_X} and α_{DR_X} denote amount scale parameters for the now/certain and the delayed/risky choice options, α_{NC_D} and α_{DR_D} denote delay scale parameters for the now/certain and the delayed/risky choice options, and α_{NC_P} and α_{DR_P} denote risk scale parameters for the now/certain and the delayed/risky choice options. $\nu_{NC} = \nu_{NC_0}$ if $X = 120$, $D = 38$, and $P\% = 5$ (the option that most favors the "now/certain" choice). $\nu_{DR} = \nu_{DR_0}$ if $X = 500$, $D = 2$, and $P\% = 95$ (the option that most favors the "delayed/risky" choice).

This results in a total of 12 parameters to be estimated: A_{NC}, A_{DR}, B, $t0$, ν_{NC_0}, α_{NC_X}, α_{NC_D}, α_{NC_P}, ν_{DR_0}, α_{DR_X}, α_{DR_D}, and α_{DR_P}. Together, these parameters should account for the distribution of response times and choice proportions for the combined inter–temporal and risky choice data.

Just as for Experiment 1, we fit two other models that test specific assumptions about the underlying choice process. The "invariant" model estimates a single ν_{NC} parameter (drift rate for the now/certain option) for all inter–temporal risky choice trials. Thus, it replaces the four free parameters (ν_{NC_0}, α_{NC_X}, α_{NC_D}, and α_{NC_P}) from the definition of ν_{NC} under the "proportional" model (see Equation 9) with one free parameter. Conceptually, this simpler model assumes that the absolute value of the now/certain option does not change with different alternatives for the delayed/risky option (i.e., same absolute value for the now/certain option across all trials). The "invariant" model has nine free parameters to be estimated. The "symmetrical" model assumes that drift rates for the now/certain option vary symmetrically (in the opposite direction) with drift rates for the delayed/risky option (i.e., $\alpha_{NC_X} = \alpha_{DR_X}$, $\alpha_{NC_D} = \alpha_{DR_D}$, and $\alpha_{NC_P} = \alpha_{DR_P}$). This model has nine free parameters to be estimated.

For Experiment 1, the "proportional" model fit the data better than both the "invariant" and the "symmetrical" models. Here, we examine if the same result holds when the inter–temporal and risky elements are combined in a single choice. Due to the small number of participants, we fit the models to individual data (as opposed to hierarchical models in Experiment 1). We used formal model comparison to find the account best supported by the data. Details on starting values, prior distributions, and number of iterations may be found in the Appendix.

Discounted Expected utility model. As in Experiment 1, we also fit a discounted expected utility model for the inter–temporal risky choice data. The model combines (in a multiplicative way) a hyperbolic discounting of time and probabilities (see Vanderveldt et al., 2015):

$$D(t, \theta) = \frac{1}{[(1 + kt)^{s_d} \times (1 + h\theta)^{s_p}]} \tag{10}$$

The first term in the denominator is identical to the two–parameter hyperbolic discounting model used in Experiment 1 (see Equation 7). The second term is the probability discounting part, where θ represents the odds against receiving a reward. It can be expressed in terms of actual probabilities as $\theta = (1-p)/p$. This form of probability discounting can be understood as reflecting similar properties of the probability weighting function (Equation 5) used in Experiment 1 (see also Vanderveldt et al., 2015). As in Experiment 1, the value of a delayed risky prospect is defined as $V = \sum D(t_i, \theta_i)u(X_i)$. We used the same power utility function (Equation 4) and the softmax rule (Equation 8) for probabilistic choice. The parameter estimation procedure was identical to that in Experiment 1, apart from the fact that we used 500 starting points for each set of parameters.

Behavioral Results

Choice and response times. All participants completed the experiment. As in Experiment 1 we excluded responses that were slower than 7s and faster than 250ms (0.47% of all trials). Preference data for Experiment 2 can be found in Figure 6. The figure shows group average proportion data. Proportions close to 1, displayed in yellow, indicate a uniform preference for the delayed/risky choice. Proportions close to 0, displayed in blue, indicate a uniform preference for the \$100 now/certain choice. The results show that participants prefer the now/certain option when the delayed/risky option does not pay very well (i.e., amounts not much higher than \$120), when the delay is long (i.e., close to 38 months), or when the payout probability is low (i.e., close to 5%). In contrast, participants prefer the delayed/risky option when it pays well (i.e., amounts close to \$500), when the delay is short (i.e., close to now), or when the payout probability is high (i.e., close to 95%).

We analyzed the data using a generalized linear mixed-effects model (Binomial distribution and logit transformation) with amount, payout probability, and delay in months of the delayed/risky option as fixed-effects predicting selection of the delayed/risky option, and random intercepts and slopes (amount, probability, and delay) for each participant. The results indicated a positive relationship between amount ($b = 1.75, z = 7.87, p < .001$) and payout probability ($b = 3.72, z = 10.61, p. < 001$) and selection of the delayed/risky option on one hand, and a negative relationship between delay ($b = -2.27, z = -2.23, p = .026$) and selection of the delayed/risky option on the other hand.

