Psychological Review

An Evaluation and Comparison of Models of Risky Intertemporal Choice

Ashley Luckman, Chris Donkin, and Ben R. Newell

Online First Publication, July 23, 2020. http://dx.doi.org/10.1037/rev0000223

CITATION

Luckman, A., Donkin, C., & Newell, B. R. (2020, July 23). An Evaluation and Comparison of Models of Risky Intertemporal Choice. *Psychological Review*. Advance online publication. http://dx.doi.org/10.1037/rev0000223

American Psychological Association

Psychological Review

© 2020 American Psychological Association
ISSN: 0033-295X

http://dx.doi.org/10.1037/rev0000223

An Evaluation and Comparison of Models of Risky Intertemporal Choice

Ashley Luckman
University of New South Wales and University of Warwick

Chris Donkin and Ben R. Newell
University of New South Wales

Risky intertemporal choices involve choosing between options that can differ in outcomes, their probability of receipt, and the delay until receipt. To date, there has been no attempt to systematically test, compare, and evaluate theoretical models of such choices. We contribute to theory development by generating predictions from 7 models for 3 common manipulations—magnitude, certainty, and immediacy—across 6 different types of risky intertemporal choices. Qualitative and quantitative comparisons of model predictions to data from an experiment involving almost 4,000 individual choices revealed that an attribute comparison-model, newly modified to incorporate risky intertemporal choices, (the risky intertemporal choice heuristic or RITCH) provided the best account of the data. Results are consistent with growing evidence in support of attribute comparison models in the risky and intertemporal choice literatures, and suggest that the relatively poorer fits of translation-based models reflect their inability to predict the differential impact of certainty and immediacy manipulations. Future theories of risky intertemporal choice may benefit from treating risk and time as independent dimensions, and focusing on attribute-comparison rather than value-comparison processes.

Keywords: risky choice, intertemporal choice, risky intertemporal choice, cognitive modeling, hierarchical Bayesian modeling

Supplemental materials: http://dx.doi.org/10.1037/rev0000223.supp

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Three elementary truths of human choice are that people tend to prefer certain rewards over risky ones, sooner rewards over later ones, and larger rewards over smaller ones. Such choices, in which one option dominates, are often straightforward. It is when the dimensions of risk, time, and magnitude combine that choices become difficult (Vanderveldt, Green, & Myerson, 2015). How does one choose between $100 with certainty today and a 25% chance of $600 in 3 months? More generally how does one choose between two options which differ in outcome (i.e., $100 vs. $600), probability of receiving those outcomes (i.e., 1.0 vs. 0.25), and delay until receipt of those outcomes (i.e., 0 months vs. 3 months).

This class of problems—known as *risky intertemporal* choices has attracted growing interest in the literature (Baucells & Heukamp, 2010, 2012; Keren & Roelofsma, 1995; Konstantinidis, van Ravenzwaaij, Güney, & Newell, 2020; Luckman, Donkin, & Newell, 2017; Sun & Li, 2010; Vanderveldt et al., 2015; Weber & Chapman, 2005a). This interest is unsurprising given that choices among risky intertemporal prospects allow a unique opportunity for the development and expansion of theories that have typically only considered the effects of risk and time in isolation. In this article we will focus on two areas of theory development where studying risky intertemporal choice may be particularly informative. The first is the development of theories of how risk and delay relate to each other in choice. Prelec and Loewenstein (1991) noted several commonalities between the effects of risk and time on choice. They argued that these choice-patterns reflect fundamental psychological properties of prospect evaluation that operate when there is uncertainty about *either* timing of prospect-receipt *or* probability of prospect-receipt. Risky intertemporal choices allow us to investigate how these psychological properties manifest when those two types of uncertainty compete for attention. The second area in which risky intertemporal choices may be particularly useful is in providing insight into how people process and integrate information across multiple attribute dimensions (i.e., risks, delay, amounts) and multiple options. There is an ongoing debate in both the risky choice and intertemporal choice literatures as to whether people make decisions primarily by comparing individual attribute values across risky/intertemporal options—that is, comparing risks to risks, amounts to amounts, and delays to delays—or whether they compare risky/intertemporal options by calculating some form of overall value or utility of each option, by

Ashley Luckman, School of Psychology, University of New South Wales, and Warwick Business School, University of Warwick; Chris Donkin and Ben R. Newell, School of Psychology, University of New South Wales

This work was supported by the Australian Research Council (Grants FT110100151, DE130100129, DP140101145, DP160101186, DP190101675), and the Economic and Social Research Council (Grant ES/P008976/1). Earlier versions of this work have been presented at the Society for Judgement and Decision Making Annual Meeting (Boston, 2016), the Australasian Mathematical Psychology Conference (Hobart, 2016), and appeared as a Chapter in AL's Thesis (UNSW, Australia, 2016). These versions included the same experimental data and model fitting. The experimental data can be accessed from https://osf.io/jsdtx/.

Correspondence concerning this article should be addressed to Ashley Luckman, Warwick Business School, University of Warwick, Scarman Road, Coventry, CV4 7AL, United Kingdom. E-mail: ashley.luckman@wbs.ac.uk

combining information across the attribute dimensions (Cheng & González-Vallejo, 2016; Vlaev, Chater, Stewart, & Brown, 2011). Because risky intertemporal choices allow more complex trade-offs between amounts and risks or amounts and delays than are possible in choices that involve only risks or only delays, they generate novel circumstances in which mechanisms that produce the same effects in standard risky or intertemporal choices, produce divergent behaviors.

To address these theoretical questions in this article we perform a comprehensive comparison of seven different models of risky intertemporal choice which utilize different mixes of attribute and utility comparisons. The models are compared based on the predictions they make for three prominent manipulations from the literature—magnitude, immediacy, and certainty—across different types of risky intertemporal choices. The remainder of the introduction is structured as follows: First we explain the distinction between models which assume that utilities are compared versus models assuming that attributes are compared. We then outline how risky intertemporal choice may be useful for discriminating between utility- and attribute-based accounts of the effect of outcome magnitude. Second, we introduce the immediacy and certainty effect, and summarize existing research into risky intertemporal choice focusing on the similarities between risk and delay, and how this research has impacted model development. Third, we perform a comprehensive review of the current research on magnitude, immediacy, and certainty effects for six different types of risky intertemporal choices. Finally, we provide details of the specifications we use for the seven models compared in this article, and outline their predictions.

Attribute Versus Utility Comparisons

In both the risky choice and intertemporal choice literatures, models of choice are often grouped together into three broad classes based upon the way in which information from different attribute dimensions and choice options is integrated (Birnbaum & Lacroix, 2008; Cheng & González-Vallejo, 2016; Reeck, Wall, & Johnson, 2017; Vlaev et al., 2011). The first of these classes, which we call utility comparison models, assume that people choose as if they have calculated the worth, or utility, of each of the options presented in a choice independently. The exact way in which the utility of an option is calculated varies from model to model, but it involves combining the attributes of an option together into a single value, such as by multiplying the outcome by a discount rate based on its delay until receipt, like in the hyperbolic discounting model of intertemporal choice (Kirby, 1997; Kirby & Marakovic, 1995; Mazur, 1987), or multiplying subjective outcomes of a gamble by weights based on their probabilities, like in prospect theory for risky choice (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992).

The second class of models, attribute-comparison models, instead assume that people directly compare attribute values across options, for instance directly comparing the delay of one option to the delay of the other, and the amount of one option to the amount of the other, like in the trade-off model of intertemporal choice (Scholten & Read, 2010). A decision is then made by comparing the differences on each dimension to each other. In other words, while utility models combine the attributes of each choice option to establish the worth of each option, then compare the options, an attribute model compares the two options on each attribute separately, then compares the attribute differences. These two classes of model can also be grouped into a broader class, often called integrative models, as they both integrate information across both attributes and choice options, just in a different order (Birnbaum & Lacroix, 2008).

The third class of models, called heuristic or lexicographic semiorder models, assume that information is not integrated across attribute dimensions, the options are compared on each attribute sequentially and independently, with a decision usually being made when any single attribute difference reaches some threshold (Brandstätter, Gigerenzer, & Hertwig, 2006). Support for models of this type is somewhat mixed in the literature, with several model comparison studies showing poor support for them, so we do not consider them further (Birnbaum & Lacroix, 2008; Rieskamp, 2008).

To date, all proposed models of risky intertemporal choice belong to the class of utility-comparison models. These models—the multiplicative hyperboloid discounting model (Vanderveldt et al., 2015), the probability and time trade-off model (Baucells & Heukamp, 2010, 2012), and a modified version of the Rachlin's hyperbolic discounting model that was used by Yi, de la Piedad, and Bickel (2006)—all build on existing utility-comparison models of risky or intertemporal choice such as prospect theory or hyperbolic discounting models. However, there is growing evidence that attribute comparison models such as the intertemporal choice heuristic (ITCH), proportional difference (PD), or trade-off models, provide better accounts of intertemporal choice behavior than classic utility based models (Cheng & González-Vallejo, 2016; Dai & Busemeyer, 2014; Ericson, White, Laibson, & Cohen, 2015; Scholten, Read, & Sanborn, 2014; but see Wulff & van den Bos, 2018). Given their success in explaining intertemporal choice data, it seems likely that attribute-comparison models may also be useful in explaining risky intertemporal choice. Furthermore, as we demonstrate in the next section, considering attribute-comparison processes in risky intertemporal choice may increase our understanding of the comparison processes underlying choice in risk-only or delay-only choices.

Magnitude Effects

A standard intertemporal choice (e.g., $50 now vs. $100 in 6 months) involves a trade-off between time and amount, as an individual chooses between a smaller sooner and a larger later outcome. A robust finding in these types of intertemporal choices is that if you increase the magnitude of the outcomes of both options by multiplying them by a common multiplier, people become much more likely to choose the larger later option (Chapman & Weber, 2006; Green & Myerson, 2004; Myerson, Green, Hanson, Holt, & Estle, 2003; Thaler, 1981; Vanderveldt et al., 2015; Vanderveldt, Green, & Rachlin, 2017). For example, people show a greater preference to wait when choosing between $500 now and $1,000 in 6 months, compared with a choice between $50 now and $100 in 6 months. This intertemporal choice effect is referred to as the magnitude effect.

One issue that arises when trying to discriminate between attribute- and utility-comparison processes is that both classes of models can capture effects like the magnitude effect, albeit using different mechanisms. For example utility-comparison models,

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

like probability and time trade-off (PTT), often assume that the rate at which people discount delayed outcomes decreases as a function of the amount of the outcome (Baucells & Heukamp, 2010; Myerson et al., 2003; Vincent, 2016). In essence PTT assumes that time delays are scaled by the amount offered, so the same delay is treated as shorter the larger the outcome is. This type of explanation, where one attribute is allowed to affect the value/processing of another, is consistent with the underlying assumption of utility-comparison models, that the delay and amount information is integrated within each option. Attribute-comparison models, like the ITCH model, do not explain the magnitude effect in the same way, because such models assume that each attribute dimension is processed separately, and are not integrated. Instead, in the case of the ITCH model, the magnitude effect is captured by assuming that people consider, as part of their decision, the absolute difference in outcome between the two options (Ericson et al., 2015). Now, when both outcomes are multiplied by a constant to produce the magnitude effect, the difference between the outcomes increases by the same constant, thus making the larger outcome relatively more attractive.

In a classic intertemporal choice the larger outcome is also the later outcome, as a consequence the two mechanisms in utility- and attribute-based models will produce the same qualitative effect, and so cannot be easily distinguished. However, in risky intertemporal choices this confound between the larger and later outcome does not exist. For instance, consider giving a participant a choice between a smaller later but safer outcome (e.g., $50 in 6 months for certain) and a larger sooner but riskier outcome ($100 now with probability 0.5). Because the PTT model assumes that the magnitude effect is caused by the delays being treated as shorter, multiplying both outcomes by 10 would predict that people shift toward preferring the smaller later safer outcome ($500 in 6 months for certain). Conversely, because the magnitude effect in the ITCH model is caused by an increased difference in amount between the two options, it instead predicts a shift toward the larger sooner riskier outcome ($1,000 now with probability 0.5). Therefore, risky intertemporal choices provide a unique opportunity to discriminate between these two types of models because attribute-comparison models, by definition, do not allow manipulation of one attribute to change the processing of another attribute.

However, we should note that utility-comparison models can produce effects of magnitude that occur purely by changing how amounts are processed. For example, consider a standard risky choice between a smaller but safer reward, and a larger but riskier reward (e.g., $50 for certain or $100 with probability 0.5). In these types of risky choices multiplying both amounts by a common multiplier (e.g., 10) has generally been shown to make people more likely to choose the safer option (e.g., $500 for certain; Green & Myerson, 2004; Holt & Laury, 2002; Markowitz, 1952; Myerson et al., 2003; Weber & Chapman, 2005b), although not always (Chapman & Weber, 2006; Vanderveldt et al., 2017). This is often referred to as the peanuts effect, as people are more willing to gamble when they are playing for peanuts, that is, small amounts. The PTT model can also capture this effect, but not by assuming that outcome amounts change how probabilities are processed. Instead the PTT model assumes that when the utility of an outcome is calculated, objective outcomes (e.g., x) are transformed into subjective values ($v(x)$) by a concave value function with decreasing elasticity (Baucells & Heukamp, 2010, 2012). Decreasing elasticity means that proportional increases in outcome magnitude, x, lead to smaller proportional increases in subjective value, $v(x)$, for larger values of x than for smaller values of x (Scholten & Read, 2014). If we multiply both outcomes by a constant, for example, 10, the subjective value of the smaller outcome, $50, will change by a greater multiplier than will the subjective value of the larger outcome, $100, leading to the smaller safer outcome becoming relatively more attractive when the outcome magnitudes are increased.

Taken together, exploring the effects of outcome magnitude in risky intertemporal choices has the potential to challenge the accounts of these effects provided by attribute-comparison models, such as ITCH or the trade-off model. If we find a pattern of magnitude effects consistent with people becoming more or less willing to wait, or more or less willing to take a risk, this will rule out an attribute-comparison explanation for the effect. If we instead find patterns consistent with a general shift either toward or away from preferring the larger option, regardless of risk or delay, this would be compatible with either attribute or utility comparison processes.

Relation Between Risk and Delay

While the distinction between attribute- and utility-comparison processes has generally not been investigated in risky intertemporal choices, the same is not true of the other area of focus of this article—the relationship between risk and delay. Empirical investigations into risky intertemporal choices have generally explicitly focused on exploring the similarities between risk and delay as a first step to developing theories of risky intertemporal choice. Of particular interest has been the relationship between the certainty effect in risky choice, and the immediacy effect in intertemporal choice. The certainty effect in risky choice—also called the common ratio effect—refers to the change in behavior observed when the probabilities of the outcomes presented in a risky choice are reduced by a common ratio (Baucells & Heukamp, 2010; Kahneman & Tversky, 1979). In general, people are more risk seeking when the probability of each outcome is reduced (by the same amount). For example, if most people prefer to take a certain $50 over $100 with probability 0.5, then fewer people prefer $50 with probability 0.2 over $100 with probability 0.1, that is, when the probabilities are divided by 5. The immediacy effect is often considered analogous to the certainty effect, but for intertemporal choice (Prelec & Loewenstein, 1991). Just as the certainty effect involves reducing probability by a common amount, the immediacy effect—also called the common difference effect[1]—involves increasing delays by a common amount, which leads to an increase in preference for the larger more delayed outcome (Kirby, 1997; Kirby & Marakovic, 1995; Loewenstein & Prelec, 1992: Thaler, 1981). For example, if most people prefer to take $50 now, rather

[1] Some researchers distinguish between the immediacy and common difference effects, with the former only applying to situations where one of the options in the initial choice set was immediate, while the latter applies to situations where both options in the original set were delayed. Similar distinctions are made between the certainty and common ratio effects (Prelec & Loewenstein, 1991). We make no such distinction and use immediacy and certainty throughout, as they refer directly to the dimensions we manipulate.

than wait 6 months for $100, fewer would prefer to take $50 in 12 months over $100 in 18 months.

The similarity between these two effects has led several researchers to test the effects of certainty on intertemporal choices, and vice versa. For instance Keren and Roelofsma (1995), among others (Weber & Chapman, 2005a), found that adding risk to an intertemporal choice, had a similar effect to adding a common delay in that participants became more likely to choose the larger later option. However, not all studies have found this relationship (Sun & Li, 2010; Weber & Chapman, 2005a). Similarly, a variety of studies have found that adding a common delay to a risky choice has a similar effect to increasing the risk, in that participants became more likely to choose the larger riskier option (Baucells & Heukamp, 2010; Oshikoji, 2012; Sagristano, Trope, & Liberman, 2002; Weber & Chapman, 2005a). Again, however, not all studies find this effect (see Weber & Chapman, 2005a).

More generally several researchers have hypothesized that intertemporal choice behavior may be driven, partially or exclusively, by considerations of risk (Epper, Fehr-Duda, & Bruhin, 2011; Sozou, 1998; Takahashi, Ikeda, & Hasegawa, 2007). For instance, Takahashi et al. (2007) had participants directly estimate the risks associated with various delayed outcomes and found that these risk estimates can partially explain intertemporal behavior. Alternatively, risky choice behavior may be driven by considerations of delay, as events that are unlikely can also be thought of in terms of when they would be expected to occur, for instance a one in 100-year flood (Rachlin, 1989; Rachlin, Raineri, & Cross, 1991). Regardless of the direction, these studies suggest that risk and delay may be treated as though they are equivalent. Two of the existing models of risky intertemporal choice, the PTT model (Baucells & Heukamp, 2010, 2012) and the hyperbolic discounting (HD) model of Yi et al. (2006), have explicitly included this concept. The PTT model assumes that participants first translate time delays into risks and then treat their decision like a standard risky choice. In the HD model, risk is translated into expected delays until payment and the decision is treated like a standard intertemporal choice.

However, the evidence that risks and delays are treated like a single attribute dimension is far from conclusive. In addition to the several studies above which do not show immediacy and certainty having the same impact on choice, a recent study found that people prefer delaying an outcome to taking a risk when the two are in direct competition, despite giving the risky and delayed outcome similar values in isolation (Luckman et al., 2017). In order to better understand how risk and delay are related, in this article we examine the effects of immediacy and certainty manipulations across a variety of different type of risky intertemporal choices.

Risky Intertemporal Choice

To understand how immediacy, certainty, and magnitude manipulations impact behavior in risky intertemporal choices, we need to acknowledge that there is no single type of risky intertemporal choice. While all risky choices, at least for single outcome gambles, involve a trade-off between probability and amount, and all intertemporal choices a trade-off between delay and amount, risky intertemporal choices can involve either, both, or neither of these trade-offs. For instance, in risky intertemporal choices there can also/instead be trade-offs between delay and probability (e.g., safer or sooner), or choices where the amount and the probability or delay are complementary, rather than in competition, such as in the example we gave in the section on magnitude effects. In this article we consider six different types of risky intertemporal choices. We define these types based upon which of the three attributes: Risk/probability, time/delay, and amount/outcome, are in competition with each other in a choice. Table 1 provides the structure and an illustrative example of each of the six choice types which we explain in more detail below. For each of these choice types we also summarize current studies on immediacy, certainty, and magnitude effects in that choice type.

Type 1: Risk vs. Amount (RvA). The first type of risky intertemporal choice we consider involves a choice between a smaller safer outcome and a larger riskier outcome, just like a standard risky choice. However, unlike a typical risky choice instead of both options occurring now, both occur after some common delay (e.g., $50 in 6 months with probability 0.8 or $100 in 6 months with probability 0.5). Typical risky choices could therefore be considered a special subset of this type of risky intertemporal choice.

To some extent all three of the effects listed above have been explored in this choice type. In Baucells and Heukamp (2010) participants were given various Risk vs. Amount choices involving different outcomes, probability levels, and common delays. Even ignoring those that could be classed as standard risky choices, in general they found that more people preferred the safer/certain option when the outcome amounts were increased—the peanuts effect—and more people preferred the riskier option when either the probabilities were reduced, the certainty effect, or the common delay was longer, the immediacy effect. The effect of common delay in Risk vs. Amount choices has also been found in other studies (Oshikoji, 2012; Sagristano et al., 2002), although not always reliably, with Weber and Chapman (2005a) finding the effect in a choice titration task, but not when using a single choice design like that in Baucells and Heukamp (2010).

Table 1
Example of Each of the Six Types of Risky Intertemporal Choices Used in This Article, and the Attributes Which Are in Competition in Each

Abbreviation	Attribute trade-off	Example
RvA	Risk vs. Amount	$50 in 6 months with probability 0.8 or $100 in 6 months with probability 0.5
DvA	Delay vs. Amount	$50 in 1 month with probability 0.5 or $100 in 6 months with probability 0.5
DvR	Delay vs. Risk	$50 in 1 month with probability 0.5 or $50 in 6 months with probability 0.8
DRvA	Delay & Risk vs. Amount	$50 in 1 month with probability 0.8 or $100 in 6 months with probability 0.5
RvAD	Risk vs. Amount & Delay	$50 in 6 months with probability 0.8 or $100 in 1 month with probability 0.5
DvAR	Delay vs. Amount & Risk	$50 in 1 month with probability 0.5 or $100 in 6 months with probability 0.8

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Type 2: Delay vs. Amount (DvA). Just as the first type of risky intertemporal choice contains standard risky choice as a subset, the second contains standard intertemporal choice. Like an intertemporal choice, Delay vs. Amount choices involve a choice between a smaller sooner outcome, and a larger later outcome, however both can be occurring with some common probability, rather than only with certainty (e.g., $50 in 1 month with probability 0.5 or $100 in 6 months with probability 0.5). Keren and Roelofsma (1995) looked at both certainty and immediacy effects in such a design. They found the usual immediacy effect observed in intertemporal choices, with participants preferring the larger later option more if a common 26-week delay was added to both options. They also found a certainty effect, with participants preferring to take the larger later option more if the common probability was lower.

Weber and Chapman (2005a) had the same issues replicating the certainty effect in Delay vs. Amount choices as they did the immediacy effect in Risk vs. Amount choices. Furthermore, Sun and Li (2010) found the opposite effect of certainty with the smaller sooner option, rather than the larger later option, becoming more attractive if the probabilities were smaller.

Type 3: Delay vs. Risk (DvR). There are also risky intertemporal choices where the amount is constant across options. These choices are between a riskier sooner outcome and a safer later outcome (e.g., $50 now with probability 0.5 or $50 in 6 months for certain). Magnitude effects have been investigated in the subset of these choices that involve a certain later option and a risky immediate option, as shown in the example. In such choices, increasing the magnitude of the outcomes results in either stronger preference for the certain later option, a willingness to wait longer for the delayed outcome, or a reduction in tolerance for risk when choosing the risky outcome—depending upon the method used (Baucells & Heukamp, 2010; Christensen, Parker, Silberberg, & Hursh, 1998; Luckman et al., 2017). All three effects are in the same direction, and all are consistent with both the magnitude effect as observed in intertemporal choice, and the peanuts effect as observed in risky choice.