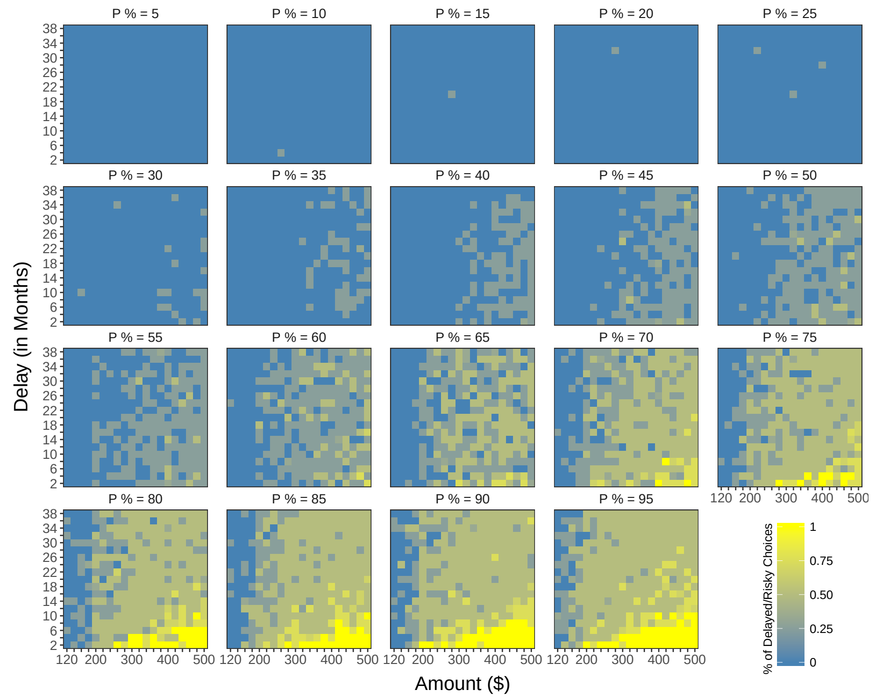

Figure 6. Aggregate choice preference data of Experiment 2. Panels represent different probability levels. A proportion of 0 (blue) indicates exclusive preference for the $100 now/certain option, a proportion of 1 (yellow) indicates exclusive preference for the delayed/risky option.

Analogous to Experiment 1, the choice data indicate that people prefer high payouts, short delays, and high probabilities. RT data for Experiment 2 can be found in Figure 7. The figure shows group average RT data, binned in five equal groups. Low RTs are closer to blue on the color spectrum and high RTs are closer to yellow on the color spectrum. The results show a clear pattern: as the probability of the delayed/risky option increases, participants tend to slow down (especially once the probability exceeds .5).

Modeling Results

Parameter convergence was satisfactory. DIC values for the three models can be found in Table 4. We obtained the same results as in Experiment 1, with the “proportional” model being the best–fitting model and the “invariant” being the worst–fitting model. Consequently, this suggests that decision makers’ absolute valuation of the now/certain option changes with different alternatives for the delayed/risky option.

Moving to the estimated parameters, we examined posterior predictive data for the linear–2A

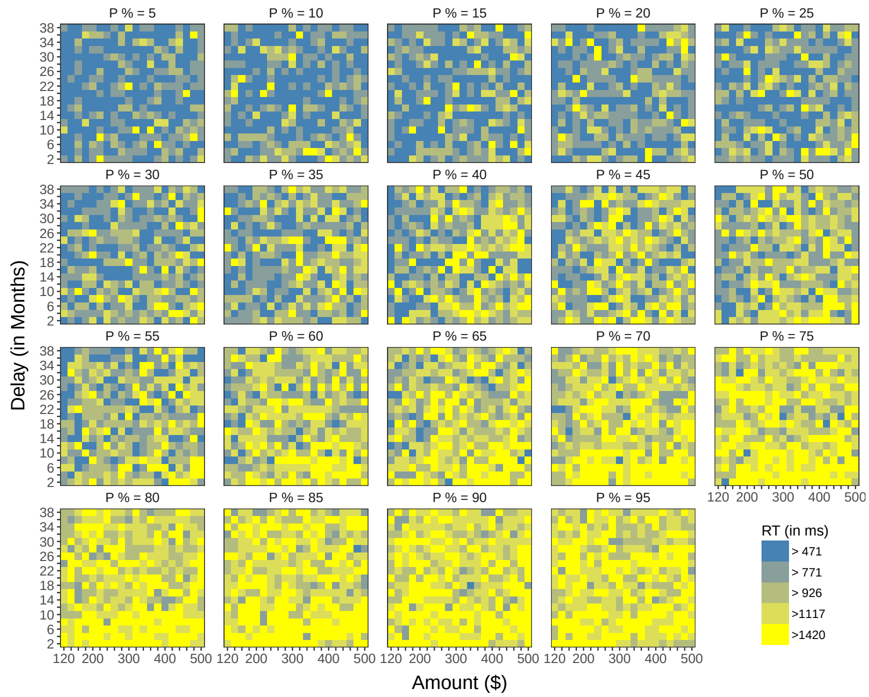

Figure 7. RT data of Experiment 2. Panels represent different probability levels. Low RTs are closer to blue on the color spectrum.

Table 4

DIC values summed over participants for all three models fit to the Experiment 2 data set. The best model for these data is the proportional model (in bold).

Model	P	Deviance	pD	DIC
Proportional	**12**	**37,028**	**41**	**37,110**
Invariant	9	47,944	30	48,003
Symmetrical	9	38,033	31	38,096

Note: P = Number of free parameters per participant; Deviance = -2 times the likelihood of the mean parameter estimate; pD = -2 times the mean likelihood of the overall model + 2 times the likelihood of the mean parameter estimate; DIC = Deviance+2pD.

“proportional” model, which are compared against the behavioral empirical data. The model fit the data well (i.e., choice proportions and RTs; see Figure 8), only slightly underestimating response times for the delayed/risky responses. Thus, extending the LBA to account for the combined effect of time and probability and implementing the same principle of accumulated preference for inter–temporal risky choices provides a parsimonious and psychologically plausible account of choice behavior in this context. Figure 8 also plots simulated choice predictions of a discounted expected utility model (multiplicative hyperboloid model; red triangle marker): compared to a discounted expected utility model that has been found to provide good fits to inter–temporal risky choice data (see Vanderveldt et al., 2015), the LBA performs well with the additional benefit of providing predictions for RT data too.[11]

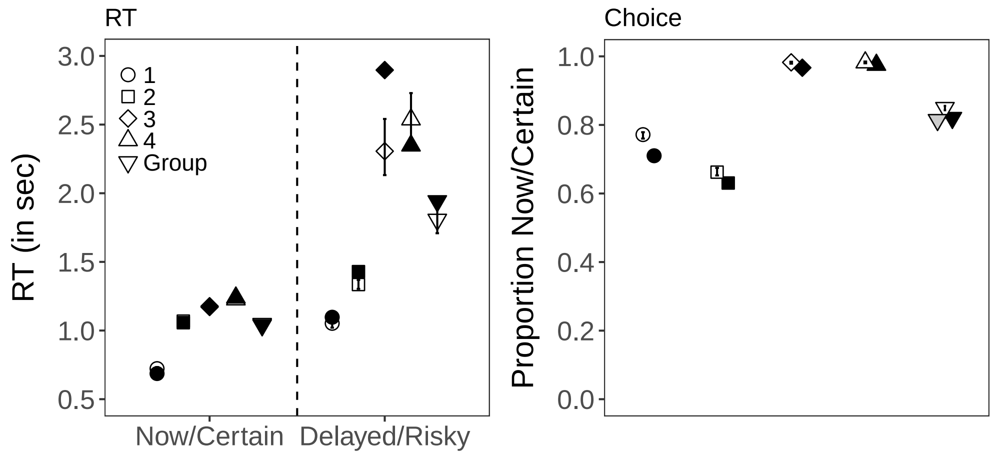

Figure 8. Posterior predictives of the linear–$2A$ proportional model for RTs (Now/Certain and Delayed/Risky options) and choice data (Proportion of Now/Certain choices) in Experiment 2 (individuals 1-4 and group results). For all panels, white-filled points represent mean posterior predictive data (error bars indicate 95% credible intervals of the posterior) and black-filled points represent experimental data. The gray triangle marker (only for the Choice panel) indicates simulated predictive data for the multiplicative hyperboloid model.

Table 5 contains median parameter values of the aggregate and individual participant parameters of the linear–$2A$ "proportional" model. One aspect that clearly stands out is that probability has a much stronger influence on the decision than either amount or delay, as evidenced by the fact that the risk scale parameters (α_{CN_R} and α_{DR_R}) are substantially larger than the amount (α_{CN_X}, α_{RD_X}) and delay scale parameters (α_{CN_D}, α_{DR_D}). It is important to note that as in Experiment 1, the larger effect of probability on participants' choices is predicated on the range of amounts and delays we used in this study.

In order to delve more deeply into this pattern, we examined individual drift rates for two representative participants (i.e., based on their model parameter values) by entering the appropriate parameters into Equation 9. Inspection of the individual parameter values (columns labelled 1 to 4 in Table 5) suggests that 2 out of 4 participants (i.e., P1 and P2) weigh probability more than delay in their decisions, whereas the remaining two participants seem to equally weigh both dimensions to make choices. The difference between the resulting drift rates for representatives of these two types of participants' (P2 and P3) risky inter–temporal choice data, $\nu_{CN} - \nu_{RD}$, can be found in Figure 9. A highly positive difference between the now/certain drift rate and the delayed/risky drift rate, displayed in blue, indicates a strong preference for the now/certain choice. A highly negative difference between the now/certain drift rate and the delayed/risky drift rate, displayed in yellow, indicates a strong preference for the delayed/risky choice. As expected based on the individual parameter values, Figure 9 shows that the two participants have quite distinct choice

[11]The corresponding *MSE*s are: LBA = 0.134; Multiplicative Hyperboloid = 0.104

Table 5

Estimated parameters of the proportional model in Experiment 2. Displayed are the median parameter values of the group (Group column) and individual (1 to 4 columns) parameters, with a 50% credible interval of the posterior presented in parentheses.

Parameter	Group	1	2	3	4
A_{NC}	0.98 (0.01, 1.74)	0.00 (0.00, 0.01)	0.01 (0.00, 0.02)	2.41 (2.35, 2.47)	1.24 (1.17, 1.31)
A_{DR}	2.39 (1.74, 4.27)	1.99 (1.92, 2.08)	1.40 (1.32, 1.49)	4.73 (4.16, 5.28)	3.86 (3.43, 4.37)
B	1.89 (1.28, 2.06)	1.04 (1.01, 1.07)	1.90 (1.87, 1.92)	1.87 (1.80, 1.95)	2.13 (2.11, 2.16)
ν_{NC_0}	3.66 (3.53, 4.92)	3.45 (3.40, 3.50)	3.72 (3.68, 3.76)	5.43 (5.37, 5.49)	3.60 (3.56, 3.65)
ν_{DR_0}	3.72 (3.15, 3.85)	3.77 (3.68, 3.85)	3.79 (3.72, 3.85)	3.81 (3.58, 4.05)	2.64 (2.48, 2.82)
α_{NC_X}	0.03 (0.01, 0.06)	0.01 (0.01, 0.01)	0.07 (0.07, 0.07)	0.05 (0.05, 0.06)	0.00 (0.00, 0.01)
α_{NC_D}	0.02 (0.01, 0.04)	0.01 (0.00, 0.01)	0.01 (0.00, 0.01)	0.05 (0.05, 0.06)	0.03 (0.03, 0.04)
α_{NC_P}	0.15 (0.11, 0.17)	0.17 (0.17, 0.18)	0.17 (0.17, 0.17)	0.13 (0.13, 0.14)	0.07 (0.07, 0.07)
α_{DR_X}	0.13 (0.10, 0.16)	0.09 (0.09, 0.09)	0.14 (0.13, 0.14)	0.20 (0.19, 0.22)	0.13 (0.11, 0.14)
α_{DR_D}	0.12 (0.01, 0.36)	0.01 (0.00, 0.01)	0.02 (0.01, 0.02)	0.45 (0.43, 0.48)	0.29 (0.27, 0.31)
α_{DR_P}	0.28 (0.19, 0.35)	0.28 (0.28, 0.29)	0.17 (0.17, 0.18)	0.50 (0.47, 0.53)	0.25 (0.23, 0.27)
$t0$	0.01 (0.00, 0.09)	0.11 (0.11, 0.12)	0.00 (0.00, 0.01)	0.04 (0.02, 0.05)	0.00 (0.00, 0.01)

profiles (see also individual parameter values in Table 5): P2's choices are almost exclusively driven by amount and payout probability as indicated by the vertical transitions in colors within and across probability levels. P3 is considerably more risk averse than P2 as there is a lot more blue than yellow in their panel. In addition, it appears that P3's choices are also impacted by delay (in addition to probability and amount) as shown by the mostly horizontal transitions in colors for probability levels greater than 55% (see Supplemental Materials for individual differences in the drift rates of the remaining two participants).