Type 4: Delay & Risk vs. Amount (DRvA). The only other choice type in which any of these effects has been investigated are choices between a smaller safer sooner outcome and a larger riskier later outcome (e.g., $50 now for certain or $100 in 6 months with probability 0.5). Vanderveldt, Green, and Myerson (2015) looked at the effects of magnitude in a specific subset of such choices, where the smaller safer sooner option is both certain and immediate, as it is in our example. In general, they found very little effect of magnitude, although there were two instances in which people behaved differently when the magnitude of the outcomes was increased. First, when the larger riskier later option was made certain, thus creating a simple intertemporal choice, they found the usual magnitude effect in intertemporal choice, with participants valuing the larger later option relatively more than the smaller sooner option as outcome magnitudes increased. However, this relative increase in value for the larger later option with increasing outcome magnitudes was not present when the riskier option was anything but certain. Second, if the later option involved a relatively small delay, that is, 1 month or 6 months (compared with delays of 2 or 5 years), then increasing the amount led to preferences in the same direction as the peanuts effect, with participants valuing the larger riskier later option relatively less.

In contrast to Vanderveldt et al. (2015); Yi et al. (2006) found that increasing the magnitude of all outcomes increased the attractiveness of the larger riskier later option. Showing an increased preference for the later option is consistent with the magnitude effect in intertemporal choice; however, because the more delayed option was also riskier, an increased preference for such a gamble is inconsistent with the peanuts effect of risky choice.

Types 5 and 6: Risk vs. Amount & Delay (RvAD); Delay vs. Amount & Risk (DvAR). We know of no studies that have investigated these final two choice types. The first type involves a choice between a smaller safer later option and a larger riskier sooner option (Risk vs. Amount & Delay; e.g., $50 in 6 months for certain or $100 now with probability 0.5). This choice type is particularly interesting for discriminating between accounts of the magnitude effect which assume it is due to changes in attitude to delay, versus those that base the magnitude effect on amount effects. The second is between a smaller riskier sooner option and a larger safer later option (Delay vs. Amount & Risk; e.g., $50 now with probability 0.5 or $100 in 6 months for certain). As we will show, existing models do make predictions for these types of choices, but their predictions are yet to be tested against empirical data.

Risky Intertemporal Choice Models

In addition to the three existing models of risky intertemporal choice—multiplicative hyperboloid discounting model (MHD; Vanderveldt et al., 2015), PTT (Baucells & Heukamp, 2010, 2012), and HD (Yi, de la Piedad, & Bickel, 2006)—we consider four additional models based on extensions of popular models of risky or intertemporal choice. The first of these is a simple combination of prospect theory and hyperbolic discounting. These are, respectively, the most popular models of risky choice and intertemporal choice in the literature, and all three existing risky intertemporal choice models include modified components of one of them. As such we include this model as a baseline comparison model, as it involves a relatively straightforward combination of popular risky and intertemporal models, with no consideration of the risky intertemporal choice literature.

The next three models we consider all involve some element of an attribute-comparison process. The first is a hybrid model proposed by Scholten and Read (2014). This model is an extension of their trade-off model of intertemporal choice (Scholten & Read, 2010) to allow it to capture risky choice. We separate this from the other three existing models of risky intertemporal choice as unlike them, it has never been fit to data. The second attribute-based model is an extension of the ITCH model (Ericson et al., 2015) to account for risky intertemporal choice. We call this model the risky intertemporal choice heuristic (RITCH) model. We include this model due to the recent success of ITCH in a comparison with other popular models of intertemporal choice (Ericson et al., 2015, although see Wulff & van den Bos, 2018 for a conflicting perspective). The third model is a version of the PD model (González-Vallejo, 2002). Other versions of this model have previously been used for both risky choice (González-Vallejo, 2002), and for intertemporal choices (Cheng & González-Vallejo, 2016; Dai & Busemeyer, 2014), but not risky intertemporal choice.

In the following section, we formally define the seven models we will compare. We will also outline the predictions each model

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

makes about the effect that immediacy, certainty, and magnitude manipulations will have in each of the six types of risky intertemporal choices. The qualitative predictions each model makes for the immediacy, certainty, and magnitude effects are given in Tables 2 to 4, respectively. By examining these predictions across all our candidate models, we build up a detailed picture of how and why (or why not) each model can explain the three signature effects. Mapping this landscape provides important insights into how people trade off risk, time, and reward.

Utility-comparison models. Utility models assume that decisions are made by comparing the utility, or worth, of each option presented. This is achieved by first calculating the utility of each option, then choosing the option with the greater utility, with some probability (Stott, 2006). Consider an option, $g(x, p)$, consisting of an outcome, x, occurring with some probability, p. According to prospect theory, the utility of this option, $U(x, p)$, is the product of a value function, $v(x)$, and a probability weighting function, $w(p)$, that is, $U(x, p) = v(x) \cdot w(p)$. Here, the value function is a means of transforming objective outcomes into subjective values, while the probability weighting function transforms objective probabilities into decision weights. Different specifications are used for these two functions throughout the literature, but the former is always a concave function, such as

$$v(x) = x^a \tag{1}$$

where $0 \leq a < 1$, and the later is always an s-shaped function, such as

$$w(p) = e^{-(-\ln p)^r} \tag{2}$$

where $0 < r \leq 1$ (Rieskamp, 2008; Stott, 2006; Tversky & Kahneman, 1992). Decreasing a increases the concavity of the value function, or the extent to which participants show diminishing sensitivity to increasing amounts, while decreasing r increases the subproportionality of probability weighting, or the degree to which small probabilities are overweighted and medium to large probabilities underweighted. When $a = 1$ the subjective value of an outcome is equal to its objective value and when $r = 1$ decision weights are equal to objective probabilities. Details on the psychological meaning and interpretation of all parameters for all models are contained in Appendix B.

Once the utility of each option has been calculated, the difference between the utilities is used to make a decision. While this decision was historically considered to be deterministic, with the option with the greater utility always chosen, it is generally now accepted that choice is stochastic, with the difference in utilities instead translated into a probability of choosing one option over the other (Hey, 1995; Loomes & Sugden, 1995). The probability of choosing the first gamble, g_1, is obtained by passing the difference in utilities through a choice function such as;

$$P(g_1 \mid \{g_1, g_2\}) = \frac{1}{1 + e^{-s(U(g_1) - U(g_2))}} \tag{3}$$

where the sensitivity parameter, $s \geq 0$ determines how deterministic choice is. Note that most choice functions used in the literature assume that the probabilistic decision rule is based on the absolute difference in utilities (Stott, 2006).

Table 2
Predicted Shift in Preferences Caused by the Magnitude Manipulation for Each Choice Type

Model	RvA†	DvA†	DvR†	RvAD†	DvAR†	DRvA†
PTHD	No Change	No Change	No Change	No Change	No Change	No Change
MHD	Smaller-**Safer**	No Change	**Safer**-Later	Smaller-**Safer**-Later	Larger-**Safer**-Later	Smaller-**Safer**-Sooner
HD	No Change	No Change	No Change	No Change	No Change	No Change
PTT	**Smaller**-Safer	*Larger*-**Later**	Safer-**Later**	**Smaller**-Safer-**Later**	*Larger*-Safer-**Later**	*Larger*-Riskier-**Later**
Trade-off	**Smaller**-Safer	*Larger*-**Later**	Safer-**Later**	**Smaller**-Safer-**Later**	*Larger*-Safer-**Later**	*Larger*-Riskier-**Later**
RITCH	**Larger**-Riskier	**Larger**-Later	No Change	**Larger**-Riskier-Sooner	**Larger**-Safer-Later	**Larger**-Riskier-Later
PD	No Change	No Change	No Change	No Change	No Change	No Change
Data	Nonsig: ($p = 1$, $p_{adj} = 1$)	Larger-Later ($p < .001$, $p_{adj} < .001$)	Safer-Later ($p < .001$, $p_{adj} < .001$)	Smaller-Safer-Later ($p < .001$, $p_{adj} < .001$)	Larger-Safer-Later ($p < .001$, $p_{adj} < .001$)	Nonsig: ($p = .801$, $p_{adj} = 1$)

Note. The predictions of each model are on a separate row. Words in bold indicate the attribute driving the effect. Words in italics indicate an attribute that predicts a shift in the opposite direction. The data row shows the observed direction in which participants' preferences shifted (on the group level) in the experiment, and whether this was significant according to a logistic general linear mixed effect model explained in the Results. PTHD = prospect theory and hyperbolic discounting; MHD = multiplicative hyperboloid discounting; HD = hyperbolic discounting; PTT = probability-time-trade-off; RITCH = risky intertemporal choice heuristic; PD = proportional difference.
† R = risk; D = delay; A = amount.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 3
Predicted Shift in Preferences Caused by the Immediacy Manipulation for Each Choice Type

Model	RvA†	DvA†	DvR†	RvAD†	DvAR†	DRvA†
PTHD	No Change	Larger-**Later**	Safer-**Later**	Smaller-Safer-**Later**	Larger-Safer-**Later**	Larger-Riskier-**Later**
MHD	No Change	Larger-**Later**	Safer-**Later**	Smaller-Safer-**Later**	Larger-Safer-**Later**	Larger-Riskier-**Later**
HD	Larger-**Riskier**	Larger-**Later**	Same as Certainty*	Same as Certainty*	Same as Certainty*	Larger-**Riskier-Later**
PTT	Larger-**Riskier**	Larger-**Later**	Same as Certainty*	Same as Certainty*	Same as Certainty*	Larger-**Riskier-Later**
Trade-off	No Change	Larger-**Later**	Safer-**Later**	Smaller-Safer-**Later**	Larger-Safer-**Later**	Larger-Riskier-**Later**
RITCH	No Change	Larger-**Later**	Safer-**Later**	Smaller-Safer-**Later**	Larger-Safer-**Later**	Larger-Riskier-**Later**
PD	No Change	Larger-**Later**	Safer-**Later**	Smaller-Safer-**Later**	Larger-Safer-**Later**	Larger-Riskier-**Later**
Data	Nonsig: Larger-Riskier ($p = .144$, $p_{adj} = .144$)	Larger-Later ($p < .001$, $p_{adj} < .001$)	Safer-Later ($p = .005$, $p_{adj} = .024$)	Nonsig: Smaller-Safer-Later ($p = .057$, $p_{adj} = .114$)	Larger-Safer-Later ($p = .009$, $p_{adj} = .034$)	Nonsig: Larger-Riskier-Later ($p = .024$, $p_{adj} = .071$)

Note. The predictions of each model are on a separate row. Words in bold indicate the attribute driving the effect. The data row shows the observed direction in which participants' preferences shifted (on the group level) in the experiment, and whether this was significant according to a logistic general linear mixed effect model explained in the Results. PTHD = prospect theory and hyperbolic discounting; MHD = multiplicative hyperboloid discounting; HD = hyperbolic discounting; PTT = probability-time-trade-off; RITCH = risky intertemporal choice heuristic; PD = proportional difference.
† R = risk; D = delay; A = amount. * PTT and HD predict participant's preferences will shift towards the option with the highest combined risk and delay. They predict this for both the immediacy and certainty manipulations and, therefore, predict the same direction of shift for both manipulations.

Table 4
Predicted Shift in Preferences Caused by the Certainty Manipulation for Each Choice Type

Model	RvA†	DvA†	DvR†	RvAD†	DvAR†	DRvA†
PTHD	Larger-**Riskier**	No Change	**Riskier**-Sooner	Larger-**Riskier**-Sooner	Smaller-**Riskier**-Sooner	Larger-**Riskier**-Later
MHD	Larger-**Riskier**	No Change	**Riskier**-Sooner	Larger-**Riskier**-Sooner	Smaller-**Riskier**-Sooner	Larger-**Riskier**-Later
HD	Larger-**Riskier**	Larger-**Later**	Same as Immediacy*	Same as Immediacy*	Same as Immediacy*	Larger-**Riskier-Later**
PTT	Larger-**Riskier**	Larger-**Later**	Same as Immediacy*	Same as Immediacy*	Same as Immediacy*	Larger-**Riskier-Later**
Trade-off	Larger-**Riskier**	Smaller-**Sooner**	**Riskier-Sooner**	Larger-**Riskier-Sooner**	Smaller-**Riskier-Sooner**	Unclear
RITCH	Larger-**Riskier**	No Change	**Riskier**-Sooner	Larger-**Riskier**-Sooner	Smaller-**Riskier**-Sooner	Larger-**Riskier**-Later
PD	No Change	No Change	No Change	No Change	No Change	No Change
Data	Larger-Riskier ($p < .001$, $p_{adj} < .001$)	Smaller-Sooner ($p = .004$, $p_{adj} = .004$)	Riskier-Sooner ($p < .001$, $p_{adj} < .001$)	Larger-Riskier-Sooner ($p < .001$, $p_{adj} < .001$)	Smaller-Riskier-Sooner ($p < .001$, $p_{adj} < .001$)	Larger-Riskier-Later ($p < .001$, $p_{adj} < .001$)

Note. The predictions of each model are on a separate row. Words in bold indicate the attribute driving the effect. The data row shows the observed direction in which participants' preferences shifted (on the group level) in the experiment, and whether this was significant according to a logistic general linear mixed effect model explained in the Results. PTHD = prospect theory and hyperbolic discounting; MHD = multiplicative hyperboloid discounting; HD = hyperbolic discounting; PTT = probability-time-trade-off; RITCH = risky intertemporal choice heuristic; PD = proportional difference.
† R = risk; D = delay; A = amount. * PTT and HD predict participant's preferences will shift towards the option with the highest combined risk and delay. They predict this for both the immediacy and certainty manipulations and, therefore, predict the same direction of shift for both manipulations.

Turning to intertemporal choices, we note that the hyperbolic discounting model is similar to prospect theory, except that the utility of an option, $g(x, t)$, where t is the delay before receiving the outcome, x, is the product of a value function and a discount function, $U(x, t) = v(x) \cdot d(t)$. A hyperbolic discounting function is used to transform objective time, t, into a subjective evaluation of the wait, $d(t)$,

$$d(t) = \frac{1}{1 + ht} \tag{4}$$

where $h \geq 0$. The discount rate, h, measures the extent to which outcomes lose their value per unit of time they are delayed. High values of h indicate that outcomes are discounted steeply as a function of time, losing value quickly when delayed, while $h = 0$ indicates no discounting, with delays essentially ignored. When applied to intertemporal choices, the value function of the hyperbolic discounting model is often set to the identity function, for instance by setting $a = 1$ in Equation 1 (Kirby, 1997; Kirby & Marakovic, 1995).

A simple way to create a utility model that can deal with risky intertemporal choices is therefore to assume that the utility is the product of all three components: the value function (Equation 1), the decision weight function (Equation 2), and the discount rate function (Equation 4), $U(x, p, t) = v(x) \cdot w(p) \cdot d(t)$ (see Vanderveldt et al., 2015, for a discussion of whether the combination of these functions should be multiplicative or additive).[2] A consequence of this combination is that it forces the assumption of a single value function for both risks and delays. The assumption of a single value function is supported in Luckman, Donkin, and Newell (2015, 2018), where risky choices and intertemporal choices were found to have the same, usually concave, value function, although some participants did exhibit linear ($a = 1$), or slightly convex ($a > 1$) functions. As such we use Equation 1 but assume only that $a \geq 0$, rather than also constraining it to be less than 1, therefore removing the constraint that the value function must be concave. We also need to specify a choice function; in this case we use Equation 3.[3] We refer to this model as the prospect theory and hyperbolic discounting model (PTHD) throughout.

Multiplicative hyperboloid discounting model (MHD). The MHD model proposed by Vanderveldt et al. (2015) is an existing model of risky intertemporal choice. Just like the PTHD model, MHD assumes that the utility of an option is the product of a discount rate, value function and decision weight, however it differs in the specification of these functions. In particular, MHD uses a two-parameter S-shaped weighting function,

$$w(p) = \frac{1}{(1 + h_r(1/p - 1))^{s_r}} \tag{5}$$

where $s_r \geq 0$ and $h_r \geq 0$, and a hyperboloid discounting function,

$$d(t) = \frac{1}{(1 + h_d t)^{s_d}} \tag{6}$$

where $s_d \geq 0$ and $h_d \geq 0$. The parameters s_d and s_r measure sensitivity to delays and risks, respectively, with values less than 1 representing diminishing sensitivity. h_r and h_d are discount rate parameters, similar to Equation 4, except that h_r measures discounting as a function of odds against receiving the outcome, rather than as a function of delay. This model can account for basic patterns of risky intertemporal choice data, but its performance has not been compared with other models of risky intertemporal choice (Vanderveldt et al., 2015). The MHD model outperforms the hyperbolic discounting models in its ability to fit to pure intertemporal choice data and pure risky choice data (Green & Myerson, 2004; Myerson et al., 2003).

This base version of MHD does not incorporate any magnitude effects. However, variants of this model have been used to investigate the magnitude effect in intertemporal choice, and the peanuts effect in risky choice. In the former, it is generally found that the discount rate parameter, h_d, changes as function of the magnitude of the reward, while in the later it is the sensitivity parameter, s_r, which changes (Myerson et al., 2003; Myerson, Green, & Morris, 2011). While no modification of the discount rate has been proposed to account for this change, Myerson, Green, and Morris (2011) proposed to make the probability weighting function depend on amount, allowing the model to account for the peanuts effect;

$$w(p, x) = \frac{1}{(1 + h_r(1/p - 1))^{s_r x^c}} \tag{7}$$

where $c \geq 0$, and measures diminishing sensitivity to amounts. For the version of MHD we use in this article, we therefore assume that the utility of an option is a multiplicative combination of Equation 1, 6, and 7.[4] Furthermore, though no choice function has been proposed for the MHD model, we will use Equation 3 to make the PTHD and MHD models more comparable.

Translation models. We now consider the two other existing utility-based models of risky intertemporal choice. These models make substantially different assumptions about how risk and delay combine in choice. In particular, rather than multiplying together a decision weight and discount rate function, these models assume that there is a function translating one of the attributes, risk or delay, into levels of the other attribute.

In the HD model of Yi et al. (2006) it is assumed that risk is transformed into a delay, which is then added to the existing level of delay before a discount function is applied, such that

$$d(t, p) = \frac{1}{1 + h(t + i(1/p - 1))} \tag{8}$$

with $i \geq 0$ and $h \geq 0$. Because risks are incorporated into the discount function, there is no probability weighting function, and we have $U(x, p, t) = v(x) \cdot d(t, p)$, with $v(x)$ calculated from Equation 1. The transformation used in this model is based on Rachlin, Raineri, and Cross (1991), who proposed that a hyperbolic

[2] Empirically Vanderveldt et al. (2015), find that a multiplicative form fits their risky intertemporal data better than an additive form such as $U(x, p, t) = v(x) - v(x) \cdot [1 - w(p)] - v(x) \cdot [1 - d(t)]$. Furthermore the additive form is theoretically problematic as, for a gamble with a positive outcome, $v(x) > 0$, it will produce a negative utility whenever $w(p) + d(t) < 1$. This is easily achieved for combinations of moderate delays and low probabilities. This issue does not occur with the multiplicative form.

[3] The combination of probability, value and choice functions given by Equations 1–3 is the combination advocated by Stott (2006) for risky choice.

[4] In the versions of MHD typically used in previous papers the exponent of the value function, a, cannot be separated from the sensitivity exponents, s_r and s_d, as the parameters of the model are normally estimated based on obtaining certainty equivalents of risky and/or intertemporal options (see Myerson et al., 2011). Because we assume participants compare the utility of the options presented, rather than their certainty equivalents, we estimate a separately from, s_r and s_d.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

discounting function could explain both intertemporal choice *and* risky choice behavior, if probability is translated into a delay, $t_p = i(1/p - 1)$ where i is the intertrial interval, that is, the amount of time between each play of the gamble, which is treated as a free parameter in our model.[5] This transformation assumes that risky choices are made by calculating the amount of time it would take, on average, to achieve the desired amount, if the gamble was repeated multiple times (Vanderveldt et al., 2017). Larger values of i, therefore, lead to a more negative view of risk as the inferred delay is longer. Vanderveldt et al. (2015) also motivate their probability weighting function in the MHD model on the same idea of transforming risks into delays, although they replace the assumption that risk and delay are interchangeable and combined, with the notion that the outcome value should be discounted separately based on the transformed risk and the delay. The discount rate parameter, h, from Equation 8 has the same interpretation as the discount rate in the pure hyperbolic model in Equation 4, except that it takes into account both the objective delay, and delay inferred from the risk.

The second translation model we fit is the PTT model, proposed by Baucells and Heukamp (2010, 2012). In the PTT, delays are combined with probabilities in the weighting function

$$w(p, t, x) = e^{-(Rt/|x| - \ln p)^{\gamma}} \tag{9}$$

where $R \geq 0$ and $0 \leq \gamma \leq 1$, which means there is no separate discounting function and therefore $U(x, p, t) = v(x) \cdot w(p, t, x)$. Here $R/|x|$ is the discount rate for time, similar to h in the hyperbolic models, with larger values of R indicating more discounting. However, unlike h, the discount rate in Equation 9 also decreases as a function of the amount. As discussed previously, by making the discount rate amount dependent the PTT model has the ability to predict intertemporal magnitude effects (Baucells & Heukamp, 2010). The second parameter, γ, measures subproportionality and has the same interpretation as r from Equation 2. Another critical difference between PTT and the other utility models is the use of a two-parameter concave value function

$$v(x) = \frac{1 - e^{-\alpha x^{1-\beta}}}{\alpha} \tag{10}$$

with $\beta \leq 1$ and $\alpha > 0$, instead of the power function employed in other models. An important difference between the power function (Equation 1) and the expo-power function (Equation 10) is that the former has constant elasticity, while the later has decreasing elasticity (Baucells & Heukamp, 2010; Scholten & Read, 2014). Decreasing elasticity means that proportional increases in outcome magnitude, x, lead to smaller proportional increases in subjective value, $v(x)$, for larger values of x than for smaller values of x (Scholten & Read, 2014). In contrast, constant elasticity means that the proportional change in subjective value, $v(x)$, is the same regardless of the value of x. In practice, the use of Equation 10 allows the PTT model to capture some "peanuts" type effects. Equation 10 can also produce constant elasticity when α approaches 0. The concavity, or extent of diminishing sensitivity to amount, in Equation 10 is controlled by both α and β with the concavity of the function decreasing as β decreases or α increases.

The translations/combinations of risk and delay proposed by both Equation 8 and Equation 9 are motivated by the observed certainty effects in intertemporal choice/Delay vs. Amount choices (Keren & Roelofsma, 1995) and immediacy effects in risky choice/Risk vs. Amount choices (Baucells & Heukamp, 2010). Both equations assume that adding a common delay has the same effect as dividing by a common probability, and vice versa. Because neither function has previously been paired with a choice function, we use Equation 3 to calculate choice probabilities to make it comparable to the other models.

Attribute-comparison models. In attribute based models of choice, the two options in any given gamble are not evaluated independently. Instead, it is assumed that participants directly compare the difference in attributes (i.e., the level of risk, delay, and amount) between the two options. These differences are then considered evidence in favor of one option or the other, depending upon the direction of the difference. Several attribute-comparison models have been proposed, such as the stochastic proportional difference model (González-Vallejo, 2002) used in risky choice, or the DRIFT (Read, Frederick, & Scholten, 2013), trade-off (Scholten & Read, 2010), and ITCH (Ericson et al., 2015) models in intertemporal choice.

To our knowledge the only existing, albeit empirically untested, model of risky intertemporal choice that uses an attribute comparison process, is the modified version of the trade-off model suggested by Scholten and Read (2014). This model can be considered a hybrid of an attribute-comparison and utility-comparison model, as it reduces to their intertemporal trade-off model (Scholten & Read, 2010)—an attribute comparison model—when dealing with pure intertemporal choices, and to a version of prospect theory—a utility comparison model—when dealing with pure risky choices (Scholten & Read, 2014).

Like the PTHD and MHD models, the trade-off model assumes that objective probabilities are transformed into decision weights by an s-shaped weighting function,

$$w_p(p) = \frac{p^{\gamma}}{(p^{\gamma} + (1 - p)^{\gamma})^{1/\gamma}} \tag{11}$$

with $0 < \gamma \leq 1$, and that objective outcomes are transformed into subjective values by a concave value function with decreasing elasticity,

$$v(x) = \frac{1}{\beta} \cdot \ln(1 + \beta x) \tag{12}$$

with $\beta > 0$. Further, these subjective values and decision weights are multiplied together as they are in a utility model, $U(x, p) = v(x) \cdot w_p(p)$. The interpretation of γ is similar to other subproportionality parameters, such as r from Equation 2, while increasing β increases diminishing sensitivity to amount.

The trade-off model differs from utility models in how delays are used in the decision. First, the trade-off model assumes that objective delays are translated into time-weights, which show diminishing sensitivity to delay,

[5] We made two alterations to the HD model as reported in Yi et al. (2006). Firstly, Yi et al. (2006) assumed a value of 35.3 for i, as this was the value found by Rachlin et al. (1991). As there is no reason to assume i should be stable across experiments or participants, we instead assume i is a free parameter. Secondly the original formulation, consistent with the standard hyperbolic discounting model, assumed an identity function for the value function. We relax this assumption, instead using a power function (Equation 1) as previous research has found little support for linear or identity value functions (Dai & Busemeyer, 2014; Luckman et al., 2015, 2018).

$$w_t(t) = \frac{1}{\tau} \cdot \ln(1 + \tau t) \tag{13}$$

with $\tau > 0$, and diminishing sensitivity to delay increasing as τ increases. Rather than multiplying the utility by these time-weights, instead a weighted difference of the time-weights is calculated,

$$Q(t_S, t_L) = \frac{\kappa}{\alpha} \ln \left(1 + \alpha\left(\frac{w_t(t_L) - w_t(t_S)}{\vartheta}\right)^{\vartheta}\right) \tag{14}$$

where t_S and t_L are the delays to the sooner and later options, respectively, and $\kappa \geq 0$, $\alpha > 0$, and $\vartheta \geq 1$. κ is a free parameter scaling differences in time to differences in utility, which could be interpreted as relative sensitivity to time versus utility differences, while α and ϑ measure subadditivity and superadditivity respectively (see Appendix B for further details on the interpretation of the parameters). In order to calculate the probability of choosing the gamble with the later outcome, $P(g_L \mid g_L, g_S)$, this weighted time difference, $Q(t_S, t_L)$ (Equation 14), is then compared with the difference in utility between the later option and sooner option,

$$U(g_L) - U(g_S) = v(x_L) \cdot w_p(p_L) - v(x_S) \cdot w_p(p_S) \tag{15}$$

where the S and L subscripts denote the attributes of the gamble with the sooner outcome, g_S, and later outcome, g_L, respectively. To be consistent with the other models under evaluation, we assume that $P(g_L \mid g_L, g_S)$ is generated by taking the difference between the utility difference (Equation 15) and the weighted time difference (Equation 14), and passing it through the logistic choice function from Equation 3, giving

$$P(g_L \mid g_L, g_S) = \frac{1}{1 + e^{-s(U(g_L) - U(g_S) - Q(t_S, t_L))}} \tag{16}$$

One advantage of using this specification for the choice function is that it can accommodate choices where both the utility difference and weighted time difference favor the same option.[6] The final two models we consider rely purely on attribute-comparisons.

Risky intertemporal choice heuristic model. We chose the ITCH model as the basis for our attribute-based risky intertemporal choice model because it is one of the few models that considers both absolute and relative differences in attribute levels, with most existing attribute comparison models considering only one or the other. For instance, the other pure attribute comparison model we evaluate, the PD model (González-Vallejo, 2002), assumes that only the relative, or proportional differences in attribute levels, are considered, while the trade-off model uses transformed absolute differences in attribute levels (Scholten & Read, 2010). The advantage of assuming that participants are sensitive to both absolute and relative differences in attributes across options is that this naturally produces a similar effect to assuming diminishing sensitivity to the attribute. For instance, the absolute difference between \$10 and \$11 is the same as the absolute difference between \$100 and \$101; however, the relative difference in amounts is much greater in the former case. A model which is sensitive to both types of differences will, overall, treat the amount differences as greater for the two smaller amounts due to the greater relative difference. This is similar to a model which assumes diminishing sensitivity to amounts, that is, a concave value function.

In the original ITCH model, it is assumed that participants consider both the absolute and relative differences in the two attributes, delay and amount, when making intertemporal choices. Specifically, they take a weighted sum of the absolute and relative difference in amount between the two options,

$$X = \beta_{xA}(x_1 - x_2) + \beta_{xR}\frac{x_1 - x_2}{x_m} \tag{17}$$

where $x_m = \frac{x_1 + x_2}{2}$, and the weighted sum of the absolute and relative differences in time between the two options,

$$T = \beta_{tA}(t_2 - t_1) + \beta_{tR}\frac{t_2 - t_1}{t_m} \tag{18}$$

where $t_m = \frac{t_1 + t_2}{2}$. Both x_m and t_m are reference points for calculating the relative difference components. The various β parameters are the weights given to each of the differences in attributes, and all $\beta \geq 0$. Psychologically, the relative values of the β parameters allow the model to place different weights, or importance, on different attributes, that is, delays or amounts, and different comparison types, that is, relative or absolute. A choice function is then used to calculate the probability of choosing Option 1 over Option 2, $P(g_1 \mid g_1, g_2) = \frac{1}{1 + e^{-(X + T + \beta_O)}}$, where β_O represents an overall bias, such that if $\beta_O > 0$ the participant has a bias toward the larger option, if $\beta_O < 0$ the participant has a bias for the sooner option.

Because the ITCH model was developed for intertemporal choice data, we had to make two modifications to allow it to deal with risks in choice. First, analogous to the way that the ITCH model deals with delay, we propose that participants also incorporate a weighted sum of the absolute and relative differences in probabilities between the two options

$$R = \beta_{pA}(p_1 - p_2) + \beta_{pR}\frac{p_1 - p_2}{p_m} \tag{19}$$

where $p_m = \frac{p_1 + p_2}{2}$, $\beta_{pA} \geq 0$ and $\beta_{pR} \geq 0$. Second, we proposed a change to the bias parameter. Rather than a single bias parameter, RITCH has three bias parameters, β_{xO}, a bias toward the larger option, β_{tO}, a bias toward the sooner option and, β_{pO}, a bias toward the safer option. We add these bias parameters to Equations 17–19, resulting in:

$$X = \beta_{xO} \cdot \mathrm{sgn}(x_1 - x_2) + \beta_{xA}(x_1 - x_2) + \beta_{xR}\frac{x_1 - x_2}{x_m} \tag{20}$$

$$T = \beta_{tO} \cdot \mathrm{sgn}(t_2 - t_1) + \beta_{tA}(t_2 - t_1) + \beta_{tR}\frac{t_2 - t_1}{t_m} \tag{21}$$

$$R = \beta_{pO} \cdot \mathrm{sgn}(p_1 - p_2) + \beta_{pA}(p_1 - p_2) + \beta_{pR}\frac{p_1 - p_2}{p_m} \tag{22}$$

All three bias parameters are constrained to be non-negative, but are multiplied by the sign of the difference in attribute value they are linked to, so that β_{xO} always increases preference for the option with the larger amount, β_{tO} preference for the option with the shorter delay and β_{pO} for the option with the least risk. If there is

[6] Scholten et al. (2014), use a different, proportional, choice function when using the intertemporal trade-off model. This choice function is integral to some of the predictions their model makes in Scholten et al. (2014). To use this function we would need to insure that both the utility difference and time-weight difference were always positive (i.e. they favor different options), which is not possible in our choice set. As such, our risky intertemporal trade-off model does not reduce to the intertemporal trade-off model when dealing with pure intertemporal choices.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

no difference in the value of an attribute across options, for example, $x_1 = x_2$, then the corresponding bias parameter is multiplied by 0, that is, $\text{sgn}(0) = 0$. Therefore, the three bias parameters are combined in a unique way for each of our six types of risky intertemporal choice.

The need for a unique bias for each of the six types of risky intertemporal choices, as they involve different trade-offs between attributes, is the reason that we use three bias parameters instead of one. In the original ITCH model a single bias parameter is sufficient, and identifiable, as only one trade-off is ever encountered by the model, that between delay and amount. Therefore β_O, this single bias, is the bias for preferring the larger later over the smaller sooner option. In our model we can also identify the bias for preferring the larger later option over the smaller sooner by taking the difference between β_{xO}, the bias for the larger option, and β_{tO}, the bias for the sooner option (this difference, like β_O, can be positive or negative depending which bias is greater). A similar single bias parameter based on amount and delay would be insufficient for risky intertemporal choices as it could not capture bias due to risk, such as the bias to prefer a larger later safer option over a smaller sooner riskier option. Further, a single bias parameter would be undefined for choices where there was no trade-off between delay and amount, either because one of these attributes does not vary between the options (DvR and RvA choices) or because delay and amount favor the same option (RvAD choices). An alternative to the three bias parameter model we propose, would be a six bias parameter model, with a unique bias for each of the six types of risky intertemporal choices. For simplicity, we use a three parameter variant.

Similar to the ITCH model the attribute differences X (Equation 20), T (Equation 21), and R (Equation 22) are summed and passed through the logistic choice function to calculate preferences.[7]

$$P(g_1 \mid g_1,\ g_2) = \frac{1}{1 + e^{-(X+T+R)}} \tag{23}$$

Proportional difference model (PD). The final model we consider is a version of the PD model. This model was chosen because, while it has not been used for risky intertemporal choices or for comparisons that involve more than two attributes, versions of it have been used for both risky choices (González-Vallejo, 2002) and intertemporal choices (Cheng & González-Vallejo, 2016; Dai & Busemeyer, 2014) separately. Furthermore, the risky intertemporal version of the PD model we propose can be considered a simplification or modification of the RITCH model above. Unlike the RITCH model the PD model assumes that people only consider the relative difference in attribute values—risks, delays, and amounts—across the two options, rather than both relative and absolute differences. It therefore presents a psychologically simpler comparison process, where each pair of attribute values is only considered once. To obtain the PD model from the RITCH model we modify the reference points– x_m, p_m and t_m – so that the maximum, rather than mean value is used,[8] and simplify Equations 20–22 so that only relative differences in attribute values are considered in the difference calculations, removing the absolute components

$$X = \beta_{xO} \cdot \text{sgn}(x_1 - x_2) + \beta_{xR}\frac{x_1 - x_2}{\max(x_1,\ x_2)} \tag{24}$$

$$T = \beta_{tO} \cdot \text{sgn}(t_2 - t_1) + \beta_{tR}\frac{t_2 - t_1}{\max(t_1,\ t_2)} \tag{25}$$

$$R = \beta_{pO} \cdot \text{sgn}(p_1 - p_2) + \beta_{pR} \times \frac{p_1 - p_2}{\max(p_1,\ p_2)} \tag{26}$$

The weights given to each relative difference—β_{xR}, β_{tR}, and β_{pR} —and the bias parameters—β_{xO}, β_{tO} and β_{pO}—are all constrained to be non-negative, as they are in Equations 20–22. Equations 24 to 26 are then summed and passed to the choice function (Equation 23), as they are in the RITCH model.

This version of the PD model differs from the more commonly used variants in three ways. First, it involves three attribute comparisons, X, T and R, rather than just two, such as X and T (González-Vallejo, 2002) or X and R (Cheng & González-Vallejo, 2016). Second, and as a consequence, it employs a different bias method, as the threshold or bias mechanism usually employed in the PD model is similar to that of the original ITCH model, and therefore is not appropriate for use in a situation that involves potential biases for three different attributes. The third major difference is that we allow the different attributes to be given different relative weights (i.e., β_{xR}, β_{pR}, and β_{tR}), which is the approach employed by Dai and Busemeyer (2014), and makes the PD model more comparable to the RITCH model. The psychological interpretation of the parameters of the PD model are similar to those of the RITCH model (Appendix B).

Experiment

In the above sections, we introduced six different types of risky intertemporal choices (see Table 1) and considered three manipulations: magnitude, certainty, and immediacy, which can be applied to them. Furthermore, we introduced seven potential models of risky intertemporal choice. In the following experiment, we test each model's ability to predict the effects of each of the three manipulations in all six choice types. This is done both quantitatively, by comparing the models fits, and qualitatively, by considering the broad predictions each model makes about the effects of each manipulation.

Qualitative Predictions

In Tables 2 to 4 we consider each of the three manipulations, magnitude, certainty, and immediacy, and outline how each model

[7] An alternative, but equivalent, specification of the RITCH model would be to constrain the six weighting parameters to sum to 1. This would reduce the number of directly estimated weighting parameters to five, as the sixth would be determined by the other five. However, this specification would then require an additional sensitivity parameter, to capture the fact that participants may have the same relative weighting of attributes, but differ in how deterministically they behave based on this weighting. Mathematically these models are equivalent. A similar alternate specification could be employed for the PD model below.

[8] We also tested a variant which used the mean value as the reference point. The maximum reference version outperformed the mean reference version for all three effects. While the maximum is often used, this is not a key feature of the general PD model. The minimum value has also been used as a reference point (Dai & Busemeyer, 2014) and the reference point can also be treated as a free parameter which might be based on the broader choice context or other considerations (Cheng & González-Vallejo, 2016; González-Vallejo, 2002).

predicts participants' preferences will change when the manipulation is applied, relative to a baseline set of choices. This is done separately for each of the six choice types. Table 2 outlines how participants' preferences are expected to change if all outcomes are multiplied by 10, that is, the magnitude manipulation. Table 3 outlines predicted preference changes when a common 12-month delay is added to all choices, the immediacy manipulation, and Table 4 the changes when all probabilities in the choice set are divided by five, the certainty manipulation.

The predictions outlined in Tables 2 to 4 for each model are described in more detail in the online supplemental materials. To provide an example we will briefly outline the predictions for the PTHD and RITCH models for each manipulation and choice type.

Magnitude. Neither the version of prospect theory nor the hyperbolic discounting model that PTHD is based on can account for either peanuts or magnitude effects, because neither the probability-weighting function (Equation 2), nor discounting function (Equation 4), are affected by outcome amount, and the power value function (Equation 1), has constant elasticity which preserves the order of utilities. As such, PTHD predicts no directional effects of the magnitude manipulation (see top row of Table 2).

The quantitative predictions of PTHD is more nuanced than reflected in Table 2, however. The PTHD model predicts that the magnitude manipulations will cause preferences to become stronger. Because the probability generated in Equation 3 is dependent upon the absolute difference between the utilities, any manipulation that changes the size of the utilities involved will change the probability of choosing one option over the other. Specifically, multiplying both outcomes by 10, as we do in the magnitude manipulation, is equivalent to multiplying the difference in utilities, U_D, from the baseline choice by 10^a. Because the probability of making a given choice increases or decreases with U_D (Equation 3), any preference shown by a participant will be exaggerated. For example, if $U_D > 0$, then $10^a \times U_D > U_D$. In essence, we would expect the strength of the participants' preference to increase as the magnitude of the outcomes increases, such that choice probabilities above 0.5 will shift toward 1 and choice probabilities below 0.5 will shift toward 0.

In Tables 2 to 4, for simplicity, we have chosen to ignore cases in which there is no consistent change in the direction of preferences across either participants or items. This approach is equivalent to predicting the option participants would choose in each condition if they were indifferent between the options in the baseline condition. Note that our quantitative comparison does take the strengthening and weakening predictions of the models into account.

Unlike PTHD, the RITCH model does predict that the magnitude manipulation will change the direction of participants' preferences. RITCH predicts a similar magnitude effect in risky intertemporal choices to the ITCH model, which was designed to capture magnitude effects in intertemporal choice (Ericson et al., 2015). Specifically, multiplying both amounts by a common factor causes the absolute difference in amount between the two options to increase, while leaving all the other differences in attributes unchanged. Because $\beta_{xA} \geq 0$, there is an overall shift in preference toward the option with the larger outcome, as X will increase when $x_1 > x_2$ and decrease when $x_1 < x_2$ (Equation 20). Therefore, for all choice types except Delay vs. Risk, where the difference in amounts is 0 regardless of the manipulation, the RITCH model predicts that participants will be more likely to choose the larger amount option in the magnitude condition than in the baseline condition (see row 6 of Table 2).

To demonstrate this formally, let $g_{b1}(x_1, p_1, t_1)$ and $g_{b1}(x_2, p_2, t_2)$ be the options presented to a participant in the baseline choice, where $x_1 < x_2$. From Equations 20 to 23 we can calculate their probability of choosing g_{b1} as follows:

$$X_b = \beta_{xO} \cdot \mathrm{sgn}(x_1 - x_2) + \beta_{xA}(x_1 - x_2) + \beta_{xR}\frac{x_1 - x_2}{\mathrm{mean}(x_1,\ x_2)}$$

$$T_b = \beta_{tO} \cdot \mathrm{sgn}(t_2 - t_1) + \beta_{tA}(t_2 - t_1) + \beta_{tR}\frac{t_2 - t_1}{\mathrm{mean}(t_1,\ t_2)}$$

$$R_b = \beta_{pO} \cdot \mathrm{sgn}(p_1 - p_2) + \beta_{pA}(p_1 - p_2) + \beta_{pR}\frac{p_1 - p_2}{\mathrm{mean}(p_1,\ p_2)}$$

$$P(g_{b1} \mid g_{b1}, g_{b2}) = \frac{1}{1 + e^{-(X_b + T_b + R_b)}}$$

Similarly, let g_{m1} and g_{m2} be the same choice presented in the magnitude condition, with both amounts multiplied by 10. Because time and risk are equivalent in the baseline and magnitude conditions, $T_b = T_m$ and $R_b = R_m$. Substituting into Equation 20, however, we see that

$$X_m = \beta_{xO} \cdot \mathrm{sgn}(10x_1 - 10x_2) + \beta_{xA}(10x_1 - 10x_2) + \beta_{xR}\frac{10x_1 - 10x_2}{\mathrm{mean}(10x_1,\ 10x_2)}$$

$$X_m = \beta_{xO} \cdot \mathrm{sgn}(x_1 - x_2) + 10 \cdot \beta_{xA}(x_1 - x_2) + \beta_{xR}\frac{10(x_1 - x_2)}{10 \cdot \mathrm{mean}(x_1,\ x_2)}$$

$$X_m = \beta_{xO} \cdot \mathrm{sgn}(x_1 - x_2) + \beta_{xA}(x_1 - x_2) + \beta_{xR}\frac{(x_1 - x_2)}{\mathrm{mean}(x_1,\ x_2)} + 9 \cdot \beta_{xA}(x_1 - x_2)$$

$$X_m = X_b + 9 \cdot \beta_{xA}(x_1 - x_2)$$

Now, we have $P(g_{m1} \mid g_{m1}, g_{m2}) = \frac{1}{1 + e^{-(X_b + T_b + R_b + 9 \cdot \beta_{xA}(x_1 - x_2))}}$ and because $x_1 < x_2$, $9 \cdot \beta_{xA}(x_1 - x_2)$ is negative. We can then conclude that $P(g_{m1} \mid g_{m1}, g_{m2}) < P(g_{b1} \mid g_{b1}, g_{b2})$. Therefore, the RITCH model predicts that multiplying both outcomes by 10 leads to a higher probability of participants choosing the option with the higher amount, Option 2.

It is important to note that the RITCH model cannot produce the previously observed effect of magnitude in the Delay vs. Risk choices, where participants became more likely to choose the delayed option when the amount of both options is increased (Baucells & Heukamp, 2010; Christensen et al., 1998; Luckman et al., 2017). Nor can the model produce a peanuts effect in Risk vs. Amount choices (Weber & Chapman, 2005b), but actually predicts the reverse of the peanuts effect (Chapman & Weber, 2006).

Immediacy. Because PTHD builds on hyperbolic discounting models, like those models, it predicts that an immediacy manipulation (i.e., increasing delays by a common duration) will increase preference for the later option in all choice types, regardless of the risks or amounts involved. This is because hyperbolic discounting functions (Equation 4) show diminishing sensitivity to delay. That is, a much higher discount rate per unit of time is applied to short time delays, than is applied to long time delays. Therefore, when both options are made more delayed, the relative difference in delay becomes less important. If both options have the same delay,

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

as they do in Risk vs. Amount choices, the immediacy manipulation will not shift preferences toward one option or the other, as there is no later option to make relatively more favorable. These predictions are shown in the top row of Table 3.

The RITCH model also produces immediacy effects, but rather than producing them through a discount function, it produces them in the same way as the ITCH model does for intertemporal choice (Ericson et al., 2015). The RITCH model assumes participants consider both absolute differences in delay, which are not affected by adding a common delay to both options, and relative differences, which are. Adding a common delay causes the relative difference in delay, $\beta_{tR}\frac{t_2 - t_1}{\text{mean}(t_1, t_2)}$, to decrease in magnitude as the numerator stays constant while the denominator, the mean delay, increases by 12 months. Reducing the relative difference in delay while leaving all other differences unaffected, causes the support for the sooner option to decrease, causing preferences to shift toward the later option. Therefore, the RITCH model predicts the same change in preferences for the immediacy manipulations as the PTHD model for all choice types, albeit for a different reason (see row 6 of Table 3).

Certainty. Due to PTHD's basis in prospect theory, the certainty manipulation (i.e., reducing the probabilities by a common ratio) will increase preferences for the riskier option in all choice types. This increase occurs because the probability weighting function (Equation 2) is s-shaped, overweighting small probabilities, making the low probability of the riskier option relatively more favorable in the certainty manipulation, and underweights moderate to large probabilities, making the riskier option relatively less favorable when the probabilities are larger. Of course, this prediction does not hold for choices where there is no riskier option—that is, the Delay vs. Amount choices (see top row of Table 4).

The RITCH model also predicts the certainty manipulation will increase preferences for the riskier option in all choice types. Dividing all probabilities by five has no effect on the relative difference in probabilities between the options, $\beta_{pR} \times \frac{p_1 - p_2}{\text{mean}(p_1, p_2)}$, because both the numerator and denominator are divided by five, but does decrease the magnitude of absolute difference in probabilities, $\beta_{pA}(p_1 - p_2)$. Reducing the absolute difference in probabilities while leaving all other differences unchanged makes the higher probability option less attractive, leading to a shift toward preferring the riskier option. Because the absolute difference in probabilities is 0 in Delay vs. Amount choices, RITCH predicts no change in preferences for this choice type (see row 6 of Table 4).

Method

Participants. One-hundred first-year psychology students from the University of New South Wales participated for course credit. All choices were hypothetical. The sample size, based on sample sizes used in model comparison studies in risky choice (Rieskamp, 2008; Stott, 2006), was decided prior to data collection.