General Discussion

The search for understanding the principles that underlie choice in inter–temporal and risky settings has been dominated by descriptive explanations of observed behavior. Choice anomalies and deviations from EUT and DUT have led to the development of a vast number of utility/subjective value-based models which propose different functional forms and additional parameters to account for observed behavioral effects. However, in recent years, decision scientists have started to adopt cognitive processing models of choice behavior, suggesting a possible paradigm shift within judgment and decision–making research (Bhatia & Mullett, 2016; Oppenheimer & Kelso, 2015). Information processing models have been applied in many areas of decision–making and have provided psychological explanations and insights into the dynamics underlying preferential choice (see e.g., Busemeyer & Townsend, 1993; Krajbich, Armel, & Rangel, 2010; Newell & Bröder, 2008; Rodriguez et al., 2014; Trueblood et al., 2014; Usher & McClelland, 2001).

In this work, we followed a similar approach using an evidence accumulation model (LBA) to account for inter–temporal and risky choices. The novelty of our approach rests on the fact that the same modeling framework can be applied to two seemingly different types of choices,

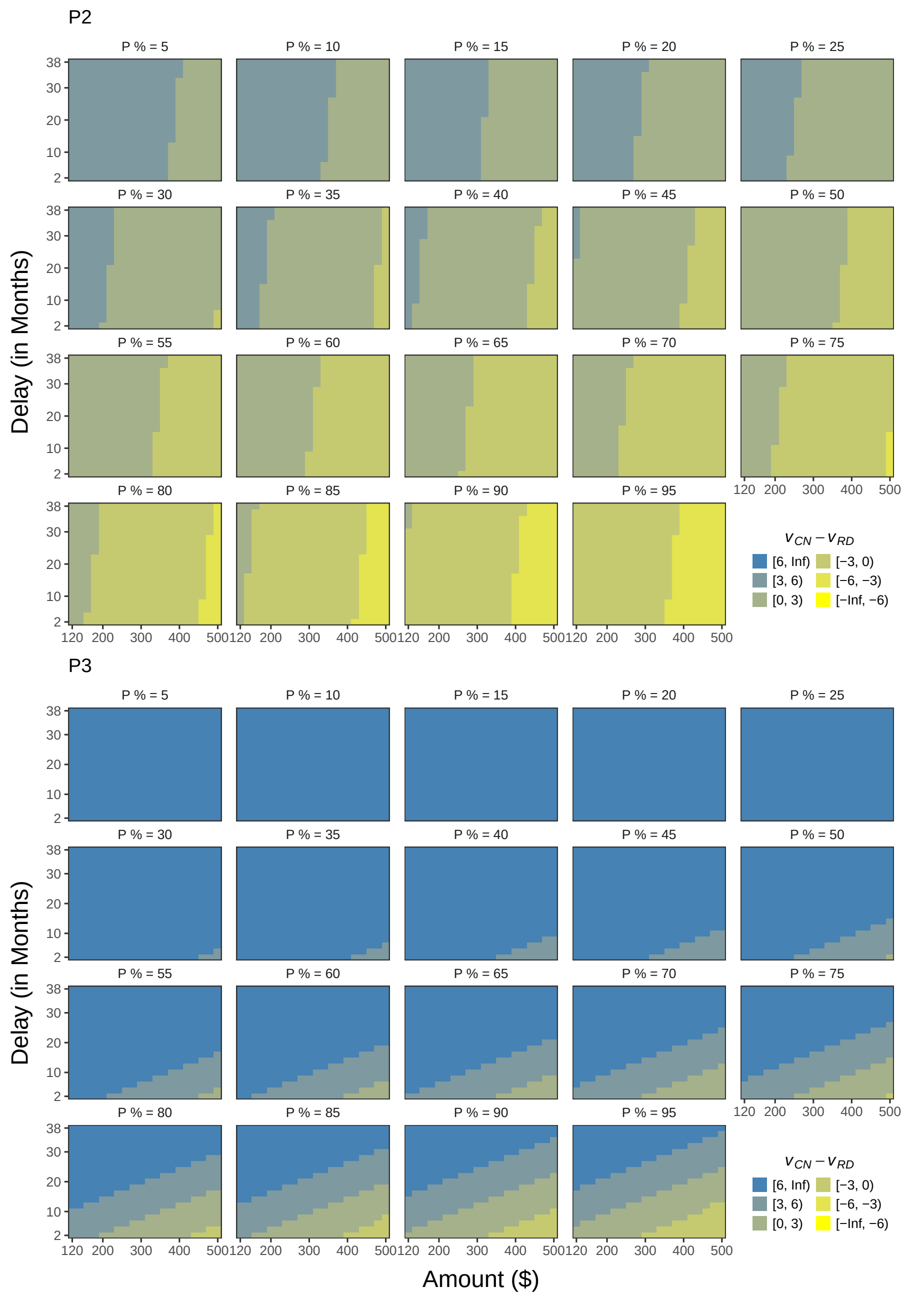

Figure 9. Absolute difference between the drift rates for the now/certain and delayed/risky options ($\nu_{CN} - \nu_{RD}$) for each choice in Experiment 2 for two participants (P2 and P3). Panels represent different probability levels. Positive drift rates reflect a preference for the now/certain option and are displayed in colors that are closer to blue on the color spectrum. Negative drift rates reflect a preference for the delayed/risky option and are displayed in colors that are closer to yellow on the color spectrum.

without relying on assumptions derived from either EU or DU models. In addition, the current work presents the first attempt to model the combined effect of probability and delay assuming an evidence accumulation framework and relying on a simple specification (weighted sum) of how preference is accumulated: drift rates provide a parsimonious and elegant measure for strength of preference, combining the information provided by choice responses and RTs.

Choice behavior in inter–temporal and risky settings

In Experiment 1, we observed that people prefer larger, sooner to later, and certain to risky payouts. A closer inspection of the results revealed that delay and payout probability had a larger effect on choice for inter–temporal and risky options, respectively, as compared to amount. Comparing inter–temporal and risky choices, amount appears to matter more in an inter–temporal setting. This pattern is consistent with observations from previous research on delay and probability discounting showing that changes in amount magnitude have a larger effect in an inter-temporal than a risky choice setting (see Greenhow, Hunt, Macaskill, & Harper, 2015; Myerson, Green, Scott Hanson, Holt, & Estle, 2003; Yi et al., 2006). For the particular set of delays, probabilities, and amounts we used, a comparison of the relative importance of probability and delay across choice settings (i.e., logit regression coefficients and model parameters) indicates that probability may have a larger effect on choice compared to delay. Vanderveldt et al. (2015) found in a task where both dimensions were combined that increasing the payout probability eliminated the effect of delay, whereas when delay was increased, the effect of probability was reduced but not completely eliminated. Nonetheless, they mentioned that the superior effect of probability may be an artifact of the amounts and the range of delays and probabilities used in their study. This could also be the case in our experiment: the larger effect of probability may have been the result of the range in which we manipulated amount and delay (probability is naturally constrained between 0 and 1). With longer delays (longer than 38 months that we used in this experiment) and different starting and ending points for the range of amounts (smaller than \$120 and larger than \$500 that we used in this experiment), the relative importance of delay could have been different. Alternatively and consistent with our results, probability may be generally more salient than delay (see also Konstantinidis, van Ravenzwaaij, & Newell, 2017).