Materials. We created 16 different instances of each of the six types of risky intertemporal choice. This total of 96 choices form the baseline set (see Appendix D for the amounts, delays and probabilities used in all baseline choices). Where possible, these choices were taken directly from those used in Luckman, Donkin, and Newell (2018) or in a pilot study completed by the participants in that experiment. In order to create enough choices of each type, some of the choices from Luckman et al. (2018), or the pilot, were modified by adding a random small delay or risk. The delays in the baseline set ranged from 0 to 54 months, the probabilities from 0.05 to 1 and the amounts from \$50 to \$475.[9]

The 96 baseline choices were modified three times to create three additional choice sets, each matching one of the three key manipulations. The 96 immediacy choices were identical to the baseline set, but with an extra 12 months added to the delays of both options. Similarly, the certainty set had the probabilities divided by five, and the magnitude set had the outcome amounts multiplied by 10.

In addition to the 384 choices of interest, we also included six choices where one option was dominated. In these dominated choices, one option was better than the other on at least one attribute, and matched on the remaining attributes. These choices were included as check questions, such that we excluded from further analysis any participants who chose the dominated option on more than one of these choices.

Procedure. Participants were told that they would be making choices that involved a mixture of risks and delays. Each choice would be between two options, and there were no correct answers, rather we were interested in their preferences. They were also told to consider each choice independently, and that all choices were hypothetical. There was no time limit.

For each choice, participants were given both a written statement of the two options, and a summary of the relevant amounts, delays, and probabilities of each option (see Appendix C). In addition to instructing participants to take their time and have breaks if necessary, the task was also divided into three sections of 130 choices each to allow breaks. The order of the 390 choices was randomized for all participants. A warning to consider their decisions carefully also appeared if a participant chose the dominated option in any of the six check questions.

Results

Exclusions. Thirty-two participants chose at least one dominated option throughout the experiment. In line with our preset exclusion criteria, 10 of these participants were excluded as they chose two or more dominated options leaving 90 participants in our analysis.

Group data. Figure 1 shows the proportion of participants choosing each option in the four choice sets (*baseline, magnitude, immediacy, certainty*). Each dot represents a single choice, while the crosses show the mean proportions for each of the six choice types (RvA, DvA, DvR, RvAD, DvAR, DRvA). (The choice proportions are also reported in Appendix D). Each panel of Figure 1 shows the responses for one of the six choice types. The *y*-axis label in each

[9] To ensure that we had some choices very similar to standard risky and standard intertemporal choices, eight of the 16 RvA choices had a delay of 0 for both options, and eight of the 16 DvA choices had a probability of 1 for both options. Furthermore, because many studies of risky intertemporal choice have focused on circumstances where at least one option contains a certain or immediate outcome, for the other four choice types half of the baseline choices had a sooner option with a delay of 0 and a safer option with a probability of 1.

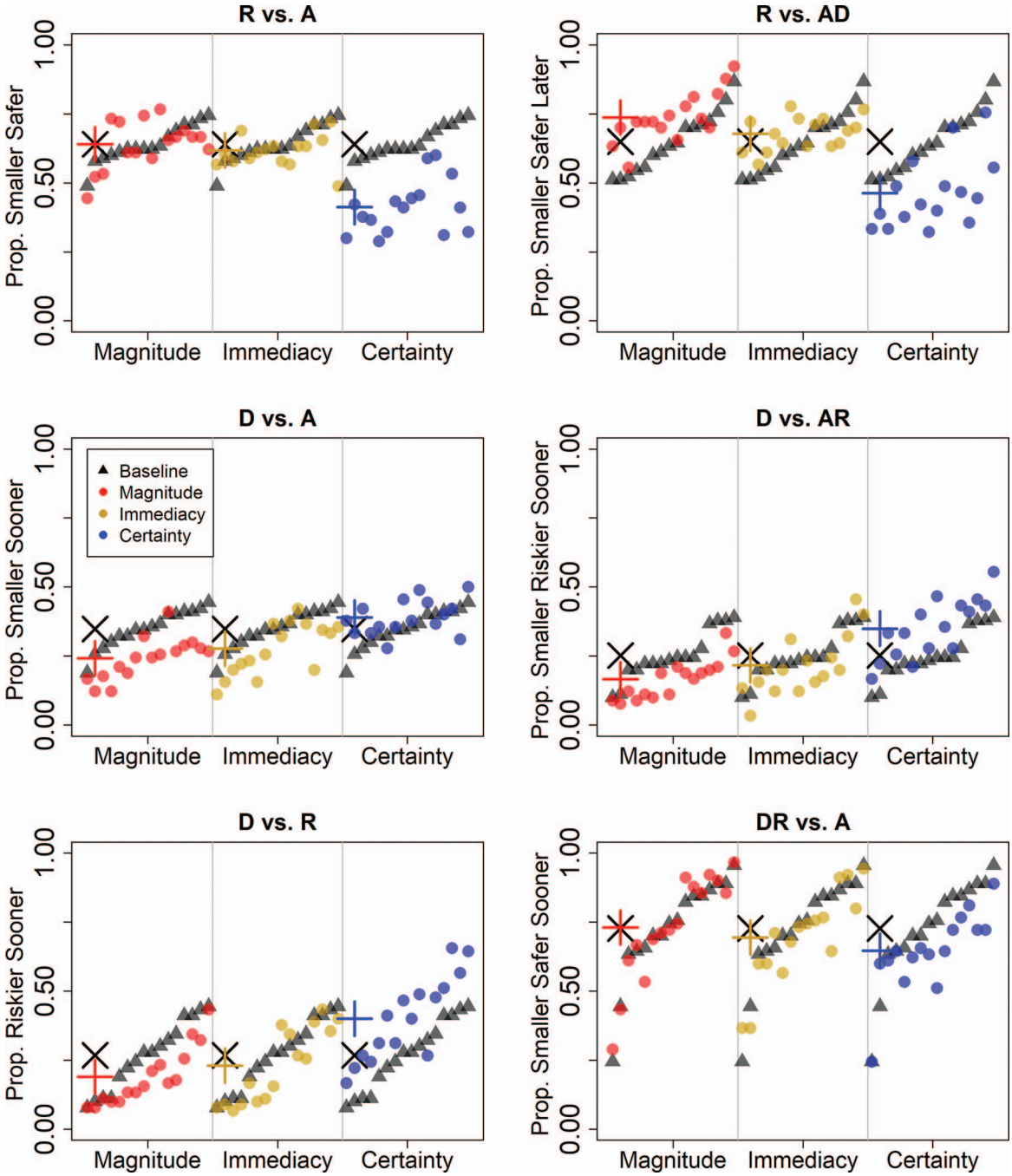


Figure 1. Proportion of participants choosing Option 1 for each choice. Choices are grouped into panels according to the choice type: Risk vs. Amount (RvA), Delay vs. Amount (DvA), Delay vs. Risk (DvR), Risk vs. Amount & Delay (RvAD), Delay vs. Amount & Risk (DvAR), and Delay & Risk vs. Amount (DRvA). Black triangles are the baseline condition and are the same across all three segments of each panel. Red dots show the magnitude condition (left segment), yellow the immediacy (middle segment), and blue the certainty (right segment) for each choice type. Crosses are the mean for each choice type for each condition. See the online article for the color version of this figure.

panel indicates whether larger proportions reflect a preference for the smaller or larger amount, safer or riskier probability, and sooner or later delay for that choice type. If an attribute has the same value for both options a descriptor is not given for that attribute. For instance, for the RvA choices in the top left panel we plot the proportion of participants choosing the smaller safer option, as both options have the same delay. Each panel is further broken into three segments, one for each of the three manipula-

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

tions: magnitude, immediacy, and certainty. To highlight the effects of each manipulation, each segment plots both the choice proportion for the effect of interest (the colored dots) and the choice proportion in the baseline condition (black triangles). Finally, as there is no a priori order of choices within each choice type, to improve clarity we have ordered choices along the *x*-axis, within each choice type, so that choice proportion in the baseline condition increases.

To provide a rather crude test of the predictions in Tables 2–4, for each manipulation and choice type we fit a logistic general linear mixed-effect model (GLMM) to the baseline and particular manipulation data. The dependent variable was whether participants chose the option indicated in the *y*-axis label in Figure 1 for each of the six choice types. The predictor was whether the trial came from the baseline or manipulation condition (the fixed effect), with random intercepts for participants and items/choices.[10] That is, we allowed different participants to have different choice proportions, and different specific choice pairs to have different choice proportions. *p* values were calculated using the likelihood ratio test method of the afex package in R (Singmann, Bolker, Westfall, Aust, & Ben-Shachar, 2019) and adjusted for multiple comparisons using the Holm-Bonferroni method (Holm, 1979).[11] Because we required a separate logistic GLMM for each choice type and manipulation, 18 separate tests were run. The last row in each table (Tables 2–4) shows the direction of the effect of the manipulation, and *p* value obtained from these tests.

Group-level magnitude effects. In the magnitude condition (left segment of each panel in Figure 1) there is a greater preference for selecting the later (delayed) option, relative to the baseline condition, for the DvA, DvAR, DvR (red points below black points in the middle and bottom left panels) and RvAD (red points above black points in the top right panel) choice types. There is no significant change in preferences, relative to the baseline condition, for the RvA or DRvA choices (red and black points overlap in top left and bottom right panels). Overall, with the exception of the DRvA choices, this pattern is very close to what we would expect if the magnitude effect was caused by participants becoming willing to wait longer when the magnitude of the outcomes increases. If the magnitude effect was instead driven by participants choosing the larger option, as predicted by pure attribute-comparison models, then we would have expected no change in preference for the DvR choices, a reversed change in the RvAD choices and a change toward the larger riskier option in the RvA choices.

Qualitatively, the pattern we observe is closest to the PTT and trade-off predictions (see Table 2), although no model perfectly predicts all the data patterns. In particular the PTT, MHD, trade-off, and RITCH models all predict preference changes in both the RvA choices and the DRvA choices, which we do not observe. Conversely the predictions of the PTHD, HD, and PD models, which do not predict magnitude effects, are correct for these two choice types, but incorrect for the other four choice types where preference changes do occur. It should be noted that the RvA (top left) choices are very similar, or sometimes identical to, pure risky choices, so the lack of difference between the magnitude (red points) and baseline (black points) could be considered a failure to replicate the usual peanuts effect.

Group-level immediacy and certainty effects. In the immediacy manipulation (middle segment), choices appear to shift toward the later option for all choices types, except RvA (top left panel) where this cannot occur. However, this shift is not significant for the RvAD (top right panel) or DRvA (bottom right panel) choices. This pattern is consistent with the immediacy effect in intertemporal choice, and is, for the most part, qualitatively consistent with the predictions of the PTHD, MHD, trade-off, PD, and RITCH models in Table 3.

The certainty manipulation (right segment) seems to shift preferences toward the riskier options, consistent with the certainty effect in risky choice. This result also mostly qualitatively matches the predictions of the PTHD, MHD, RITCH, and trade-off models, as shown in Table 4. However, we also see a moderate shift toward the sooner option in the DvA choice (blue points higher than black points, middle left panel), which is only predicted by the trade-off model. Both PTT and HD, due to the translation process they assume, predict that the certainty and immediacy manipulations should cause preferences to shift in the same direction for all six choice types (Table 3 and 4, rows 3 and 4). However, focusing on the three choice types where risk and delay are traded off against each other (bottom left, top right, and middle right panels) we instead see that the immediacy (yellow points) and certainty (blue points) preferences shift in opposite directions relative to the baseline (black points). In the certainty condition they shift toward the riskier option, while in the immediacy condition they shift toward the later option, or not at all.

In summary, the magnitude manipulation seemed to shift participants' preferences toward the later option, except when the later option was riskier. Similarly, the immediacy manipulation shifted preferences toward the later option, although not significantly for all choice types. The certainty manipulation, in contrast, shifted preferences toward the riskier option, except when there was no riskier option, in which case the sooner option became more preferred. These results do not perfectly match the predictions of any one model, although the trade-off models' predictions are qualitatively the best across all three manipulations (see Tables 2, 3, and 4 and associated explanation in the online supplemental materials). The immediacy and certainty manipulation results also seem to best match the predictions of the nontranslation models (excluding the PD model which does not predict certainty effects), rather than the two translation models, PTT and HD. However, predicting the correct direction of the effect does not necessarily mean that the model is predicting the correct change in behavior. For this reason, the next section will focus on the models' quantitative predictions.

[10] Using an alpha level of 0.05, the same results are obtained if we calculate for each choice type the proportion of times each participant chose Option 1 in each condition and compare each manipulation to the baseline using Wilcoxin signed rank tests.

[11] To control the family wise error rate across our six tests of the magnitude effect we used the Holm-Bonferroni method (Holm, 1979) to calculate adjusted *p*-values. Using this method the unaltered *p*-values are first ordered from lowest to highest, then adjusted values calculated using $p_{adj(i)} = min(1, \max_{j \leq i} ([m - j + 1] \cdot p_{(j)}))$, where m is the total number of tests, in this case six, $p_{(j)}$ is the jth lowest p-value, and $p_{adj(i)}$ is the adjusted value of the ith lowest p-value. The same method was used to calculate adjusted *p*-values for the immediacy and certainty effect tests.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Model Comparison

Fitting procedure. In order to test the models' ability to predict the effects of each of the three critical manipulations, we need to know which parameter values are most appropriate for each model. In standard Bayesian model selection using Bayes factors, the predictions of the models come from the joint specification of the likelihood function for the model and the prior distributions placed on parameters. In our situation, however, our knowledge of each of the models differs tremendously, making it unclear how best to define prior distributions for some of the models. For example, some of the models have never been fit to risky intertemporal data, making it difficult to know which parameter values are likely to provide reasonable predictions for participants' choices. Our solution to this problem is to use the choices in the baseline condition to obtain posterior distributions for the parameters of all models. We then use these posterior distributions as prior distributions for their respective models to generate predictions for each of the three manipulation data sets.

Prediction as a natural penalty for model complexity. We chose to use this Bayesian method of model comparison because Bayes factors, by their very nature, include a penalty for model complexity. As we will describe in detail shortly, Bayes factors are based on how well the model fits the data across their whole parameter space, weighted by the priors placed on those parameters. So, unlike methods that take into account only the best-fitting combination of parameter values, Bayes factors have a built-in penalty for more complex models. Therefore, a complex model that fits the data very well for one particular subset of parameter combinations, but poorly for many other combinations (permitted by the priors) will perform worse than a simpler model that fits moderately well for all its possible parameter combinations.

Our approach should provide a balanced approach to the issue of the different parametric complexity of our set of models. For instance, consider a relatively flexible model that captures the observed baseline behavior with a very limited set of parameter values. Such a model will have its parameter space restricted by the resultant prior distribution used to predict data in the manipulation conditions. If behavior in the manipulation choice sets are then consistent with the model's predictions based on baseline behavior, then this model will perform well, despite the fact that it could have predicted other patterns with other parameter values if the baseline behavior was different. Conversely, if the manipulation behavior is inconsistent with the baseline behavior, but is consistent with other possible parameterizations of the model, this model will perform poorly, as the prior distributions will neglect these other parts of the model's parameter space.

Estimating parameters for Baseline behavior. We fit hierarchical versions of each of the models to the baseline data. Hierarchical models represent a compromise between the independent estimation of parameters for each individual, and pooling the data from all individuals to obtain a single parametrization of each model. In hierarchical modeling, while individual parameter values are estimated for each participant, these individual-level parameters are assumed to be drawn from a group-level distribution. Hierarchical models allow individual differences to be captured, while also using the group performance to constrain individual estimates (see Nilsson, Rieskamp, & Wagenmakers, 2011). At the individual level, model fit is given by the likelihood of the observed response for each of the 96 choices, such that no aggregation was done at the level of choice type.

For all parameters in all models, the individual-level parameters were assumed to be drawn from normal distributions, $N(\mu, \sigma)$, where μ and σ were group-level parameters—hyperparameters—which are also estimated from the data. Depending upon the particular parameter, these normal distributions were truncated to match the restrictions placed on the values particular parameters could take.[12] Appendix A presents the distributions used for each parameter for each model in the current analysis.

The parameters for all models were estimated using Bayesian methods (in JAGS; Plummer, 2003). Rather than estimating a single best value for each parameter, Bayesian parameter estimation results in a posterior distribution of values for each parameter. To create these posterior distributions for each model we ran four chains of 200,000 iterations each, saving every fourth iteration of each chain. Bayesian parameter estimation requires that we set a prior distribution for each of our hyperparameters. When estimating parameters for the baseline condition, because we were only interested in parameter estimation, we used vague, broad uniform prior distributions for all hyperparameters. Appendix A lists the priors for the hyperparameters. For bounded parameters we placed a uniform prior on μ equal to the possible range of values. For exponent parameters a (Equation 1), s_d (Equation 6), s_r and c (Equation 7), we placed a uniform prior on μ of $U(0, 10)$, as in all cases we expected the mean value to be below 1, but wanted to allow the possibility of increasing sensitivity. For all other μ parameters we based the prior roughly on the literature, by allowing for values from the minimum allowed, usually 0, up to approximately 100 times those reported previously in the literature (Baucells & Heukamp, 2010; Ericson et al., 2015; Luckman et al., 2018; Myerson et al., 2011; Scholten et al., 2014; Vanderveldt et al., 2015). The prior on all σ parameters was always set from 0 to the width of the prior on the corresponding μ parameter.

Once the posterior distributions for all parameters were estimated from the baseline data, they were then used to test the predictions of each model for behavior in the manipulation data sets. In short, the posteriors from the baseline set were used as prior distributions on parameters in the manipulation data sets. As such, each model generates a full distribution of predicted choices for each of the three data sets. To compare the overall predictive accuracy of the models, we use the marginal likelihood of each model, in each of the three manipulation data sets. The marginal likelihood is the probability of the data, integrating over the prior probability of all possible parametrizations of the model. The marginal likelihood is then used to calculate Bayes factors comparing particular pairs of models.

[12] A version was also run using more complex assumptions about the distributions from which parameters were drawn. Discount rate parameters are commonly found to be positively skewed (Myerson, Green, & Warusawitharana, 2001), with logarithmic transformations often proposed (Kim & Zauberman, 2009), for this reason all discount rate parameters (h^{PTHD}, h_r^{MHD}, h_d^{MHD}, h^{HD}, i^{PTHD}) were lognormally, rather than normally, distributed. We also followed the method employed by Nilsson et al. (2011) for dealing with parameters which are constrained to be between zero and one (r^{PTHD}, γ^{PTT}). This involves transforming these parameters to the probit scale before assuming they are drawn from unconstrained normal distributions. The results from these versions were similar to the results we report here.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Estimating marginal likelihoods for manipulation conditions. There are many methods for calculating the marginal likelihood of a model, the method we employed is similar to the sampling method outlined in Vandekerckhove, Matzke, and Wagenmakers (2015). In this method, values for each parameter for each individual are randomly sampled from the appropriate prior distribution (here, the posterior distributions estimated from the baseline data), and the likelihood is calculated for each parameter set. Note that although the parameters used for the three manipulation data sets are the same as those estimated from the baseline data, the data used to evaluate the likelihood of those parameters come from the manipulation data sets. That is, our process tests how well the parameters estimated for the baseline set can predict the data observed in the three manipulation sets.

Once sufficiently many samples are drawn, we average the entire set of likelihoods over parameter values, which yields the marginal likelihood of the model. Here, we sampled parameter sets directly, and exhaustively, from the posterior distributions obtained for the baseline data set, thus preserving the covariance between parameters, and the hierarchical structure of our models. This means that each marginal likelihood is based on 200,000 samples (i.e., the number of iterations conducted to build the posterior distributions of each parameter).[13] Each sampled parameter set consisted of a sample of participant level parameter values for each individual participant. Repeating this sampling procedure separately for each of the three manipulations, resulted in three separate marginal likelihoods for each model.

As well as calculating the likelihood for each sample parameter set, we also calculated prior predictives for each model in each manipulation data set. The prior predictives are simply the predicted distribution of choices generated from each parameter set. Practically, creating prior predictives involved simulating the behavior of each individual for each choice, given the parameters sampled for that particular individual. We also obtained posterior predictives for the baseline dataset using the same method.

Baseline. Figure 2 shows the posterior predictives of each model for the baseline data. As with Figure 1 each choice type is plotted in a separate panel, and each panel is divided into multiple segments. Each segment in Figure 2 shows both the observed choice proportion for each choice, given by the black points, and the predictions for one of the seven models, given by the gray crosses and lines (see the *y*-axis label for a definition of choice proportion for each choice type). The black points match the triangles shown in Figure 1. As the models were directly fit to this data, the predictions here are posterior predictives; that is, they are predictions based on the posterior distributions of parameters after they have been fit to these data. The crosses show the median prediction from each model, while the lines give the range from .01 to 0.99 quantile. As expected, given that the models were fit to this dataset, all models appear to capture the observed behavior in the baseline dataset.

Magnitude effect. Figure 3 gives a visual impression of how well each model captures the data from the magnitude manipulation condition. Similar to Figure 2, the black points are the choice proportions for the given choice type in the magnitude condition, while the gray lines and crosses in each segment are the choice proportions predicted by a specific model. Unlike the gray in Figure 2, we now plot prior predictives; that is, we plot the predictions of each model based on the parameter settings estimated from the baseline dataset. To be clear, we now plot pure predictions from each model. The parameters of each model were not estimated based on the choices participants made in the magnitude condition.

In order to compare the performance of the models we calculated Bayes factors from the marginal likelihoods of each model. Bayes factors are the ratio of the marginal likelihoods for two models, and so display how much more likely the observed responses are under one model than the other. For instance, a Bayes factor of 2 indicates that the observed responses are twice as likely under one model than another. We calculated Bayes factors comparing each model to the PTHD model. Table 5 shows the natural log of these Bayes factors (lnBF). As we have taken the log, positive values indicate that the listed model performed better than the PTHD model, while negative values indicate that it performed worse than the PTHD model. Because the same reference model, PTHD, has been used for calculating all Bayes factors you can calculate the lnBF comparing any other two models (e.g., RITCH and MHD), by taking the difference of the lnBFs reported in the table. For instance, the lnBF comparing the RITCH model to the MHD model for the magnitude manipulation is 227, meaning that the evidence is stronger for the RITCH model than the MHD model. Further, because of the common reference model, the model with the highest lnBF in any column is the model with the best predictions, and the model with the lowest lnBF gave the worst predictions. If we chose a different model to act as the comparison model the ordering of models would be preserved.

In addition to calculating BFs comparing each model to the PTHD model, we also calculated two BFs comparing classes of model. Row 8 of Table 5 shows the lnBFs for each condition comparing the class of attribute-comparison models (RITCH and PD) to the class of utility-comparison models (PTHD, MHD, HD, PTT). We exclude the trade-off model from either grouping, as it uses a mixture of attribute-comparisons and utility calculations. Positive values indicate stronger evidence for an attribute-comparison process while negative values indicate evidence for utility-comparison process. Row 9 similarly shows the lnBFs comparing the class of nontranslation utility-comparison models (PTHD and MHD) to the class of utility-comparison models with a translation of risk into delay (HD and PTT). As this comparison focuses on the presence or absence of a translation process within a utility model, the three attribute-comparison models are excluded. Negative values indicate support for the translation of risks into delays. For these two rows of BFs we assumed equal priors for each class of models, and calculated the marginal likelihood by averaging the marginal likelihood for all models in a class. These

[13] One risk with sampling directly from the prior is that there may be sets of parameter values which greatly affect the marginal likelihood, because they provide a good fit to the data, but are unlikely to be sampled, as they are unlikely under the prior. This can result in an unreliable estimate of the marginal likelihood, as the estimate will vary substantially depending upon whether this unlikely parameter set was included in the sample or not. This risk becomes worse the fewer samples you take from the prior, as the chance of these sets being excluded increases. In the online supplemental materials we include the results from several analyses we performed to check if this may be an issue/if our number of samples from the prior was sufficient for reliable estimates. We also conducted a limited model recovery study for three of the models, PTT, RITCH, and MHD, to test our method further.

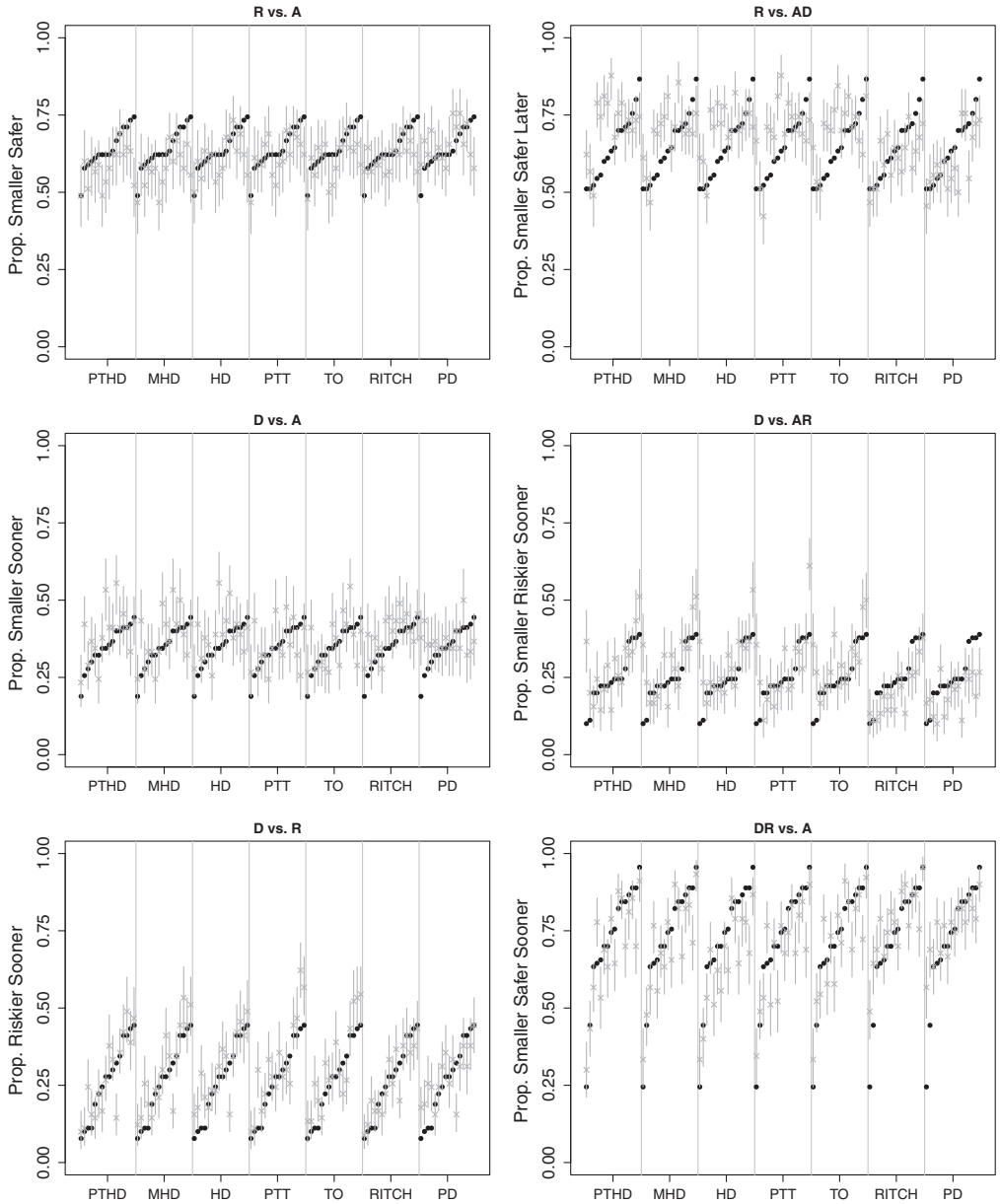


Figure 2. Observed and predicted proportion of participants choosing Option 1 for each choice in the baseline condition. Choices are grouped into panels according to the choice type: Risk vs Amount (RvA), Delay vs. Amount (DvA), Delay vs. Risk (DvR), Risk vs. Amount & Delay (RvAD), Delay vs. Amount & Risk (DvAR), Delay & Risk vs. Amount (DRvA). Each panel is divided into seven segments. Each segment shows the actual data for that choice type (black points) against the predicted proportions from one of the models (gray crosses and lines). The crosses are the median predicted proportions from the model, while the lines are 1 and 99 quantiles. The order of segments/model predictions is: prospect theory with hyperbolic discounting (PTHD), multiplicative hyperboloid discounting (MHD), hyperbolic discounting (HD), probability and time trade-off (PTT), trade-off (TO), the risky intertemporal choice heuristic (RITCH), and proportional difference (PD).

BFs cannot be compared with each other, or with the BFs for individual models.

Column 2 of Table 5 shows the performance of the seven models in the magnitude data. Overall, the RITCH model performs best with the MHD, PD, trade-off, and HD models also performing better than the PTHD model (i.e., they have positive lnBF). Two of the three best performing models, RITCH and MHD (excluding PD), predict that the magnitude will have a directional effect on at least some choice types. However, as shown in Table 2, they do not predict the same effects in all choice types. This difference of prediction is also seen in Figure 3, where one model tends to visually perform worse than the other for certain choice types. For

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

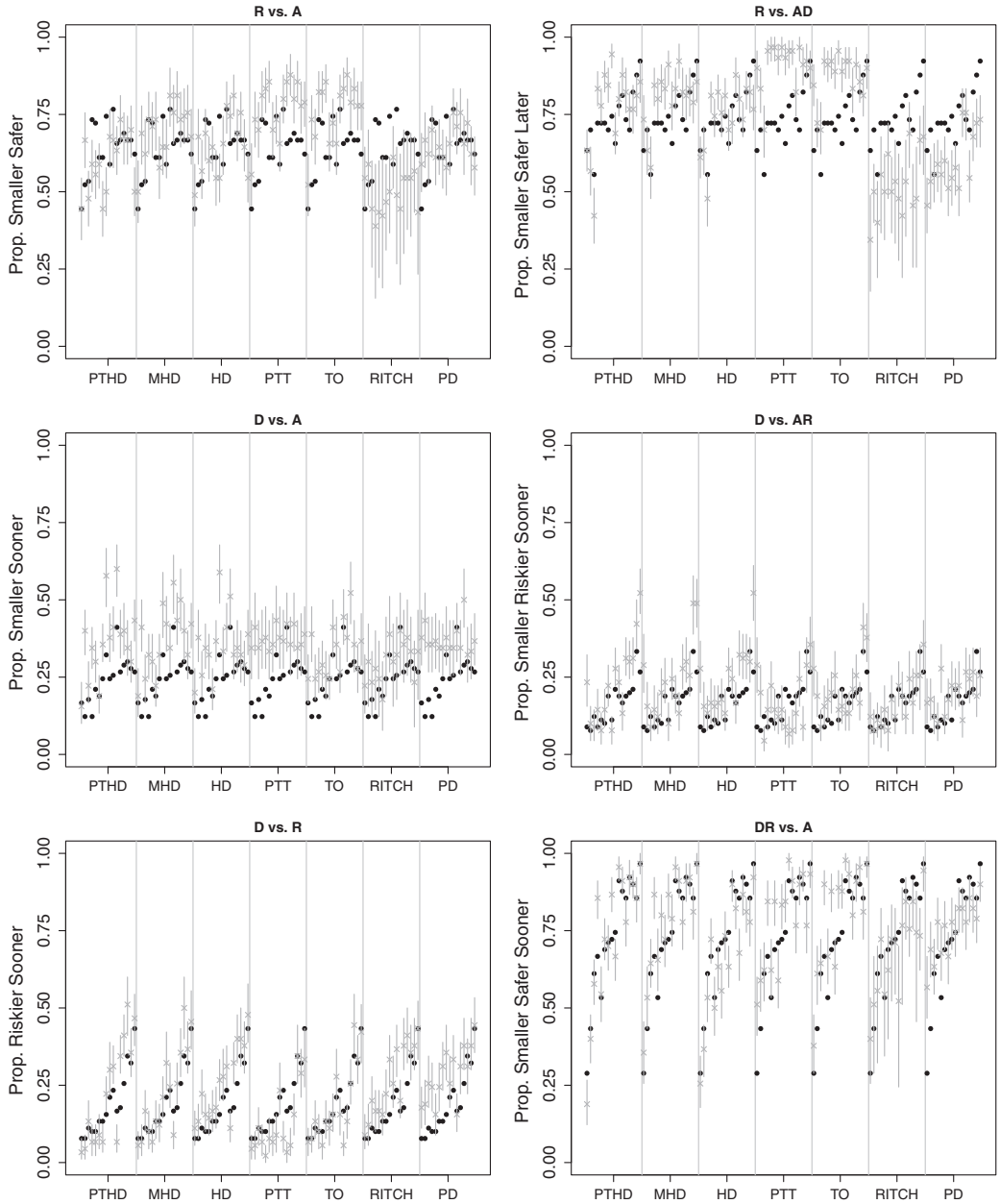


Figure 3. Observed and predicted proportion of participants choosing Option 1 for each choice in the magnitude condition. Choices are grouped into panels according to the choice type: Risk vs Amount (RvA), Delay vs. Amount (DvA), Delay vs. Risk (DvR), Risk vs. Amount & Delay (RvAD), Delay vs. Amount & Risk (DvAR), Delay & Risk vs. Amount (DRvA). Each panel is divided into seven segments. Each segment shows the actual data for that choice type (black points) against the predicted proportions from one of the models (gray crosses and lines). The crosses are the median predicted proportions from the model, while the lines are 1 and 99 quantiles. The order of choices within each segment matches Figure 1. The order of models on the *x*-axis matches Figure 2. PTHD = prospect theory and hyperbolic discounting; MHD = multiplicative hyperboloid discounting; HD = hyperbolic discounting; PTT = probability-time-trade-off; RITCH = risky intertemporal choice heuristic; PD = proportional difference.

instance, in the DvA choices (middle left panel) RITCH appears to perform much better than MHD, as it correctly predicts that participants' preferences will shift toward the more delayed option. In contrast, RITCH does not perform as well as the MHD model in the RvAD choices (top right panel). Here, RITCH predicts a shift toward the larger option, while MHD (qualitatively, but not necessarily quantitatively) predicts the observed shift toward the smaller safer later option.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 5
Logarithm of the Bayes Factor for Each Model Compared With the PTHD Model

Model	All conditions	Magnitude	Immediacy	Certainty
PTHD	0	0	0	0
MHD	1198	923	48	200
HD	242	521	−103	−199
PTT	−396	−159	−207	−73
Trade-off	160	527	−30	−358
RITCH	**1646**	**1148**	**145**	**321**
PD	1148	1134	64	−148
Attribute versus utility	449	226	97	122
Nontranslation utility versus translation utility	956	403	151	273

Note. Higher values indicate better performance, with positive values indicating the model performed better than PTHD and negative values indicating it performed worse. PTHD has a lnBF of 0, as comparing a model to itself results in a BF of 1. The best-performing model in each condition is bolded. We assume an equal prior probability for all models in these calculations. The final two rows show lnBFs comparing classes of models. Row 8 shows the lnBF comparing the attribute class of models (RITCH and PD) to the utility class (PTHD, MHD, HD and PTT), while row 9 shows the lnBF comparing the nontranslation utility models (PTHD and MHD) to the translation utility class (HD and PTT). For these calculations we assume equal prior probabilities for each model within a class, and equal priors for each class. PTHD = prospect theory and hyperbolic discounting; MHD = multiplicative hyperboloid discounting; HD = hyperbolic discounting; PTT = probability-time-trade-off; RITCH = risky intertemporal choice heuristic; PD = proportional difference.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

The width of the predictions of the models should also be noted. While the median predicted proportions generated by the MHD model in the RvAD choices is much closer to the observed proportions than the median proportions generated by the RITCH model, MHD also makes a much more limited range of predictions. Because RITCH predicts a much larger range of proportions, the RITCH model will not be penalized as heavily for incorrect predictions. This highlights the importance of considering the quantitative, not just qualitative predictions of the model. It should also be noted that Figure 3 displays the group level predictions of each model, while the marginal likelihoods, and therefore Bayes factors, are based on each model's ability to predict individual data points. Therefore, while Figure 3 provides a reasonable visualization of the models' performances, it does not precisely reflect the Bayes factors.

The PTT and trade-off models also predict magnitude effects, although they perform worse than MHD and RITCH. This is despite them making the correct qualitative prediction for the majority of choices (see Table 2). In the case of PTT, it also performs worse than both PTHD and HD, which do not predict magnitude effects (see Table 2). This poor performance suggests that the way PTT captures magnitude effects is flawed, as it appears to predict too large an effect in some choice types (e.g., bottom left and top right panels) and too small an effect for other choice types (middle left). The trade-off model shows a similar misfit, though to a lesser extent. The HD model also performs better than the PTHD model, despite the two models making qualitatively similar predictions regarding magnitude effects.

Comparing the performance of the RITCH and PD models is also informative, as PD is very similar to RITCH except without the absolute comparisons, and using the maximum value as the reference point. These are the two best performing models, despite making qualitatively different predictions, with RITCH predicting magnitude effects, due to the consideration of absolute differences in outcomes, while PD does not. The good performance of both models is also reflected in the positive lnBF comparing the classes of attribute- and utility-comparison models (column 2, row 8). The strong performance of the PD model could partly be due to those choice types in which RITCH incorrectly predicts the magnitude effect, such as the RvA or RvAD choices (top panels). However, both HD and PTHD models make the same qualitative predictions as the PD model for all choice types (see Table 2), and neither of these utility-based models performs well. The poor performance of the HD and PTHD models, relative to PD, may indicate an overall advantage for the attribute-based comparison process over a utility-based process. Alternatively, the PD model may be performing so well, despite failing to predict magnitude effects, because the data is better captured by using the maximum as a reference point as opposed to the mean.

Individual fits. As the models were fit hierarchically, in addition to comparing the models at the group level, we can also compare the performance of the models for each individual by calculating the marginal likelihood, and therefore BFs, for each participant for each model. In Table 6 we report the number of participants for which each model had the highest BF. For the magnitude manipulation, the individual level results broadly match the group level BFs, with the majority of participants best fit by one of the RITCH, PD, or MHD models, although the PD model accounts for the single largest number.

Immediacy effect. Figure 4 shows the prior predictives for the immediacy manipulation choices. Similar to the magnitude effect, the RITCH, MHD, and PD models perform better than PTHD, while PTT performs worse. Unlike the magnitude manipulation, PTHD outperforms the HD and the trade-off model in the immediacy manipulation. Looking at column 3 of Table 5, there appears to be a relatively clear divide between the five models which predict that the immediacy manipulation affects the intertemporal component of the choice—RITCH, MHD, PTHD, PD, and trade-off (see Table 3), and the translation models, which predict the immediacy manipulation also affects the risky component of the choice. This pattern is also present in individual level fits in Table 6, with only 16 out of 90 participants best fit by the translation

Table 6
Number of Participants (out of 90) Best Fit by Each Model According to Individual Level BFs

Model	All conditions	Magnitude	Immediacy	Certainty
PTHD	1	5	8	4
MHD	19	16	8	19
HD	1	7	8	3
PTT	2	8	8	1
Trade-off	12	11	9	7
RITCH	**38**	19	**29**	**41**
PD	17	**24**	20	15
Attribute versus utility	63	54	59	62
Nontranslation utility versus translation utility	82	62	64	73

Note. The best-performing model in each condition is bolded. The second last row shows the number of participants who are best fit by the attribute model over the utility model, if we consider the two attribute models (RITCH and PD) a single model with equal prior and the four utility models (PTHD, MHD, HD and PTT) a single model with equal prior. The last row shows the number of participants (out of 90) best fit by a combined nontranslation utility model (PTHD and MHD) compared with a combined translation utility model (HD and PTT). PTHD = prospect theory and hyperbolic discounting; MHD = multiplicative hyperboloid discounting; HD = hyperbolic discounting; PTT = probability-time-trade-off; RITCH = risky intertemporal choice heuristic; PD = proportional difference.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

models. If we look at the class level lnBF comparing the two types of utility models the evidence also favors no translation of risks into delays (Table 5: column 3, row 9).

As with the magnitude condition, if we compare the attribute- and utility-comparison model classes we find support for the attribute-comparison process (Table 5: column 3, row 8). At the level of individual fits the two attribute models also account for 49 participants between them. Comparing the RITCH model and PD model suggests that the former model, which uses both relative and absolute differences in attribute levels, is necessary, at least for many participants. At the group level, the RITCH model performs much better than PD, despite both qualitatively predicting the same effects, suggesting that incorporating absolute differences into a model's behavior may be necessary to temper the effects predicted by the relative changes in attribute values. However, it may not be that all participants consider absolute values, as the PD model does the second-best job of predicting individual participant's choices.

Certainty effect. The results of the certainty choices are similar to the immediacy choices. As with the immediacy choices the MHD and RITCH models outperform PTHD, while the trade-off, PTT, and HD models are outperformed by PTHD (see Figure 5 and column 4 of Table 5). However, unlike for the immediacy manipulation, the PD model is outperformed by PTHD. It should be noted that the PD model is the only model that does not predict any certainty effects. Of the models that do predict certainty effects, the models that locate the effect in the risky component of choice (MHD, RITCH, and PTHD; see Table 4), outperform the two translation models—PTT and HD, which assume certainty also affects the intertemporal component. Isolating the effect of certainty to risk also gives rise to the superior performance of nontranslation over translation models at a class level (Table 5: column 4, row 9). The trade-off model, which performs the worst, also assumes the certainty manipulation affects intertemporal preferences, albeit as a consistent preference for the sooner option.

The individual-level results are again broadly consistent with the group level (see Table 6, column 4). RITCH is best fitting for almost half of participants, with the MHD model accounting for the next highest number of participants. Unlike the group level results, the PD model accounts for the third largest number of participants, perhaps suggesting that there is a small subset of participants who do not show a certainty effect.

Overall. One issue with comparing model performance on each manipulation separately is that a model may perform well for each effect, but poorly for all three at once. Column 1 of Tables 5 and 6 shows the model comparison results when BFs are calculated based on the fits to all three manipulation data sets at once. The RITCH model has the highest lnBF at the group level, and best predicts data for the largest number of participants, with the MHD and PD models being the next best performing. The class-level lnBFs are also consistent with the separate results from the three manipulations, with attribute-comparisons favored over utility models, and models assuming no translation of risk and delay being supported over models with a translation.

Parameter estimates. Table 7 shows the posterior parameter estimates for the group-level μ values for each of the parameters of each model obtained from the fits to the baseline data (an equivalent table for σ can be found in the online supplemental materials). For each μ parameter we report the median and the central 95% credible interval of the posterior distribution for that parameter. Table 7 thus shows what values we believe the mean of each parameter to be, having now seen the baseline data. In the table the parameters for each model are on a separate row, but we group together parameters from similar functions (e.g., value functions, discounting functions, etc.) into columns. These estimates reveal several important insights. First, focusing on the models that use a power value function—PTHD, HD, and MHD—we can see in the value column that μ_a is less than 1, resulting in a concave value function at the group level. The concavity of the value functions indicates that, at the group level, the participants show diminishing sensitivity to amounts, (e.g., they are more sensitive to the difference between \$1 and \$2 than between \$101 and \$102). Diminishing sensitivity to amount is consistent with prior research into risky choice (Tversky & Kahneman, 1992; Stott, 2006), although not necessarily intertemporal choice (Abdellaoui, Bleichrodt, l'Haridon & Paraschiv, 2013). In pure risky-choice situations, this pattern is often characterized as risk aversion. The value

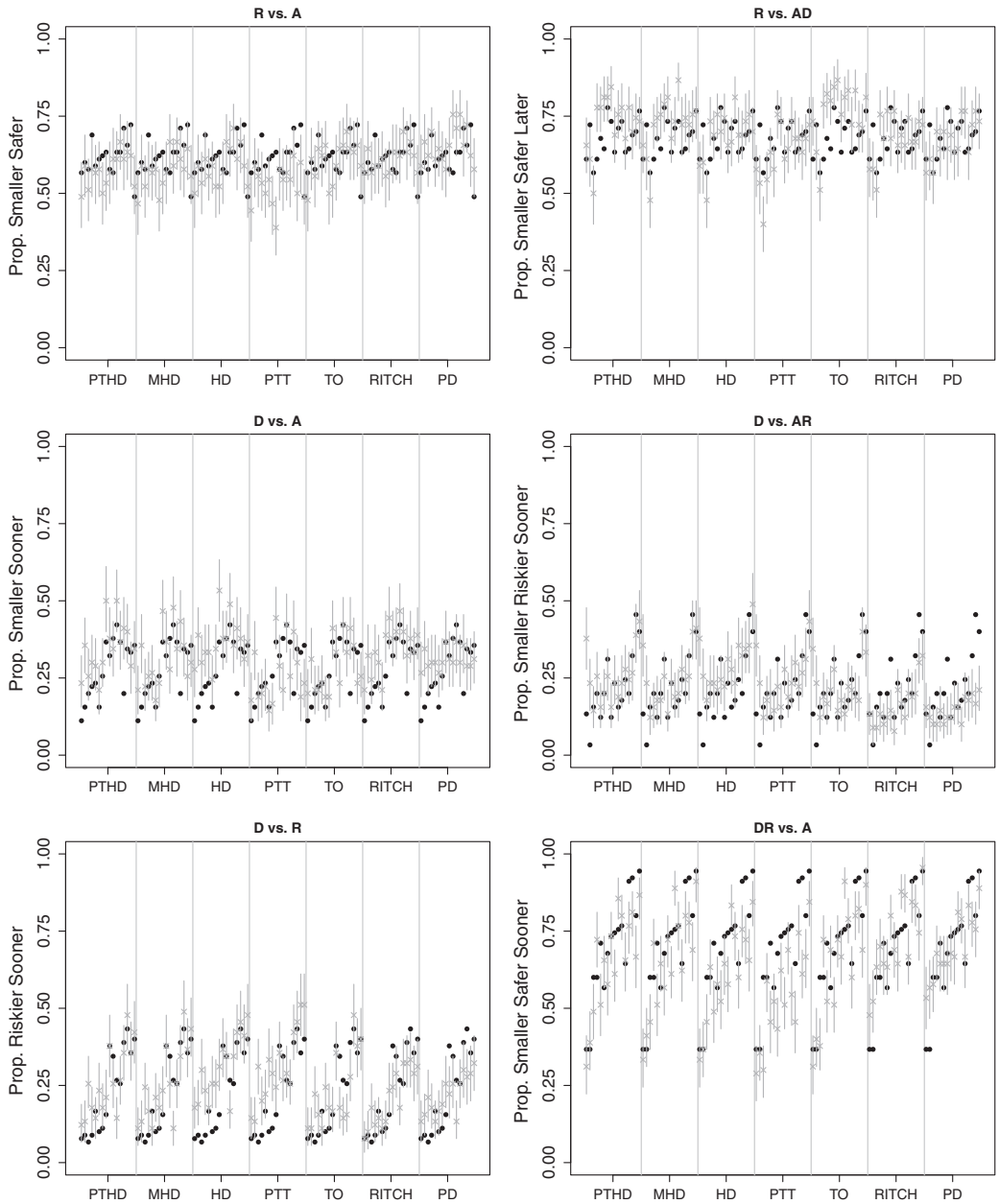


Figure 4. Observed and predicted proportion of participants choosing Option 1 for each choice in the immediacy condition. Choices are grouped into panels according to the choice type: Risk vs Amount (RvA), Delay vs. Amount (DvA), Delay vs. Risk (DvR), Risk vs. Amount & Delay (RvAD), Delay vs. Amount & Risk (DvAR), Delay & Risk vs. Amount (DRvA). Each panel is divided into seven segments. Each segment shows the actual data for that choice type (black points) against the predicted proportions from one of the models (gray crosses and lines). The crosses are the median predicted proportions from the model, while the lines are 1 and 99 quantiles. The order of choices within each segment matches Figure 1. The order of models on the *x*-axis matches Figure 2. PTHD = prospect theory and hyperbolic discounting; MHD = multiplicative hyperboloid discounting; HD = hyperbolic discounting; PTT = probability-time-trade-off; RITCH = risky intertemporal choice heuristic; PD = proportional difference.

functions used by the other models—PTT and trade-off—explicitly assume diminishing sensitivity to amount, so it is reassuring that this assumption is supported by those models that have the flexibility to refute it.