RT data showed that risky choices were made on average faster than inter–temporal choices. Participants' responses were slower when risky or delayed options were more attractive than the default option of \$100 now or with certainty. In addition, clear preferences in terms of proportion produced shorter RTs. These results are suggestive of the dynamic nature of inter–temporal and risky choice, indicating that the use of static and descriptive models of choice may hinder our understanding of how preferences and choices are formed (see Dai & Busemeyer, 2014).

The purpose of Experiment 2 was to elicit preferences for the factorial combination of a wide range of amounts, payout probabilities, and delays. This resulted in a very large amount of delayed risky options being offered as choice alternatives to a fixed amount now-certain choice option. While this design allowed us to provide more accurate individual model parameter estimates, the small sample size in this experiment (i.e., 4) does not allow strong conclusions to be drawn regarding the generalizability of the behavioral patterns found in the data. Figure 9 shows that there is considerable variation in participants' choice patterns. However, the cognitive modeling analyses and the way that drift rates are defined (weighted sums and scaling parameters) allowed us to quantify and explain individual differences, and the extent to which participants weighed each dimension in their decision-making behavior.

Perspectives on modeling probability and delay

We used formal model comparison to pit three different variants of LBA against each other that differed in the assumptions they make about the absolute evaluation of the now/certain choice option.[12] In both experiments, the "proportional" model (with as many starting point parameters as the available choice options and linear drift rates) fit the data best, suggesting that the absolute attractiveness of the now/certain choice option cannot be judged in a vacuum. This goes against classic expected and discounted utility models which assume that the subjective value of an option is fixed and the product of a utility function paired with a discounting function (inter–temporal choice) or a probability weighting function (risky choice). Our cognitive process model makes no such assumptions; instead the definition of the drift rates suggests that preference for each option is formed through a weighted sum of its attributes (money, delay, and probability) and it is dependent on the numerical value of the attributes of the alternative option. However, one can assume different functional forms for the definition of drift rates and the way amount, probability, and delay combine. We tested this assumption by allowing the drift rates to have nonlinear forms, but this led to poorer fits compared to the linear forms of the drift rates. A complementary approach is to assume that drift rates incorporate the functional forms from existing models (such as those used in PT and HD) in determining preferences for choice alternatives (e.g., Dai & Busemeyer, 2014). In this sense, accumulated preference over time is governed by discounted or subjective utility valuations of delayed risky prospects (e.g., Rodriguez et al., 2014). Future research can determine the practical and theoretical advantages of implementing and testing such models. The objective of the present work was to present a process model which takes into account response times and assumes the *same* modeling and processing framework for both types of preferential choice.

We also compared the performance of our cognitive process model against standard approaches, such as Prospect Theory in risky choice and Hyperbolic Discounting in inter–temporal choice. We also attempted a combination of these two models (i.e., the multiplicative hyperboloid model) when probability and delay appear in the same choice option (Vanderveldt et al., 2015). In both Experiments 1 and 2, we observed that predictions from these models were almost identical to those of the LBA, indicating that assuming a modeling framework akin to attribute-wise models and accounting for choice response times provide an equally plausible account of choice behavior.

[12] Our model also assesses "relative" attractiveness and preferences since the numerical value of the now/certain option is also dependent on the numerical value of the delayed/risky option.

The additional benefits from using our cognitive process model are that the LBA provides predictions on two streams of behavioral data, choice proportions and RTs (while standard approaches are only concerned with the former, and in most cases at the aggregate level) and it also provides an economic way (weighted sum of attribute values) to model the effect of probability and delay when they are treated independently from each other, but also when combined. Glöckner and Herbold (2011) showed that Cumulative Prospect Theory can provide adequate ("reasonably good", p. 94) descriptions of aggregate choice behavior in risky choice tasks, but they suggested that accounting for individual choice behavior and the underlying choice process, models such as DFT (an evidence accumulation model; see Busemeyer & Townsend, 1993) are more appropriate.

The fact that our instantiation of the LBA assumes weighted comparisons between choice options makes it analogous to attribute–based models of inter–temporal and risky choice. These models assume that preferential choice between options is not necessarily based on underlying delay or probability discounting functions, but it is rather driven by direct comparisons between the attributes of each option (see e.g., Brandstätter et al., 2006; Cheng & González-Vallejo, 2016; González-Vallejo, 2002; Read, Frederick, & Scholten, 2013; Scholten & Read, 2010). Attentional focus or importance placed on each attribute is instantiated by weights. The scaling parameters for amount, delay, and probability in the drift rates of the LBA can be conceived as serving the same purpose (for similar ideas see, Dai & Busemeyer, 2014; Read et al., 2013).

Our modeling analysis also adds to recent attempts that employed evidence accumulation models to account for effects in risky and inter–temporal choice. For example, Dai and Busemeyer (2014) found that an attribute–wise diffusion model, based on absolute comparisons between the dimensions of money and time, could account for three inter–temporal choice effects. Rodriguez et al. (2014) used the LBA in an inter–temporal choice setting and concluded that delayed decision–making can be also explained by sequential sampling mechanisms. Our current work extends these theoretical and practical observations and presents LBA as a model which accounts for inter–temporal and risky decision–making independently but also when the two dimensions combine in a singe choice option. The model fits the combined choice data well (see Figures 3 and 8) without incorporating trade–offs between probability and time (as is required in the probability and time trade–off model Baucells & Heukamp, 2010, 2012), and without assuming any particular functional form for probability and delay discounting (as is required in the multiplicative hyperboloid model; Vanderveldt et al., 2015). In addition, the LBA naturally accounts for response times and implements them in the decision process as an important component of developing a preferential strength for each option. Taking all these facets together, our work presents the first attempt to model the combined effect of probability and delay through an evidence accumulation process, and to provide psychological explanations about preferential choice that rely on the simultaneous examination of choice and RT data.