Looking further at Table 7 we can also compare the probability-weighting function parameters to those commonly reported in the literature. For models with a one-parameter probability weighting function (PTHD, PTT, trade-off), we see that the $\mu_{\gamma/r}$ parameters

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

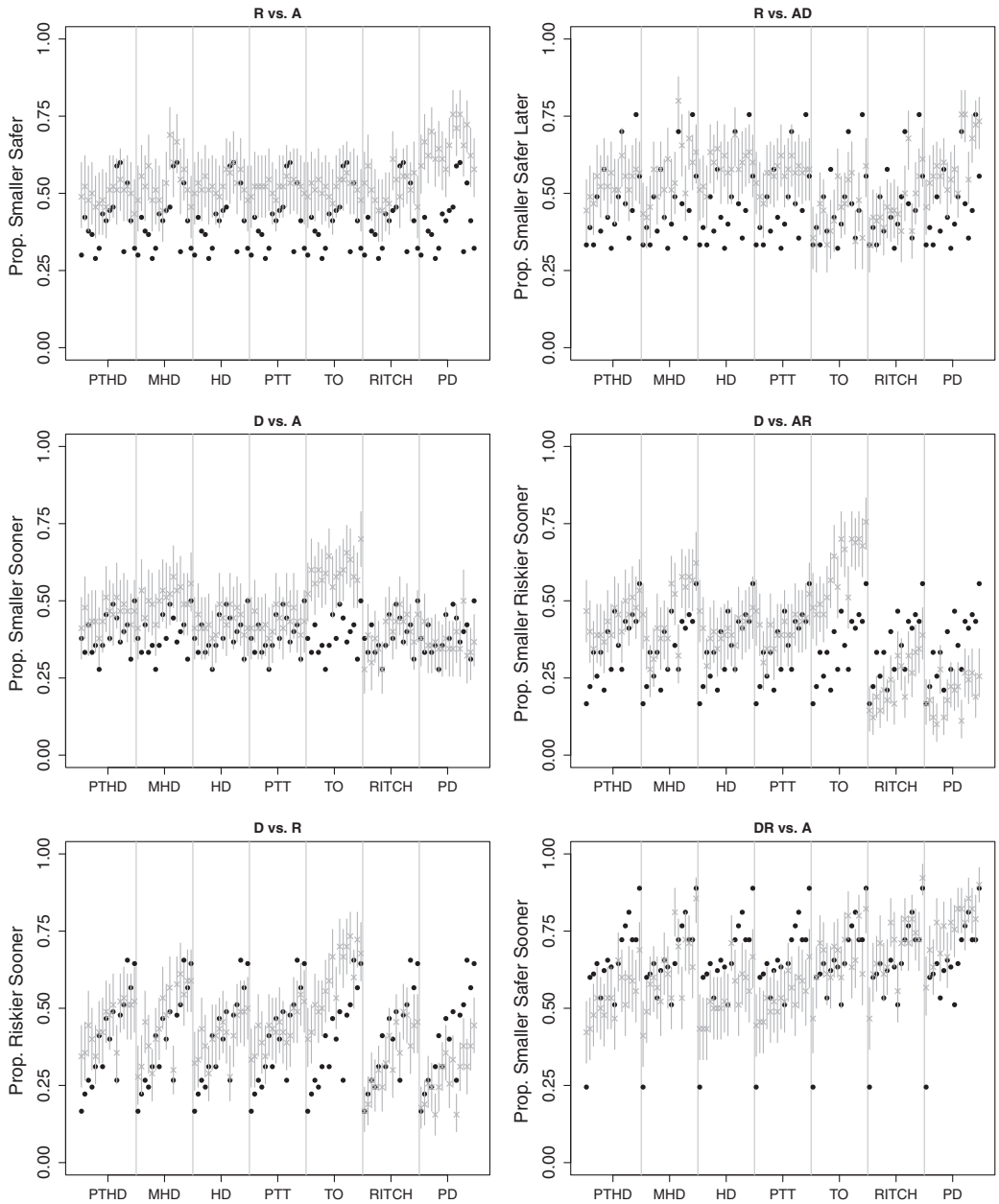


Figure 5. Observed and predicted proportion of participants choosing Option 1 for each choice in the certainty condition. Choices are grouped into panels according to the choice type: Risk vs Amount (RvA), Delay vs. Amount (DvA), Delay vs. Risk (DvR), Risk vs. Amount & Delay (RvAD), Delay vs. Amount & Risk (DvAR), Delay & Risk vs. Amount (DRvA). Each panel is divided into seven segments. Each segment shows the actual data for that choice type (black points) against the predicted proportions from one of the models (gray crosses and lines). The crosses are the median predicted proportions from the model, while the lines are 1 and 99 quantiles. The order of choices within each segment matches Figure 1. The order of models on the *x*-axis matches Figure 2. PTHD = prospect theory and hyperbolic discounting; MHD = multiplicative hyperboloid discounting; HD = hyperbolic discounting; PTT = probability-time-trade-off; RITCH = risky intertemporal choice heuristic; PD = proportional difference.

are generally estimated to be close to 1, suggesting that at the group level, participants use almost linear/objective weighting of probabilities. Based on the literature, we might expect larger deviations from linear weighting due to participants overweighting small probabilities and underweighting moderate ones (Tversky & Kahneman, 1992; Nilsson et al., 2011), however, such small departures are not unprecedented (Luckman et al., 2018; Stott, 2006).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 7
Posterior Distributions for Group Level (μ) Parameters for Each Model

	μ values (median [95% CI])								
	Value		Probability weighting			Discounting		Choice
Model	a		r			h		s
PTHD	0.35 [0.21, 0.41]		0.97 [0.88, 1.00]			0.0018 [7e-05, 0.0086]		0.20 [0.008, 0.87]
	a		h_r	s_r	c	h_d	s_d	s
MHD	0.17 [0.12, 0.21]		1.71 [0.08, 4.49]	0.10 [0.06, 0.16]	0.12 [0.07, 0.17]	0.022 [0.0009, 0.074]	0.060 [0.003, 0.200]	7.57 [4.85, 12.47]
	a		i			h		s
HD	0.28 [0.12, 0.34]		21.60 [0.88, 79.62]			0.0035 [0.0001, 0.0120]		0.42 [0.02, 1.52]
	α	β	γ			R		s
PTT	0.013 [0.0006, 0.036]	0.32 [0.26, 0.38]	0.96 [0.88, 1.00]			0.28 [0.01, 1.33]		0.29 [0.02, 0.61]

	Value	Prob. weighting	Time weighting				Choice
	β	γ	τ	ϑ	κ	α	s
Trade-off	0.038 [0.002, 0.114]	0.97 [0.90, 1.00]	0.17 [0.007, 0.57]	2.80 [1.13, 4.09]	60.61 [6.48, 112.21]	34.04 [1.59, 93.41]	0.048 [0.002, 0.197]

	Bias			Relative weight			Absolute weight		
	β_{xO}	β_{tO}	β_{pO}	β_{xR}	β_{tR}	β_{pR}	β_{xA}	β_{tA}	β_{pA}
RITCH	0.30 [0.01, 0.72]	0.021 [0.0009, 0.091]	0.27 [0.01, 0.86]	1.42 [0.26, 1.92]	0.045 [0.002, 0.190]	0.61 [0.03, 1.62]	0.00021 [9e-06, 0.00087]	0.012 [0.0005, 0.040]	0.0063 [0.0003, 0.0229]
PD	0.16 [0.007, 0.49]	0.058 [0.003, 0.185]	0.21 [0.009, 0.74]	0.38 [0.02, 1.30]	0.12 [0.005, 0.51]	0.97 [0.04, 2.89]			

Note. Median and 95% equal-tailed credible interval reported for each parameter of each model. Parameters for the utility and trade-off model are grouped according to the type of function (e.g. value function) they are associated with. For attribute models, they are grouped according to the type of difference (bias, relative, absolute) they are weighting. PTHD = prospect theory and hyperbolic discounting; MHD = multiplicative hyperboloid discounting; HD = hyperbolic discounting; PTT = probability-time-trade-off; RITCH = risky intertemporal choice heuristic; PD = proportional difference.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Turning to attribute-based models, we now focus on the group-level parameter estimates from the best-fitting model, the RITCH model. Focusing on the bias parameters—$\mu_{\beta_{xo}}$, $\mu_{\beta_{po}}$, and $\mu_{\beta_{tO}}$—we see more relative bias, independent of the magnitude of attribute differences, to prefer either the larger or safer option, than the sooner option. This result is similar to the equivalent parameters in the PD model although the estimates for that model suggest that the bias toward the safer option may be stronger than that toward the larger option.

Staying with the RITCH model, if we turn to the weights given to relative differences in each of the attributes—$\mu_{\beta_{xR}}$, $\mu_{\beta_{pR}}$, and $\mu_{\beta_{tR}}$—we see that greater weight is given to relative differences in amount (median = 1.42), compared with relative differences in probability (median = 0.61), and that the least weight is given to relative differences in delay (median = 0.045). On its own, this pattern might suggest that differences in outcome amount are considered the most important when making decisions, and differences in delay the least important. However, an interpretation focusing only on the relative differences ignores the absolute differences in each attribute level in the RITCH model. If we look at the absolute-difference weighting parameters—$\mu_{\beta_{xA}}$, $\mu_{\beta_{pA}}$, and $\mu_{\beta_{tA}}$—we see that a 1-month difference in delay is given more weight (median = 0.012) than a 1% difference in probability (median = 0.0063), which in turn is given more weight than a $1 difference in amount (median = 0.00021). It is worth noting that these parameters should not be interpreted as a measure of the importance given to each attribute, because the weight parameters are sensitive to the scale of measurement used for each attribute. However, the parameters are useful for indicating the approximate equivalence of dollars, months, and percent chance for the specific set of gambles used in the experiment.

As a final point, we note that we focused our attention on models that did not allow parameters to vary freely across the manipulation conditions. We made this choice because we wanted to test the models' ability to predict behavior in such conditions without arbitrary changes in parameter values. However, researchers may wish to expand the models tested here to account for these manipulations. One way of extending such models would be to allow the parameters to vary as a function of the properties we manipulated. To that end, we provide posterior estimates for each model for each of the three manipulation conditions in the online supplemental materials. However, care should be taken when attempting to understand these parameter values, as the attribute ranges in these data sets are quite often extreme as they were not designed for use in estimating the parameters of the various functions. For instance, the certainty dataset contains no choices where the probability was greater than 0.2, therefore the parameter estimates for probability-weighting functions are not penalized if they poorly capture behavior for moderate to high-probabilities.

Discussion

The goal of this article was to better understand how people make risky intertemporal choices. In particular we sought to understand (a) whether choices are made by focusing on comparing attribute values, or comparing utilities, and (b) how risk and delay information is combined. To do this we compared and evaluated the performance of several models of risky intertemporal choice by examining the effects of magnitude, immediacy, and certainty manipulations on a range of risky intertemporal choices. This comparison included three existing utility-comparison models of risky intertemporal choice from the literature. All three of these models, the PTT model (Baucells & Heukamp, 2010), the MHD model (Vanderveldt et al., 2015), and the HD model (Yi et al., 2006) have been argued to provide adequate fits to risky intertemporal data. However, this article documents the first attempt to compare their relative performance, directly test the differing assumptions the models make, and the effects that they predict.

In line with the literature, we found that all three of these models, as well as an untested model from the literature and three additional models we propose, appear to capture behavior when fit directly to risky intertemporal choice data. Crucially, however, we found that when predicting behavior from these models there are clear differences in how well they perform. Across all three manipulations—magnitude, immediacy, and certainty—we found that an attribute-comparison model we proposed, RITCH, which is a modified version of the existing ITCH (Ericson et al., 2015) model of intertemporal choice, predicted participants' behavior better than any other model. The next best performing models tended to be those that made qualitatively similar predictions to the RITCH model in many cases, such as the MHD model, which was the second best performing overall, or PD model. Models which assumed that risk and delay information is combined into a single attribute dimension, the PTT and HD models, performed very poorly.

Attribute Versus Utility Models

Our results add to the growing discussion about whether utility-comparison models are the appropriate way to model choice behavior (Dai & Busemeyer, 2014; Ericson et al., 2015; Scholten & Read, 2010; Stewart, Chater, Stott, & Reimers, 2003; Vlaev, Chater, & Stewart, 2007; Vlaev et al., 2011). Attribute-comparison models, where participants are assumed to directly compare attributes across options and the worth of an option is dependent upon the other options presented alongside it, are often argued to provide a better reflection of the processes underlying choice (e.g., Konstantinidis et al., 2020). Overall, our results suggest that such an attribute comparison process may better capture participants' behavior when making risky intertemporal choices. The RITCH model, an attribute-comparison model, outperformed all of the utility models, even those that made qualitatively similar predictions about the magnitude, immediacy, and certainty manipulations. It also outperformed the trade-off model, which included a mixture of attribute and utility processes. Furthermore, the other attribute comparison model we considered, PD, outperformed all utility models in the magnitude manipulation, despite predicting no effects of magnitude, which suggests that even when utility models predict the correct changes in behavior they may not accurately reflect the processes underlying participants' decisions.

Magnitude effect. Though the RITCH model performed best in the magnitude manipulation, the actual pattern of results we observed is problematic for a pure attribute-comparison model. As we outlined in the Introduction, pure attribute-comparison models assume that manipulating the amount dimension only affects the evaluation of that dimension, and thus fail to explain changes in behavior that imply that other dimensions were affected. For instance, RITCH predicts that the magnitude manipulation will

increase participants' preference for the larger option. However, the actual pattern we observe is more consistent with a change in how willing participants were to wait, than an increased attraction to the larger outcome. If we focus on the three choice types where the later option and larger option are not the same option, we find preferences shift toward: (a) the later option over the larger option when the two are in competition (the RvAD choices); (b) the later option when there is no larger option (the DvR choices; see also Baucells & Heukamp, 2010; Christensen et al., 1998; Luckman et al., 2017); and (c) no change in preference when there is a larger option but no later option (the RvA choices). The only case where participants do not shift toward preferring the later option is the DRvA choices, where rather than shifting toward the larger later riskier option (cf. Yi et al., 2006), they show no change in preference.

Leaving aside the DRvA choices, the shift toward preferring the later option we observe is more consistent with the explanation for the magnitude effect suggested by the PTT model, where discount rates decrease as a function of the amount offered. As such we might expect that the PTT model would capture the magnitude data better than RITCH. There are three plausible reasons why this might not be the case: (a) the benefit of assuming an attribute-comparison process may outweigh the benefit of predicting the correct magnitude effect, (b) the detriment of assuming that delays are translated into risks might outweigh the benefit of correctly predicting the magnitude effect, and (c) the PTT model is penalized for incorrectly predicting a peanuts effect in the RvA choices. To eliminate Explanation b and explore Possibilities a and c further, we fit two additional variants of the MHD model to our data, as MHD does not assume a translation of delays into risks. These variants assumed that the discount rate for delays decreases as a function of amount, similar to the PTT model, using the relationship proposed by Vincent (2016):

$$d(t,x) = \frac{1}{(1+x^m h_d t)^{s_d}} \tag{27}$$

where $m < 0$. As the original MHD model also predicts a peanuts effect, to eliminate Explanation c, one of the new variants also dropped the assumption that risk sensitivity changed as a function of amount (see Equation 5). In Table 8 we report the lnBFs for these two new models, as well as the lnBFs for the original MHD and RITCH models. As we can see in column 2, the two new MHD models perform better than the original model for the magnitude manipulation. The MHD variant that does not predict a peanuts effect also outperforms the RITCH model. However, they still perform worse than RITCH model when all three effects are taken into account (column 1). Taken together, these results suggest that we cannot draw the simple conclusion that attribute-comparison models provide a better account of our data than utility-comparison models. Rather, our results suggest that the strict distinction between the two types of processes may not be helpful. While the RITCH model, an attribute-comparison model, clearly provides the best account of our data, the magnitude manipulation results suggest that we may need to weaken the assumption that each attribute is compared separately, as outcome magnitude seems to reduce the importance of delays.

Relative versus absolute differences. Within the literature on attribute-comparison models there is also discussion as to whether attributes are compared at an absolute, or direct, level (Dai & Busemeyer, 2014; Scholten & Read, 2010) or whether relative differences in attribute levels are considered (Cheng & González-Vallejo, 2016; Dai & Busemeyer, 2014; González-Vallejo, 2002). Because the RITCH model we proposed uses both absolute and relative differences, we must ask whether they are both necessary. A comparison of the fit for the PD and RITCH models suggests that consideration of relative difference alone (PD model) is not sufficient, particularly when considering immediacy and certainty effects (compare rows 6 and 7 of Table 5). Although not reported in the main analysis, we also tested a restricted version of the RITCH model that considered only absolute differences in attribute values. This restricted variant performed relatively well for the certainty manipulation (lnBF = 204), though worse than the full RITCH model, moderately for the immediacy manipulation (lnBF = −15), but very poorly in the magnitude condition (lnBF = −4626), even though it used the same mechanism for predicting magnitude effects as the full RITCH model. Overall, this pattern suggests that consideration of both absolute and relative differences are necessary for the RITCH model to capture behavior, with the loss of either leading to worse performance for all three manipulations.

Translation of Delays Into Risks

A recurring theme in the risky intertemporal choice literature is the extent to which risks and time delays are treated as a single attribute dimension. This is reflected in existing models of risky intertemporal choice, PTT and HD, which assume that risks and delays can be transformed into a single dimension, by transforming delays into perceived risks, or probabilities into inferred delays until receipt (Baucells & Heukamp, 2010; Yi et al., 2006). The

Table 8
Logarithm of the Bayes Factor for Each of the Three Variants of MHD, and the RITCH Model Compared With the PTHD Model

Model	Overall	Magnitude	Immediacy	Certainty
MHD-peanuts (original)	1198	923	48	200
MHD-peanuts and magnitude	1434	1144	37	245
MHD-magnitude	1511	**1173**	12	257
RITCH	**1646**	1148	**145**	**321**

Note. Higher values indicate better performance. The best-performing model in each condition is bolded. PTHD = prospect theory and hyperbolic discounting; MHD = multiplicative hyperboloid discounting; RITCH = risky intertemporal choice heuristic.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

primary implication of this assumption, which we test in our experiment, is that increasing either the riskiness of options, such as through the certainty manipulation, or delay of the options, through the immediacy manipulation, should produce similar effects. In our experiment we find no support for this assumption. Instead we found that for all choice types the certainty manipulation shifted participants' preferences toward the riskier option, and the immediacy manipulation usually shifted preferences toward the later option. This result means that when the risk and delay are in competition, as they are in panels 3, 4, and 5 of Figure 1, the two manipulations shift preferences in opposite directions. This result is consistent with risk and delay being treated as separate dimensions, with the immediacy effect acting only on the time dimension and the certainty effect acting only on the probability dimension. If risks and delays are treated as a single dimension we would instead expect the shift in these panels to be in the same direction (see rows 3 and 4 of Tables 3 and 4). The only scenario in which we find the two manipulations affect preferences in the same direction is for the DRvA choices, where the riskier and later options are the same option.

The modeling results reinforce the conclusion from the behavioral results. For both the certainty and immediacy manipulations the models which assume a translation of risks and delays into a single dimension, PTT and HD, perform worse than those which assume they are treated as separate dimensions. The only models they outperform are the PD and trade-off models when they, respectively, predict no and nonstandard certainty effects (see rows 5 and 7 of Table 4). If we focus just on models which assume utility comparisons (PTT, HD, PTHD, MHD), we also see that, as a class, models which treat risks and delays as separate dimensions outperform those that assume they are a single dimension (row 9 of Table 5). To further reinforce the poor performance of the translation assumption, the translation models also perform poorly for the magnitude effect (column 2 of Table 5) and overall only three participants show behavior which is best explained by one of the two translation models (column 1 of Table 6).

Altogether these results suggest that there is not a special relationship between risks and delays to the extent that they can be combined into a single dimension. Instead, it appears that people treat them as two separate attributes with independent effects, just as they do risks and amounts or delays and amounts. This is problematic for any account of intertemporal choice that seeks to explain discounting as due to primarily considerations of the risks inherent in waiting (Sozou, 1998), or an account of risky choice that explains behavior by considering only the average time it would take to achieve an outcome through repeated plays of a gamble (Rachlin et al., 1991; Yi et al., 2006).

However, our results do not rule out considerations of delay playing some part in risky choice, or risk some part in intertemporal choice. For instance it would be perfectly consistent with our results for someone to consider the risks associated with waiting in a decision involving time delays, such as considering the likelihood of the fund provider going bankrupt when choosing how much to invest in a retirement fund. However, we would expect either these risks to not be the primary consideration driving their choice, or to be treated as separate to risks that are external to the delay, such as if the fund provider offered investment options with different risk profiles. Similarly, we might expect someone making a risky decision, such as whether to build on a floodplain, to consider the timescale over which a flood is expected to occur. This would also be consistent with our results as long as it is not the primary way in which they are assessing the risk, or is not treated the same as a guaranteed certain delay.

Conflicting Results

Immediacy and certainty. Existing research into certainty and immediacy effects in risky intertemporal choice is not entirely consistent with our results. According to the majority of the literature, we should have expected the immediacy manipulation to increase the proportion of participants choosing the riskier option in RvA choices (Baucells & Heukamp, 2010; Oshikoji, 2012; Sagristano et al., 2002). Likewise, some literature suggests that the certainty manipulation increases the proportion choosing the later option in DvA choices (Keren & Roelofsma, 1995; Weber & Chapman, 2005a). Instead, we found that the immediacy manipulation had no effect on RvA choices and the certainty manipulation increased preferences for sooner options in the DvA choices. These results are more similar to the choice results of Weber and Chapman (2005a) or the results of Sun and Li (2010). Ignoring the effect of certainty in DvA choices, our results are consistent with explanations given for the immediacy effect in the intertemporal choice literature, such as hyperbolic discounting, diminishing sensitivity to delay (Scholten & Read, 2010), or sensitivity to relative changes in delay (Ericson et al., 2015). Our results are also consistent with explanations for the certainty effect in the risky choice literature, such as overweighting of small probabilities (Kahneman & Tversky, 1979).

When interpreting these different results it is important to remember that the choice context of our task is different to other risky intertemporal choice experiments. In previous tasks, participants were generally asked to consider only a single choice type, RvA or DvA, and were not required to shift between situations where differences in probability were present, to those where differences in delay were present. Our task, on the other hand, required participants to adapt constantly to whether probability or delay differences were present. It could be that in a situation where one attribute is consistently nondiscriminating between options, reducing certainty or immediacy have similar effects on behavior, while in situations where both attributes frequently discriminate, as they do in our task, the two have distinct influences. No existing theory of the task would provide an explanation of why that would be the case, however.

Limitations

By design, we sought to conduct a quantitative comparison of various models of risky intertemporal choice. In order to fit these various models effectively, and discriminate between them, we needed each participant to complete a large number of risky intertemporal choices. This many choices makes our experiment different from those in the literature that use small choice sets or one-off choices. Therefore, our conclusions may only generalize to similar experimental designs. For example, we observe certain "benchmark" effects that are robustly observed in the literature (e.g., magnitude effects, standard immediacy, and certainty effects), but do not observe peanut effects, which appear less reliably in the literature (compare Baucells & Heukamp, 2010; Markowitz,

1952; Myerson et al., 2003; Weber & Chapman, 2005b to Chapman & Weber, 2006; Vanderveldt et al., 2017). If the discrepancy between our findings and some existing empirical results are a consequence of how many trials the participants complete, then our conclusions about model superiority may not apply to shorter trial sequences. Until we know more about the role of experimental design on such effects, we need to maintain care when generalizing our results to smaller choice-set experiments (Li, Wall, Johnson, & Toubia, 2016). That said, because many experiments do use multiple-trial experiments, our results do speak to a large section of the literature. Furthermore, should the results differ when participants make fewer choices, any overarching theory will have to explain why the number of choices matters.

Wider Implications

Many of the difficult choices we face—as individuals and as a society—comprise the explicit and combined evaluation of risks over time. Whether it be a decision to act on emission reductions now in order to offset a future probabilistic benefit (some chance of avoiding a $< 2C$ rise in global temperature) or simply deciding whether to take that second piece of cake (potentially reducing the onset of Type 2 diabetes) there is a pressing need to understand and *predict* how people will choose when faced with uncertainty about the timing and the receipt of different outcomes.

Our comprehensive model evaluation provides important evidence regarding the psychological mechanisms underlying these kinds of risky-intertemporal choices. These results could open the door to interventions that may improve decision-making in such situations. Whether via a "nudge" toward a healthier, more sustainable, or more financially responsible option (Thaler & Sunstein, 2008) or a "boost" to the decision-making competency of an individual (Hertwig & Grüne-Yanoff, 2017), low-cost, subtle changes to information presentation can yield large effects. For example, the clear evidence for the superiority of attribute-comparison models in our evaluation, suggests that presenting information about risks, delays, and outcomes in ways that facilitates comparisons and amplifies both relative and absolute differences between options may make important choices simpler for people (e.g., Bateman, Dobrescu, Ortmann, Newell, & Thorp, 2016; Reeck et al., 2017). By the same token our results suggest that assuming people can translate disparate attributes into a common underlying psychoeconomic scale of utility may be detrimental when designing choice environments (cf. Vlaev et al., 2011).

Conclusion

Overall, we find that magnitude, immediacy, and certainty manipulations all affect participants' preferences across a range of different risky intertemporal choice types. Models that assumed that the immediacy and certainty manipulations respectively affect the intertemporal and risky components of choice, performed better than those which assumed each affected both components. This result suggests that future model development should focus on those models that treat risks and delays separately, rather than treating them as a single attribute dimension. Overall, we found that a model assuming an attribute-comparison process, RITCH, performed best, both overall and for all three manipulations. However, we also note that attribute-comparison models, such as RITCH, cannot capture completely the effect that outcome magnitude had on participants' preferences. Modifying attribute-comparison models to predict an increased preference for the later option when outcome magnitude is increased would be a promising avenue for future research. Finally, we demonstrate a Bayesian method for testing the competing predictions of a range of existing, and new, models of risky intertemporal choice. This same method is readily applicable to new models, or other model comparison problems, where there is limited or unbalanced information for determining priors for the models under comparison.

References

Abdellaoui, M., Bleichrodt, H., l'Haridon, O., & Paraschiv, C. (2013). Is there one unifying concept of utility? An experimental comparison of utility under risk and utility over time. *Management Science, 59,* 2153–2169. http://dx.doi.org/10.1287/mnsc.1120.1690

Bateman, H., Dobrescu, I., Ortmann, A., Newell, B. R., & Thorp, S. (2016). As easy as pie: How retirement savers use prescribed investment disclosures. *Journal of Economic Behavior & Organization, 121,* 60–76. http://dx.doi.org/10.1016/j.jebo.2015.10.020

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

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

Birnbaum, M. H., & LaCroix, A. R. (2008). Dimension integration: Testing models without trade-offs. *Organizational Behavior and Human Decision Processes, 105,* 122–133. http://dx.doi.org/10.1016/j.obhdp.2007.07.002

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

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. http://dx.doi.org/10.3758/BF03193582

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

Christensen, J., Parker, S., Silberberg, A., & Hursh, S. (1998). Tradeoffs in choice between risk and delay depend on monetary amounts. *Journal of the Experimental Analysis of Behavior, 69,* 123–139. http://dx.doi.org/10.1901/jeab.1998.69-123

Dai, J., & Busemeyer, J. R. (2014). A probabilistic, dynamic, and attribute-wise model of intertemporal choice. *Journal of Experimental Psychology: General, 143,* 1489–1514. http://dx.doi.org/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. http://dx.doi.org/10.1007/s11166-011-9129-x

Ericson, K. M. M., White, J. M., Laibson, D., & Cohen, J. D. (2015). Money earlier or later? Simple heuristics explain intertemporal choices better than delay discounting does. *Psychological Science, 26,* 826–833. http://dx.doi.org/10.1177/0956797615572232

González-Vallejo, C. (2002). Making trade-offs: A probabilistic and context-sensitive model of choice behavior. *Psychological Review, 109,* 137–155. http://dx.doi.org/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. http://dx.doi.org/10.1037/0033-2909.130.5.769

Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science, 12,* 973–986. http://dx.doi.org/10.1177/1745691617702496

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Hey, J. D. (1995). Experimental investigations of errors in decision-making under risk. *European Economic Review, 39,* 633–640. http://dx.doi.org/10.1016/0014-2921(09)40007-4

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6,* 65–70.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *The American Economic Review, 92,* 1644–1655. http://dx.doi.org/10.1257/000282802762024700

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

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

Kim, B. K., & Zauberman, G. (2009). Perception of anticipatory time in temporal discounting. *Journal of Neuroscience, Psychology, and Economics, 2,* 91–101. http://dx.doi.org/10.1037/a0017686

Kirby, K. N. (1997). Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology: General, 126,* 54–70. http://dx.doi.org/10.1037/0096-3445.126.1.54

Kirby, K. N., & Marakovic, N. N. (1995). Modeling myopic decisions - evidence for hyperbolic delay-discounting within-subjects and amounts. *Organizational Behavior and Human Decision Processes, 64,* 22–30. http://dx.doi.org/10.1006/obhd.1995.1086

Konstantinidis, E., van Ravenzwaaij, D., Güney, Ş., & Newell, B. R. (2020). Now for sure or later with a risk? Modeling risky intertemporal choice as accumulated preference. *Decision, 7,* 91–120. http://dx.doi.org/10.1037/dec0000103

Li, Y., Wall, D., Johnson, E., & Toubia, O. (2016). Cognitively optimized measurement of preferences. In P. Moreau & S. Puntoni (Eds), *Advances in consumer research* (Vol. 44, pp. 532–535). Duluth, MN: Association for Consumer Research.

Loewenstein, G., & Prelec, D. (1992). Anomalies in intertemporal choice—evidence and an interpretation. *Quarterly Journal of Economics, 107,* 573–597.

Loomes, G., & Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review, 39,* 641–648. http://dx.doi.org/10.1016/0014-2921(94)00071-7

Luckman, A., Donkin, C., & Newell, B. R. (2015). Exploring the concept of utility: Are separate value functions required for risky and intertemporal choice? In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (pp. 1440–1446). Austin, TX: Cognitive Science Society.

Luckman, A., Donkin, C., & Newell, B. R. (2017). People wait longer when the alternative is risky: The relation between preferences in risky and inter-temporal choice. *Journal of Behavioral Decision Making, 30,* 1078–1092. http://dx.doi.org/10.1002/bdm.2025

Luckman, A., Donkin, C., & Newell, B. R. (2018). Can a single model account for both risky choices and inter-temporal choices? Testing the assumptions underlying models of risky inter-temporal choice. *Psychonomic Bulletin & Review, 25,* 785–792. http://dx.doi.org/10.3758/s13423-017-1330-8

Markowitz, H. (1952). The utility of wealth. *Journal of Political Economy, 60,* 151–158. http://dx.doi.org/10.1086/257177

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: The effect of delay and of intervening events on reinforcement value* (pp. 55–73). Hillsdale, NJ: Erlbaum, Inc.

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

Myerson, J., Green, L., & Morris, J. (2011). Modeling the effect of reward amount on probability discounting. *Journal of the Experimental Analysis of Behavior, 95,* 175–187. http://dx.doi.org/10.1901/jeab.2011.95-175

Myerson, J., Green, L., & Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *Journal of the Experimental Analysis of Behavior, 76,* 235–243. http://dx.doi.org/10.1901/jeab.2001.76-235

Nilsson, H., Rieskamp, J., & Wagenmakers, E. J. (2011). Hierarchical Bayesian parameter estimation for cumulative prospect theory. *Journal of Mathematical Psychology, 55,* 84–93. http://dx.doi.org/10.1016/j.jmp.2010.08.006

Oshikoji, K. (2012). *Risky intertemporal choice in the loss domain* (Master's thesis). University of Waterloo, Waterloo, ON.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In K. Hornik, F. Leisch, & A. Zeileis (Eds.), *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*. Retrieved from https://www.r-project.org/conferences/DSC-2003/Proceedings/Plummer.pdf

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

Rachlin, H. (1989). *Judgment, decision, and choice: A cognitive/behavioral synthesis*. New York, NY: WH Freeman.

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

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. http://dx.doi.org/10.1037/a0029177

Reeck, C., Wall, D., & Johnson, E. J. (2017). Search predicts and changes patience in intertemporal choice. *Proceedings of the National Academy of Sciences of the United States of America, 114,* 11890–11895. http://dx.doi.org/10.1073/pnas.1707040114

Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1446–1465. http://dx.doi.org/10.1037/a0013646

Sagristano, M. D., Trope, Y., & Liberman, N. (2002). Time-dependent gambling: Odds now, money later. *Journal of Experimental Psychology: General, 131,* 364–376. http://dx.doi.org/10.1037/0096-3445.131.3.364

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

Scholten, M., & Read, D. (2014). Prospect theory and the "forgotten" fourfold pattern of risk preferences. *Journal of Risk and Uncertainty, 48,* 67–83. http://dx.doi.org/10.1007/s11166-014-9183-2

Scholten, M., Read, D., & Sanborn, A. (2014). Weighing outcomes by time or against time? Evaluation rules in intertemporal choice. *Cognitive Science, 38,* 399–438. http://dx.doi.org/10.1111/cogs.12104

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2019). *Analysis of factorial experiments. R package version 0.24–1*. Retrieved from https://CRAN.R-project.org/package=afex

Sozou, P. D. (1998). On hyperbolic discounting and uncertain hazard rates. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 265,* 2015–2020.

Stewart, N., Chater, N., Stott, H. P., & Reimers, S. (2003). Prospect relativity: How choice options influence decision under risk. *Journal of Experimental Psychology: General, 132,* 23–46. http://dx.doi.org/10.1037/0096-3445.132.1.23

Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty, 32,* 101–130. http://dx.doi.org/10.1007/s11166-006-8289-6

Sun, Y., & Li, S. (2010). The effect of risk on intertemporal choice. *Journal of Risk Research, 13,* 805–820. http://dx.doi.org/10.1080/13669871003606224

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Takahashi, T., Ikeda, K., & Hasegawa, T. (2007). A hyperbolic decay of subjective probability of obtaining delayed rewards. *Behavioral and Brain Functions, 3,* 52. http://dx.doi.org/10.1186/1744-9081-3-52

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

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5,* 297–323. http://dx.doi.org/10.1007/BF00122574

Vandekerckhove, J., Matzke, D., & Wagenmakers, E. J. (2015). Model comparison and the principle of parsimony. In J. R. Busemeyer, J. Townsend, Z. J. Wang, & A. Aidels (Eds.), *Oxford handbook of computational and mathematical psychology* (pp. 300–319). Oxford, UK: Oxford University Press.

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. http://dx.doi.org/10.1037/xlm0000029

Vanderveldt, A., Green, L., & Rachlin, H. (2017). Discounting by probabilistic waiting. *Journal of Behavioral Decision Making, 30,* 39–53.

Vincent, B. T. (2016). Hierarchical Bayesian estimation and hypothesis testing for delay discounting tasks. *Behavior Research Methods, 48,* 1608–1620. http://dx.doi.org/10.3758/s13428-015-0672-2

Vlaev, I., Chater, N., & Stewart, N. (2007). Financial prospect relativity: Context effects in financial decision-making under risk. *Journal of Behavioral Decision Making, 20,* 273–304.

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

Weber, B. J., & Chapman, G. B. (2005a). 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. http://dx.doi.org/10.1016/j.obhdp.2005.01.001

Weber, B. J., & Chapman, G. B. (2005b). Playing for peanuts: Why is risk seeking more common for low-stakes gambles? *Organizational Behavior and Human Decision Processes, 97,* 31–46. http://dx.doi.org/10.1016/j.obhdp.2005.03.001

Wulff, D. U., & van den Bos, W. (2018). Modeling choices in delay discounting. *Psychological Science, 29,* 1890–1894.

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

(*Appendices follow*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix A

Model and Prior Distribution Specification

Model	Function	Participant parameter	Hyperparameters
PTHD	$w(p) = e^{-(-\ln p)^r}$	$r \sim N(\mu_r, \sigma_r)T(0, 1)$	$\mu_r \sim U(0, 1)$
			$\sigma_r \sim U(0, 1)$
	$d(t) = \frac{1}{1 + ht}$	$h \sim N(\mu_h, \sigma_h)T(0, \infty)$	$\mu_h \sim U(0, 5)$
			$\sigma_h \sim U(0, 5)$
MHD	$d(t) = \frac{1}{(1 + h_d t)^{s_d}}$	$h_d \sim N(\mu_{hd}, \sigma_{hd})T(0, \infty)$	$\mu_{hd} \sim U(0, 5)$
			$\sigma_{hd} \sim U(0, 5)$
		$s_d \sim N(\mu_{sd}, \sigma_{sd})T(0, \infty)$	$\mu_{sd} \sim U(0, 10)$
			$\sigma_{sd} \sim U(0, 10)$
	$w(p, x) = \frac{1}{(1 + h_r(1/p - 1))^{s_r x^c}}$	$h_r \sim N(\mu_{hr}, \sigma_{hr})T(0, \infty)$	$\mu_{hr} \sim U(0, 360)$
			$\sigma_{hr} \sim U(0, 360)$
		$s_r \sim N(\mu_{sr}, \sigma_{sr})T(0, \infty)$	$\mu_{sr} \sim U(0, 10)$
			$\sigma_{sr} \sim U(0, 10)$
		$c \sim N(\mu_c, \sigma_c)T(0, \infty)$	$\mu_c \sim U(0, 10)$
			$\sigma_c \sim U(0, 10)$
HD	$d(t, p) = \frac{1}{1 + h(t + i(1/p - 1))}$	$h \sim N(\mu_h, \sigma_h)T(0, \infty)$	$\mu_h \sim U(0, 5)$
			$\sigma_h \sim U(0, 5)$
		$i \sim N(\mu_i, \sigma_i)T(0, \infty)$	$\mu_i \sim U(0, 360)$
			$\sigma_i \sim U(0, 360)$
PTHD, MHD, and HD	$v(x) = x^a$	$a \sim N(\mu_a, \sigma_a)T(0, 100)$	$\mu_a \sim U(0, 10)$
			$\sigma_a \sim U(0, 10)$
PTT	$v(x) = \frac{1 - e^{-\alpha x^{1-\beta}}}{\alpha}$	$\alpha \sim N(\mu_\alpha, \sigma_\alpha)T(10^{-10}, \infty)$	$\mu_\alpha \sim U(10^{-10}, 0.2)$
			$\sigma_\alpha \sim U(0, 0.5)$
		$\beta \sim N(\mu_\beta, \sigma_\beta)T(-\infty, 1)$	$\mu_\beta \sim U(-1, 1)$
			$\sigma_\beta \sim U(0, 2)$
	$w(p, t, x) = e^{-(Rt/\lvert x\rvert - \ln p)^\gamma}$	$R \sim N(\mu_R, \sigma_R)T(0, \infty)$	$\mu_R \sim U(0, 150)$
			$\sigma_R \sim U(0, 150)$
		$\gamma \sim N(\mu_\gamma, \sigma_\gamma)T(0, 1)$	$\mu_\gamma \sim U(0, 1)$
			$\sigma_\gamma \sim U(0, 1)$
Utility	$P(g_1 \mid \{g_1, g_2\}) = \frac{1}{1 + e^{-s(U(g_1) - U(g_2))}}$	$s \sim N(\mu_s, \sigma_s)T(0, \infty)$	$\mu_s \sim U(0, 100)$
			$\sigma_s \sim U(0, 100)$
Trade-off	$v(x) = \frac{1}{\beta} \cdot \ln(1 + \beta x)$	$\beta \sim N(\mu_\beta, \sigma_\beta)T(0, \infty)$	$\mu_\beta \sim U(0, 20)$
			$\sigma_\beta \sim U(0, 20)$
	$w_p(p) = \frac{p^\gamma}{(p^\gamma + (1 - p)^\gamma)^{1/\gamma}}$	$\gamma \sim N(\mu_\gamma, \sigma_\gamma)T(0, 1)$	$\mu_\gamma \sim U(0, 1)$
			$\sigma_\gamma \sim U(0, 1)$
	$w_t(t) = \frac{1}{\tau} \cdot \ln(1 + \tau t)$	$\tau \sim N(\mu_\tau, \sigma_\tau)T(0, \infty)$	$\mu_\tau \sim U(0, 90)$
			$\sigma_\tau \sim U(0, 90)$
	$Q(t_S, t_L) = \frac{\kappa}{\alpha}\ln\left(1 + \alpha\left(\frac{w_t(t_L) - w_t(t_S)}{\vartheta}\right)^\vartheta\right)$	$\kappa \sim N(\mu_\kappa, \sigma_\kappa)T(0, \infty)$	$\mu_\kappa \sim U(0, 200)$
			$\sigma_\kappa \sim U(0, 200)$
		$\alpha \sim N(\mu_\alpha, \sigma_\alpha)T(0, \infty)$	$\mu_\alpha \sim U(0, 100)$
			$\sigma_\alpha \sim U(0, 100)$
		$\vartheta \sim N(\mu_\vartheta, \sigma_\vartheta)T(1, \infty)$	$\mu_\vartheta \sim U(1, 100)$
			$\sigma_\vartheta \sim U(0, 100)$
	$P(g_1 \mid \{g_1, g_2\}) = \frac{1}{1 + e^{-s(U(g_L) - U(g_S) - Q(t_S, t_L))}}$	$s \sim N(\mu_s, \sigma_s)T(0, \infty)$	$\mu_s \sim U(0, 100)$
			$\sigma_s \sim U(0, 100)$
RITCH	$X = \beta_{xO} \cdot sgn(x_1 - x_2) + \beta_{xA}(x_1 - x_2) + \beta_{xR}\frac{x_1 - x_2}{x_m}$	$\beta_{xO} \sim N(\mu_{\beta xO}, \sigma_{\beta xO})T(0, \infty)$	$\mu_{\beta xO} \sim U(0, 100)$
			$\sigma_{\beta xO} \sim U(0, 100)$
		$\beta_{xA} \sim N(\mu_{\beta xA}, \sigma_{\beta xA})T(0, \infty)$	$\mu_{\beta xA} \sim U(0, 100)$
			$\sigma_{\beta xA} \sim U(0, 100)$
		$\beta_{xR} \sim N(\mu_{\beta xR}, \sigma_{\beta xR})T(0, \infty)$	$\mu_{\beta xR} \sim U(0, 100)$
			$\sigma_{\beta xR} \sim U(0, 100)$
	$T = \beta_{tO} \cdot sgn(t_2 - t_1) + \beta_{tA}(t_2 - t_1) + \beta_{tR}\frac{t_2 - t_1}{t_m}$	$\beta_{tO} \sim N(\mu_{\beta tO}, \sigma_{\beta tO})T(0, \infty)$	$\mu_{\beta tO} \sim U(0, 100)$
			$\sigma_{\beta tO} \sim U(0, 100)$

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix A (*continued*)

Model	Function	Participant parameter	Hyperparameters
PD	$R = \beta_{pO} \cdot sgn(p_1 - p_2) + \beta_{pA}(p_1 - p_2) + \beta_{pR}\frac{p_1 - p_2}{p_m}$	$\beta_{tA} \sim N(\mu_{\beta tA}, \sigma_{\beta tA})T(0, \infty)$	$\mu_{\beta tA} \sim U(0, 100)$
			$\sigma_{\beta tA} \sim U(0, 100)$
		$\beta_{tR} \sim N(\mu_{\beta tR}, \sigma_{\beta tR})T(0, \infty)$	$\mu_{\beta tR} \sim U(0, 100)$
			$\sigma_{\beta tR} \sim U(0, 100)$
		$\beta_{pO} \sim N(\mu_{\beta pO}, \sigma_{\beta pO})T(0, \infty)$	$\mu_{\beta pO} \sim U(0, 100)$
			$\sigma_{\beta pO} \sim U(0, 100)$
		$\beta_{pA} \sim N(\mu_{\beta pA}, \sigma_{\beta pA})T(0, \infty)$	$\mu_{\beta pA} \sim U(0, 100)$
			$\sigma_{\beta pA} \sim U(0, 100)$
		$\beta_{pR} \sim N(\mu_{\beta pR}, \sigma_{\beta pR})T(0, \infty)$	$\mu_{\beta pR} \sim U(0, 100)$
			$\sigma_{\beta pR} \sim U(0, 100)$
	$X = \beta_{xO} \cdot sgn(x_1 - x_2) + \beta_{xR}\frac{x_1 - x_2}{\max(x_1, x_2)}$	$\beta_{xO} \sim N(\mu_{\beta xO}, \sigma_{\beta xO})T(0, \infty)$	$\mu_{\beta xO} \sim U(0, 100)$
			$\sigma_{\beta xO} \sim U(0, 100)$
		$\beta_{xR} \sim N(\mu_{\beta xR}, \sigma_{\beta xR})T(0, \infty)$	$\mu_{\beta xR} \sim U(0, 100)$
			$\sigma_{\beta xR} \sim U(0, 100)$
	$T = \beta_{tO} \cdot sgn(t_2 - t_1) + \beta_{tR}\frac{t_2 - t_1}{\max(t_1, t_2)}$	$\beta_{tO} \sim N(\mu_{\beta tO}, \sigma_{\beta tO})T(0, \infty)$	$\mu_{\beta tO} \sim U(0, 100)$
			$\sigma_{\beta tO} \sim U(0, 100)$
		$\beta_{tR} \sim N(\mu_{\beta tR}, \sigma_{\beta tR})T(0, \infty)$	$\mu_{\beta tR} \sim U(0, 100)$
			$\sigma_{\beta tR} \sim U(0, 100)$
	$R = \beta_{pO} \cdot sgn(p_1 - p_2) + \beta_{pR}\frac{p_1 - p_2}{\max(p_1, p_2)}$	$\beta_{pO} \sim N(\mu_{\beta pO}, \sigma_{\beta pO})T(0, \infty)$	$\mu_{\beta pR} \sim U(0, 100)$
			$\sigma_{\beta pR} \sim U(0, 100)$
		$\beta_{pR} \sim N(\mu_{\beta pR}, \sigma_{\beta pR})T(0, \infty)$	$\mu_{\beta pO} \sim U(0, 100)$
			$\sigma_{\beta pO} \sim U(0, 100)$
RITCH and PD	$P(g_1 \mid g_1, g_2) = \frac{1}{1 + e^{-(X + T + R)}}$		

Note. $N(\mu, \sigma)T(a, b)$ is a truncated normal distribution with mean μ, standard deviation σ, lowerbound a and upperbound b. $U(a, b)$ is a uniform distribution with lowerbound a, and upperbound b. PTHD = prospect theory and hyperbolic discounting; MHD = multiplicative hyperboloid discounting; HD = hyperbolic discounting; PTT = probability-time-trade-off; RITCH = risky intertemporal choice heuristic; PD = proportional difference.