References

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole Americaine. *Econometrica*, *21*, 503–546. doi: 10.2307/1907921

Appelt, K. C., Hardisty, D. J., & Weber, E. U. (2011). Asymmetric discounting of gains and losses: A query theory account. *Journal of Risk and Uncertainty*, *43*, 107–126. doi: 10.1007/s11166-011-9125-1

Baucells, M., & Heukamp, F. H. (2010). Common ratio using delay. *Theory and Decision*, *68*, 149–158. doi: 10.1007/s11238-008-9130-2

Baucells, M., & Heukamp, F. H. (2012). Probability and time trade-off. *Management Science*, *58*, 831–842. doi: 10.1287/mnsc.1110.1450

Bhatia, S., & Mullett, T. L. (2016). The dynamics of deferred decision. *Cognitive Psychology*, *86*, 112–151. doi: 10.1016/j.cogpsych.2016.02.002

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, *113*, 409–432. doi: 10.1037/0033-295X.113.2.409

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*, 153–178. doi: 10.1016/j.cogpsych.2007.12.002

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*, 432–459. doi: 10.1037/0033-295X.100.3.432

Cassey, P., Gaut, G., Steyvers, M., & Brown, S. D. (2015). *A generative joint model for spike trains and saccades during perceptual decision making.* Manuscript submitted for publication.

Cassey, P., Heathcote, A., & Brown, S. D. (2014). Brain and behavior in decision-making. *PLoS Computational Biology*, *10*, e1003700. doi: 10.1371/journal.pcbi.1003700

Chapman, G. B., & Weber, B. J. (2006). Decision biases in intertemporal choice and choice under uncertainty: Testing a common account. *Memory & Cognition*, *34*, 589–602. doi: 10.3758/BF03193582

Cheng, J., & González-Vallejo, C. (2016). Attribute-wise vs. aternative-wise mechanism in intertemporal choice: Testing the proportional difference, trade-off, and hyperbolic models. *Decision*, *3*, 190–215. doi: 10.1037/dec0000046

Dai, J., & Busemeyer, J. R. (2014). A probabilistic, dynamic, and attribute-wise model of intertemporal choice. *Journal of Experimental Psychology: General*, *143*, 1489–1514. doi: 10.1037/a0035976

Epper, T., Fehr-Duda, H., & Bruhin, A. (2011). Viewing the future through a warped lens: Why uncertainty generates hyperbolic discounting. *Journal of Risk and Uncertainty*, *43*, 169–203. doi: 10.1007/s11166-011-9129-x

Forstmann, B. U., Anwander, A., Schäfer, A., Neumann, J., Brown, S., Wagenmakers, E.-J., … Turner, R. (2010). Cortico-striatal connections predict control over speed and accuracy in perceptual decision making. *Proceedings of the National Academy of Sciences*, *107*, 15916–15920. doi: 10.1073/pnas.1004932107

Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., Von Cramon, D. Y., Ridderinkhof, K. R., & Wagenmakers, E.-J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences*, *105*, 17538–17542. doi: 10.1073/pnas.0805903105

Gamerman, D., & Lopes, H. F. (2006). *Markov chain Monte Carlo: Stochastic simulation for Bayesian inference*. Boca Raton, FL: Chapman & Hall/CRC Press.

Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (1996). *Markov chain Monte Carlo in practice.* Boca Raton, FL: Chapman & Hall/CRC Press.

Glöckner, A., & Herbold, A.-K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making*, *24*, 71–98. doi: 10.1002/bdm.684

Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, *38*, 129–166. doi: 10.1006/cogp.1998.0710

González-Vallejo, C. (2002). Making trade-offs: A probabilistic and context-sensitive model of choice behavior. *Psychological Review*, *109*, 137–155. doi: 10.1037/0033-295X.109.1.137

Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, *130*, 769–792. doi: 10.1037/0033-2909.130.5.769

Greenhow, A. K., Hunt, M. J., Macaskill, A. C., & Harper, D. N. (2015). The effect of reinforcer magnitude on probability and delay discounting of experienced outcomes in a computer game task in humans. *Journal of the Experimental Analysis of Behavior*, *104*, 186–197. doi: 10.1002/jeab.166

Hawkins, G. E., Marley, A., Heathcote, A., Flynn, T. N., Louviere, J. J., & Brown, S. D. (2014). Integrating cognitive process and descriptive models of attitudes and preferences. *Cognitive Science*, *38*, 701–735. doi: 10.1111/cogs.12094

Ho, T. C., Brown, S., & Serences, J. T. (2009). Domain general mechanisms of perceptual decision making in human cortex. *The Journal of Neuroscience*, *29*, 8675–8687. doi: 10.1523/JNEUROSCI.5984-08.2009

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291. doi: 10.2307/1914185

Keren, G., & Roelofsma, P. (1995). Immediacy and certainty in intertemporal choice. *Organizational Behavior and Human Decision Processes*, *63*, 287–297. doi: 10.1006/obhd.1995.1080

Konstantinidis, E., van Ravenzwaaij, D., & Newell, B. R. (2017). Exploring the decision dynamics of risky intertemporal choice. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th annual conference of the cognitive science society* (pp. 694–699). Austin, TX: Cognitive Science Society.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*, 1292–1298. doi: doi:10.1038/nn.2635

Loewenstein, G., & Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics*, *107*, 573–597. doi: 10.2307/2118482

Luce, R. D. (1986). *Response times.* New York, NY: Oxford University Press.

Myerson, J., & Green, L. (1995). Discounting of delayed rewards: Models of individual choice. *Journal of the Experimental Analysis of Behavior*, *64*, 263–276. doi: 10.1901/jeab.1995.64-263

Myerson, J., Green, L., Scott Hanson, J., Holt, D. D., & Estle, S. J. (2003). Discounting delayed and probabilistic rewards: Processes and traits. *Journal of Economic Psychology*, *24*, 619–635. doi: 10.1016/S0167-4870(03)00005-9

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, *7*, 308–313. doi: 10.1093/comjnl/7.4.308

Newell, B. R., & Bröder, A. (2008). Cognitive processes, models and metaphors in decision research. *Judgment and Decision Making*, *3*, 195–204.

Newell, B. R., Lagnado, D. A., & Shanks, D. R. (2015). *Straight choices: The psychology of decision making* (2nd ed.). Hove, UK: Psychology Press.

Oppenheimer, D. M., & Kelso, E. (2015). Information processing as a paradigm for decision making. *Annual Review of Psychology*, *66*, 277–294. doi: 10.1146/annurev-psych-010814-015148

Pitt, M. A., & Myung, I. J. (2002, oct). When a good fit can be bad. *Trends in Cognitive Sciences*, *6*, 421–425. doi: 10.1016/S1364-6613(02)01964-2

Prelec, D., & Loewenstein, G. (1991). Decision making over time and under uncertainty: A common approach. *Management Science*, *37*, 770–786. doi: 10.1287/mnsc.37.7.770

Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior*, *55*, 233–244. doi: 10.1901/jeab.1991.55-233

Rae, B., Heathcote, A., Donkin, C., Averell, L., & Brown, S. D. (2014). The hare and the tortoise: Emphasizing speed can change the evidence used to make decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1226–1243. doi: 10.1037/a0036801

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108. doi: 10.1037/0033-295X.85.2.59

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, *111*, 159–182. doi: 10.1037/0033-295X.111.1.159

Ratcliff, R., Thapar, A., & McKoon, G. (2006). Aging, practice, and perceptual tasks: A diffusion model analysis. *Psychology and Aging*, *21*, 353–371. doi: 10.1037/0882-7974.21.2.353

Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive Psychology*, *60*, 127–157. doi: 10.1016/j.cogpsych.2009.09.001

Read, D., Frederick, S., & Scholten, M. (2013). DRIFT: An analysis of outcome framing in intertemporal choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 573–588. doi: 10.1037/a0029177

Rodriguez, C. A., Turner, B. M., & McClure, S. M. (2014). Intertemporal choice as discounted value accumulation. *PloS ONE*, *9*, e90138. doi: 10.1371/journal.pone.0090138

Samuelson, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies*, *4*, 155–161. doi: 10.2307/2967612

Scholten, M., & Read, D. (2010). The psychology of intertemporal tradeoffs. *Psychological Review*, *117*, 925–944. doi: 10.1037/a0019619

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, *49*, 11–36. doi: 10.1016/0010-0277(93)90034-S

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society*, *64*, 583–639. doi: 10.1111/1467-9868.00353

Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology*, *53*, 1–26. doi: 10.1016/j.cogpsych.2005.10.003

Thaler, R. H. (1981). Some empirical evidence on dynamic inconsistency. *Economic Letters*, *8*, 201–207. doi: 10.1016/0165-1765(81)90067-7

Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological Review*, *121*, 179–205. doi: 10.1037/a0036137

Turner, B. M., Sederberg, P. B., Brown, S. D., & Steyvers, M. (2013). A method for efficiently sampling from distributions with correlated dimensions. *Psychological Methods*, *18*, 368–384. doi: 10.1037/a0032222

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323. doi: 10.1007/bf00122574

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*, 550–592. doi: 10.1037/0033-295X.108.3.550

van Ravenzwaaij, D., Cassey, P., & Brown, S. D. (2018). A simple introduction to Markov Chain Monte–Carlo. *Psychonomic Bulletin & Review*, *25*, 143–154. doi: 10.3758/s13423-016-1015-8

van Ravenzwaaij, D., Dutilh, G., & Wagenmakers, E.-J. (2012). A diffusion model decomposition of the effects of alcohol on perceptual decision making. *Psychopharmacology*, *219*, 1017–1025. doi: 10.1007/s00213-011-2435-9

van Ravenzwaaij, D., Provost, A., & Brown, S. D. (2017). A confirmatory approach for integrating neural and behavioral data into a single model. *Journal of Mathematical Psychology*, *76*, 131–141.

van Ravenzwaaij, D., van der Maas, H. L. J., & Wagenmakers, E.-J. (2011). Does the name-race implicit association test measure racial prejudice? *Experimental Psychology*, *58*, 271–277. doi: 10.1027/1618-3169/a000093

Vanderveldt, A., Green, L., & Myerson, J. (2015). Discounting of monetary rewards that are both delayed and probabilistic: Delay and probability combine multiplicatively, not additively. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 148–162. doi: 10.1037/xlm0000029

Vlaev, I., Chater, N., Stewart, N., & Brown, G. D. (2011). Does the brain calculate value? *Trends* *in Cognitive Sciences*, *15*, 546–554. doi: 10.1016/j.tics.2011.09.008

Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language*, *58*, 140–159. doi: 10.1016/j.jml.2007.04.006

Weber, B. J., & Chapman, G. B. (2005). The combined effects of risk and time on choice: Does uncertainty eliminate the immediacy effect? Does delay eliminate the certainty effect? *Organizational Behavior and Human Decision Processes*, *96*, 104–118. doi: 10.1016/j.obhdp.2005.01.001

Weber, E. U., Johnson, E. J., Milch, K. F., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric discounting in intertemporal choice: A query-theory account. *Psychological Science*, *18*, 516–523. doi: 10.1111/j.1467-9280.2007.01932.x

Yi, R., de la Piedad, X., & Bickel, W. K. (2006). The combined effects of delay and probability in discounting. *Behavioural Processes*, *73*, 149–155. doi: 10.1016/j.beproc.2006.05.001

Appendix
Distributional Choices

Experiment 1

Starting values for the MCMC chains for individual parameters were drawn from the following distributions: $B_I \sim N(1.2, 0.12)|(0, \infty)$, $B_R \sim N(1, 0.1)|(0, \infty)$, $A_D \sim N(2, 0.2)|(0, \infty)$, $A_R \sim N(1.75, 0.175)|(0, \infty)$, $\nu_{N_0} \sim N(3, 0.3)|(0, \infty)$, $\nu_{D_0} \sim N(3, 0.3)|(0, \infty)$, $\nu_{C_0} \sim N(3.5, 0.35)|(0, \infty)$, $\nu_{R_0} \sim N(3, 0.3)|(0, \infty)$, $\alpha_{N_X} \sim N(0.11, 0.311)|(-3, \infty)$, $\alpha_{N_D} \sim N(0.15, 0.315)|(-3, \infty)$, $\alpha_{D_X} \sim N(0.15, 0.315)|(-3, \infty)$, $\alpha_{D_D} \sim N(0.3, 0.33)|(-3, \infty)$, $\alpha_{C_X} \sim N(0.06, 0.036)|(-3, \infty)$, $\alpha_{C_R} \sim N(0.16, 0.316)|(-3, \infty)$, $\alpha_{R_X} \sim N(0.1, 0.31)|(-3, \infty)$, $\alpha_{R_R} \sim N(0.4, 0.34)|(-3, \infty)$, and $t0 \sim N(0.15, 0.015)|(0, \infty)$. In case of any of the 4-A models, starting values for both the now/delayed and the certain/risky starting points were drawn from the same distribution as indicated above. In case of the non-linear models, β_{N_X}, β_{N_D}, β_{D_X}, $\beta_{D_D} \sim N(-0.3, 0.27)|(-3, \infty)$ and β_{C_X}, β_{C_R}, β_{R_X}, $\beta_{R_R} \sim N(0, 0.3)|(-3, \infty)$. The notation $\sim N(,)$ indicates that values were drawn from a normal distribution with mean and standard deviation parameters given by the first and second number between parentheses, respectively. The notation $|(,)$ indicates that the values sampled from the normal distribution were truncated between the first and second numbers in parentheses.

The hierarchical set–up prescribes that all individual parameters come from a truncated Gaussian group–level distribution (truncated to positive values only). Thus, for each parameter to be estimated, we are estimating a group level mean parameter and a group level standard deviation parameter. All group level mean parameters are normally distributed, both B_μs $\sim N(1, 0.3)|(0, \infty)$, $A_D\mu \sim N(2, 1)|(0, \infty)$, $A_R\mu \sim N(1.75, 1)|(0, \infty)$, all ν_μs $\sim N(3, 1.5)|(0, \infty)$, all α_μs $\sim N(0, 1)|(0, \infty)$, and $t0_\mu \sim N(0.15, 0.07)|(0, \infty)$. In case of any of the 4-A models, both the now/delayed and the certain/risky options had the same prior for starting point as indicated above. In case of the nonlinear models, all β_μs $\sim N(0, 1)|(-3, \infty)$. All group level standard deviation parameters are gamma distributed, with a shape and a scale parameter of 1. Starting values for the MCMC chains for group level μ parameters were drawn from the same distributions as those for the individual parameters, and starting values for group level σ parameters were derived from starting value distributions for the individual parameters by dividing the mean by 10 and the standard deviation by 2. These prior settings are quite uninformative, and are based on previous experience with parameter estimation for the LBA model. As a result, the specific settings will not have a large influence on the shape of the posterior. For more details on distributional choices for the priors, we refer the reader to Turner et al. (2013).

For sampling, we used 32 interacting Markov chains and ran each for 1,000 burn-in iterations followed by 1,000 iterations after convergence. The two tuning parameters of the differential evolution proposal algorithm were set to standard values used in previous work: random perturbations were added to all proposals drawn uniformly from the interval [-.001, .001]; and the scale of the difference added for proposal generation was set to $\gamma = 2.38 \times (2K)^{-0.5}$, where K is the number of parameters per participant. The MCMC chains blocked proposals separately for each participant's parameters, and also blocked the group–level parameters in $\{\mu, \sigma\}$ pairs.

Experiment 2

Starting values for the MCMC chains for individual parameters were drawn from the following distributions: $B \sim N(2, 0.2)|(0, \infty)$, both As $\sim N(2, 0.2)|(0, \infty)$, $\nu_{NC_0} \sim N(6, 0.6)|(0, \infty)$, $\nu_{DR_0} \sim N(6, 0.6)|(0, \infty)$, $\alpha_{NC_X} \sim N(0.1, 0.01)|(0, \infty)$, $\alpha_{NC_D} \sim N(0.15, 0.015)|(0, \infty)$, $\alpha_{NC_R} \sim N(0.2, 0.02)|(0, \infty)$, $\alpha_{DR_X} \sim N(0.1, 0.01)|(0, \infty)$, $\alpha_{DR_D} \sim N(0.15, 0.015)|(0, \infty)$, $\alpha_{DR_R} \sim N(0.2, 0.02)|(0, \infty)$, and $t0 \sim N(0.2, 0.02)|(0, \infty)$.

Priors for all individual parameters are normally distributed, $B \sim N(2, 2)|(0, \infty)$, both As $\sim N(2, 2)|(0, \infty)$, $\nu_{NC_0} \sim N(6, 6)|(0, \infty)$, $\nu_{DR_0} \sim N(6, 6)|(0, \infty)$, $\alpha_{NC_X} \sim N(0.1, 0.1)|(0, \infty)$, $\alpha_{NC_D} \sim N(0.15, 0.15)|(0, \infty)$, $\alpha_{NC_R} \sim N(0.2, 0.2)|(0, \infty)$, $\alpha_{DR_X} \sim N(0.1, 0.1)|(0, \infty)$, $\alpha_{DR_D} \sim N(0.15, 0.15)|(0, \infty)$, $\alpha_{DR_R} \sim N(0.2, 0.2)|(0, \infty)$, and $t0 \sim N(0.2, 0.2)|(0, \infty)$.

For sampling, we used 32 interacting Markov chains and ran each for 1,000 burn-in iterations followed by 1,000 iterations after convergence. The two tuning parameters of the differential evolution proposal algorithm were set to standard values used in previous work: random perturbations were added to all proposals drawn uniformly from the interval [-.001, .001]; and the scale of the difference added for proposal generation was set to $\gamma = 2.38 \times (2K)^{-0.5}$, where K is the number of parameters per participant. The MCMC chains blocked proposals separately for each participant's parameters, and also blocked the group–level parameters in $\{\mu, \sigma\}$ pairs.