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix B

Parameter Interpretations

Function	Model	Equation	Range	Interpretation of parameter
Value	PTHD MHD HD	$v(x) = x^a$	$a \geq 0$	a measures how sensitive participants are to changes in amount, as a function of amount. If $0 < a < 1$ then the value function is concave, with participants showing diminishing sensitivity to amounts, that is as the amount of money increases the effect that a unit increase in money has is diminished, e.g. the difference between \$1 and \$2 is much greater than the difference between \$101 and \$102. The lower a is the greater the diminishing sensitivity. If $a = 1$, participants display constant sensitivity, i.e. objective amounts and subjective values are linearly related with a \$1 increase in amount having the same effect on value across the range of amounts. If $a > 1$, the participant exhibits increasing sensitivity/a convex relationship, with subjective value changing by more when amounts are high compared with low. These nonlinear sensitivities to amount could be interpreted as due to the subjective perception of amount, or due to considerations of utility. In risky choice $a < 1$ is often interpreted as indicating risk aversion, while $a > 1$, indicates risk seeking, but this mapping is not valid for risky intertemporal choices.
	Trade-off	$v(x) = \frac{1}{\beta} \cdot \ln(1 + \beta x)$	$\beta > 0$	β also measures sensitivity to changes in amount, however it can only measure changes in the degree of diminishing sensitivity, not capture increasing sensitivity. For $\beta > 0$, participants exhibit diminishing sensitivity, with higher values of β indicating greater diminishing sensitivity. As β goes to 0 constant sensitivity is observed (linear relationship), while as β goes to infinity insensitivity is observed.
	PTT	$v(x) = \frac{1 - e^{-\alpha x^{1-\beta}}}{\alpha}$	$\beta \leq 1$	α and β both control sensitivity to amount, similar to a in the value function above. When α approaches 0 (i.e. the exponential component reduces to a linear function) then $1 - \beta$ has equivalent interpretation to a in the power function above. $0 < \beta < 1$ produces diminishing sensitivity to amount, with diminishing sensitivity increasing as β increases. $\beta < 0$ produces increasing sensitivity, $\beta = 0$ produces linear sensitivity and $\beta = 1$ produces insensitivity to amounts. If α is not approaching 0, then when $\beta < 0$ participants show increasing sensitivity for small amounts which shifts to diminishing sensitivity as amounts increase.

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix B (*continued*)

Function	Model	Equation	Range	Interpretation of parameter
			$\alpha > 0$	α effects the degree of diminishing sensitivity, and also the elasticity of the function (which is important for explaining the magnitude effect). When $\beta = 0$ (i.e. the power component is a linear function), α alone controls the concavity of the function, with stronger diminishing sensitivity the higher α is. When α approaches infinity, participants become insensitive to amounts. When α approaches 0 the value function shifts from being decreasingly elastic to constant elasticity (i.e. it reduces to the power function), and will become linear if $\beta = 0$.
Probability weighting	PTHD	$w(p) = e^{-(-\ln p)^r}$	$0 \leq r \leq 1$	γ/r measures γ/r subproportionality or the extent to which participants overweight small probability events and underweight moderate to large probability events. Small γ/r values mean participants greatly overweight small probabilities and underweight moderate to large probabilities, while values close to 1 indicate little under/overweighting. $\gamma/r = 1$ means no over/underweighting occurs, and decision-weights are identical to objective probabilities. Values close to 0 mean all probabilities are given equal weight.
	Trade-off	$w_p(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}}$	$0 < \gamma \leq 1$	
	PTT	$w(p, t, x) = e^{-\langle Rt/\lvert x\rvert - \ln p\rangle_\gamma}$	$0 \leq \gamma \leq 1$	
			$R \geq 0$	Discount rate parameter. High values indicate that delayed rewards lose a large amount of their worth per month they are delayed, while low values indicate they lose little of their value. If $\gamma = 1$ then $R/\lvert x\rvert$ is an exponential discount rate/compound interest rate, with the rate of discounting decreasing as the amount increases.
	MHD	$w(p, x) = \frac{1}{(1 + h_r(1/p - 1))^{s_r x^c}}$	$h_r \geq 0$	Can be interpreted in two ways. Firstly, it can be related to under and overweighting of probabilities. When $s_r x^c = 1$, low values of h_r indicate participants overweight all probabilities, with this overweighting reducing and shifting to underweighting as h_r increases. When $s_r x^c < 1$, then the probability weighting function is s-shaped, with h_r still corresponding to increasing/decreasing overweighting, but also controlling the inflection point, i.e. the probability where participants go from underweighting to overweighting probabilities. As h_r increases the inflection point decreases (i.e. more moderate probabilities are underweighted). Alternatively, it can be interpreted as a measure of discounting based on the inferred delay until receipt (see h below).

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix B (*continued*)

Function	Model	Equation	Range	Interpretation of parameter
			$s_r \geq 0$	Measures sensitivity to risk, similar to a for amounts in the value function above. $0 < s_r < 1$ (the expected range) indicates diminishing sensitivity to risk/odds against receipt, with participants more sensitive to changes in risk when risk is low (i.e. probability is high) compared with when risk is high (i.e. probability low). $s_r > 1$ indicates increasing sensitivity to risk, with participants more sensitive to changes in odds against receipt when risk is high, than when risk is low. $s_r = 0$ indicates insensitivity and $s_r = 1$ constant sensitivity to odds against receipt.
			$c \geq 0$	The extent to which sensitivity to risk changes as function of amount. $c = 0$ means amount has no effect on sensitivity to risk ($s_r x^c = s_r$). $c > 0$ means that sensitivity to risk increases as the amount offered increases, with participants more sensitive to risk for large amounts than for small.
Discount	HD	$d(t, p) = \frac{1}{1 + h(t + i(1/p - 1))}$	$i \geq 0$	Scale parameter for converting odds against receipt of a reward into expected delay until that reward. Can be interpreted as the number of months participants expect they would need to wait between each opportunity to play the gamble.
	PTHD	$d(t) = \frac{1}{1 + ht}$	$h \geq 0$	h measures the discount rate. As h increases participants discount rewards more steeply as a function of time, i.e. if h is small then a reward retains more of its value per month it is delayed than if h is large. If $h = 0$ then outcomes are not discounted, i.e. $\$x$ now is the same as $\$x$ in the future. h can also be interpreted as the simple (as opposed to compound) interest rate participants apply when calculating the future value of an immediate reward.
	MHD	$d(t) = \frac{1}{(1 + h_d t)^{s_d}}$	$h_d \geq 0$	
			$s_d \geq 0$	Measures sensitivity to delay, similar to s_r for risk above. $0 < s_d < 1$ (expected range) indicates diminishing sensitivity to time/delay, with participants more sensitive to changes in delay when delays are small, compared with when delays are long. $s_d > 1$ indicates increasing sensitivity to delay, with participants more sensitive to changes in delay when delays are long, than when delays are short. $s_d = 0$ indicates insensitivity and $s_d = 1$ constant sensitivity to delay.
Time weighting	Trade-off	$w_t(t) = \frac{1}{\tau} \cdot \ln(1 + \tau t)$	$\tau > 0$	τ measures diminishing sensitivity to delays. Interpretation is the same as the β from the trade-off value function, but for delay.

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix B (*continued*)

Function	Model	Equation	Range	Interpretation of parameter
Time weighting	Trade-off	$Q(t_S, t_L) = \frac{\kappa}{\alpha}\ln\left(1 + \alpha\left(\frac{w_t(t_L) - w_t(t_S)}{\vartheta}\right)^{\vartheta}\right)$	$\kappa \geq 0$	Represents time sensitivity (relative to amount/utility). As κ increases participants become more sensitive to delays/less sensitive to utilities.
			$\alpha > 0$	Measures subadditivity. Additivity is a feature of many discounting models, whereby it is assumed that discounting over a time period (e.g. 1 month) should be equivalent regardless of whether it is assessed as a single interval, or broken into subintervals when assessing discounting (e.g. assess discounting for each of the 4 weeks of the month). Subadditivity is a violation of additivity whereby discounting is greater when assessed over the subintervals (e.g. 4 separate weeks) than when assessed over the whole interval (1 month). In the limit as α approaches 0 subadditivity disappears.
			$\vartheta \geq 1$	Measures superadditivity. The reverse of subadditivity, whereby discounting is greater when assessed over the single interval, relative to the subintervals. The model predicts a shift from superadditivity to subadditivity as the interval length increases. If $\vartheta = 1$ superadditivity disappears.
Choice	PTHD MHD HD PTT Trade-off	$P(g_1 \mid \{g_1, g_2\}) = \frac{1}{1 + e^{-s(U(g_1) - U(g_2))}}$ $P(g_L \mid g_L, g_S) = \frac{1}{1 + e^{-s(U(g_L) - U(g_S) - Q(t_S, t_L))}}$	$s \geq 0$	Choice scaling parameter or sensitivity parameter, it is inverse to the noise in the participants decisions. Determines how deterministic a participant's preferences are based on the utility differences they observe. High values indicate very strong preferences, with preferences becoming binary/deterministic as s approaches infinity and preferences becoming weak/insensitive to utility differences as s approaches 0.

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix B (*continued*)

Function	Model	Equation	Range	Interpretation of parameter
Amount difference	RITCH	$X = \beta_{xO} \cdot \mathrm{sgn}(x_1 - x_2) + \beta_{xA}(x_1 - x_2) + \beta_{xR}\frac{x_1 - x_2}{x_m}$	$\beta_{xA} \geq 0$	Measures the relative weight given to a $1 difference in amount between the two options. The higher this value the more influence a difference in amount will have on the participants preference/the more importance that participant gives to the absolute outcomes. Can be compared with β_{tA} and β_{pA} to see how much importance is given to a $1 difference in amount relative to a 1-month difference in delay relative to a 1% difference in probability. However, these comparisons should be made with an awareness of β_{xR}, β_{tR}, and β_{pR}, as differences in amount may still have a very large impact on a participants' preferences, even if β_{xA} is relatively low, if β_{xR} (i.e. the weight given to proportional differences in amount) is high enough.
	PD	$X = \beta_{xO} \cdot \mathrm{sgn}(x_1 - x_2) + \beta_{xR}\frac{x_1 - x_2}{\max(x_1, x_2)}$	$\beta_{xR} \geq 0$	Measures the relative weight given to a proportional change in amount. Higher values indicate more importance is being given to proportional differences in amount, but is only interpretable as a relative measure (i.e. when compared to other weight parameters) not in absolute terms.
			$\beta_{xO} \geq 0$	Bias towards preferring the larger amount. Higher values indicate a greater preference for the larger amount, regardless of the size of the amount difference. Can be compared with β_{tO} and β_{pO} to see strength of bias for each of the dimensions.
Delay difference	RITCH	$T = \beta_{tO} \cdot \mathrm{sgn}(t_2 - t_1) + \beta_{tA}(t_2 - t_1) + \beta_{tR}\frac{t_2 - t_1}{t_m}$	$\beta_{tA} \geq 0$	Measures the relative weight given to a 1 month difference in delay between the 2 options. The higher this value the more influence a difference in delay will have on the participants preference/the more importance that participant gives to absolute delays. See β_{xA} for information on interpretation.
	PD	$T = \beta_{tO} \cdot \mathrm{sgn}(t_2 - t_1) + \beta_{tR}\frac{t_2 - t_1}{\max(t_1, t_2)}$	$\beta_{tR} \geq 0$	Measures the relative weight given to a proportional change in delay. Higher values indicate more importance is being given to proportional differences in delay, but is only interpretable as a relative measure (i.e. compared with other weight parameters) not in absolute terms.
			$\beta_{tO} \geq 0$	Bias towards preferring the sooner outcome. Higher values indicate a greater preference for the sooner outcome, regardless of the size of the delay difference. Can be compared with β_{pO} and β_{xO}.

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

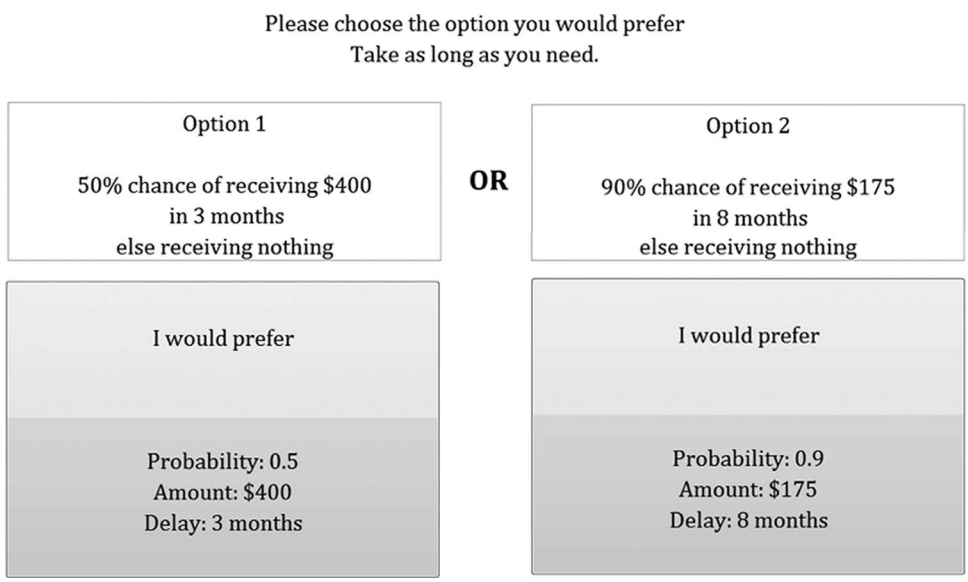
Appendix B (*continued*)

Function	Model	Equation	Range	Interpretation of parameter
Probability difference	RITCH	$R = \beta_{pO} \cdot \mathrm{sgn}(p_1 - p_2) + \beta_{pA}(p_1 - p_2) + \beta_{pR}\frac{p_1 - p_2}{p_m}$	$\beta_{pA} \geq 0$	Measures the relative weight given to a 1% difference in probability between the two options. The higher this value the more influence a difference in probability will have on the participants preference/the more importance that participant gives to absolute probabilities. See β_{xA} for information on interpretation.
	PD	$R = \beta_{pO} \cdot \mathrm{sgn}(p_1 - p_2) + \beta_{pR}\frac{p_1 - p_2}{\max(p_1, p_2)}$	$\beta_{pR} \geq 0$	Measures the relative weight given to a proportional change in probability. Higher values indicate more importance is being given to proportional differences in probability, but is only interpretable as a relative measure (i.e. compared with other weight parameters) not in absolute terms.
			$\beta_{pO} \geq 0$	Bias towards preferring the safer outcome. Higher values indicate a greater preference for the safer outcome, regardless of the size of the probability difference. Can be compared with β_{tO} and β_{xO}.

Note. PTHD = prospect theory and hyperbolic discounting; MHD = multiplicative hyperboloid discounting; HD = hyperbolic discounting; PTT = probability-time-trade-off; RITCH = risky intertemporal choice heuristic; PD = proportional difference. Parameters with similar meaning are grouped together. Allowed ranges of all parameters are specified.

Appendix C

Screenshot of Trial



(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix D

Stimuli and Choice Proportions

Type	Option 1	Option 2	Base.	Mag.	Imm.	Cert.
Risk vs. Amount	\$75 in 3 months with $p = .5$	\$100 in 3 months with $p = .45$	44 (0.49)	40 (0.44)	51 (0.57)	27 (0.3)
	\$250 in 8 months with $p = .5$	\$325 in 8 months with $p = .4$	52 (0.58)	47 (0.52)	54 (0.6)	38 (0.42)
	\$300 in 0 months with $p = .5$	\$475 in 0 months with $p = .4$	53 (0.59)	48 (0.53)	52 (0.58)	34 (0.38)
	\$125 in 9 months with $p = .45$	\$475 in 9 months with $p = .2$	54 (0.6)	66 (0.73)	62 (0.69)	33 (0.37)
	\$150 in 0 months with $p = .9$	\$400 in 0 months with $p = .6$	55 (0.61)	65 (0.72)	53 (0.59)	26 (0.29)
	\$150 in 0 months with $p = .85$	\$475 in 0 months with $p = .5$	56 (0.62)	55 (0.61)	55 (0.61)	29 (0.32)
	\$125 in 0 months with $p = .6$	\$250 in 0 months with $p = .45$	56 (0.62)	55 (0.61)	56 (0.62)	39 (0.43)
	\$100 in 1 months with $p = .95$	\$200 in 1 months with $p = .75$	56 (0.62)	67 (0.74)	57 (0.63)	37 (0.41)
	\$325 in 5 months with $p = .55$	\$350 in 5 months with $p = .5$	56 (0.62)	53 (0.59)	52 (0.58)	40 (0.44)
	\$50 in 4 months with $p = .25$	\$325 in 4 months with $p = .05$	57 (0.63)	69 (0.77)	51 (0.57)	41 (0.46)
	\$100 in 0 months with $p = .5$	\$425 in 0 months with $p = .2$	60 (0.67)	59 (0.66)	57 (0.63)	53 (0.59)
	\$50 in 6 months with $p = .5$	\$375 in 6 months with $p = .1$	62 (0.69)	60 (0.67)	57 (0.63)	54 (0.6)
	\$275 in 0 months with $p = .7$	\$400 in 0 months with $p = .55$	64 (0.71)	62 (0.69)	64 (0.71)	28 (0.31)
	\$50 in 0 months with $p = .6$	\$275 in 0 months with $p = .2$	64 (0.71)	60 (0.67)	59 (0.66)	48 (0.53)
	\$325 in 2 months with $p = .95$	\$400 in 2 months with $p = .85$	66 (0.73)	60 (0.67)	65 (0.72)	37 (0.41)
	\$200 in 0 months with $p = .95$	\$400 in 0 months with $p = .75$	67 (0.74)	56 (0.62)	44 (0.49)	29 (0.32)
Delay vs. Amount	\$100 in 1 months with $p = 1$	\$200 in 5 months with $p = 1$	17 (0.19)	15 (0.17)	10 (0.11)	34 (0.38)
	\$325 in 2 months with $p = 1$	\$475 in 9 months with $p = 1$	23 (0.26)	11 (0.12)	14 (0.16)	30 (0.33)
	\$100 in 3 months with $p = 1$	\$225 in 12 months with $p = 1$	25 (0.28)	16 (0.18)	18 (0.2)	38 (0.42)
	\$125 in 3 months with $p = 1$	\$325 in 21 months with $p = 1$	27 (0.3)	11 (0.12)	20 (0.22)	30 (0.33)
	\$175 in 4 months with $p = 1$	\$350 in 15 months with $p = 1$	29 (0.32)	19 (0.21)	21 (0.23)	32 (0.36)
	\$100 in 1 months with $p = .9$	\$325 in 12 months with $p = .9$	29 (0.32)	17 (0.19)	14 (0.16)	25 (0.28)
	\$75 in 1 months with $p = 1$	\$300 in 30 months with $p = 1$	31 (0.34)	22 (0.24)	23 (0.26)	32 (0.36)
	\$350 in 30 months with $p = 1$	\$425 in 42 months with $p = 1$	31 (0.34)	29 (0.32)	33 (0.37)	41 (0.46)
	\$175 in 12 months with $p = .7$	\$350 in 30 months with $p = .7$	32 (0.36)	22 (0.24)	29 (0.32)	34 (0.38)
	\$125 in 6 months with $p = 1$	\$325 in 30 months with $p = 1$	33 (0.37)	23 (0.26)	34 (0.38)	44 (0.49)
	\$300 in 13 months with $p = .8$	\$425 in 30 months with $p = .8$	36 (0.4)	37 (0.41)	38 (0.42)	40 (0.44)
	\$75 in 4 months with $p = .65$	\$225 in 30 months with $p = .65$	36 (0.4)	24 (0.27)	33 (0.37)	33 (0.37)
	\$350 in 0 months with $p = .75$	\$450 in 4 months with $p = .75$	37 (0.41)	26 (0.29)	18 (0.2)	36 (0.4)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix D (*continued*)

Type	Option 1	Option 2	Base.	Mag.	Imm.	Cert.
	\$75 in 15 months with $p = .6$	\$175 in 36 months with $p = .6$	37 (0.41)	27 (0.3)	31 (0.34)	38 (0.42)
	\$150 in 6 months with $p = .85$	\$450 in 27 months with $p = .85$	38 (0.42)	25 (0.28)	30 (0.33)	28 (0.31)
	\$125 in 0 months with $p = .95$	\$350 in 24 months with $p = .95$	40 (0.44)	24 (0.27)	32 (0.36)	45 (0.5)
Delay vs. Risk	\$250 in 3 months with $p = .45$	\$250 in 6 months with $p = .85$	7 (0.08)	7 (0.08)	7 (0.08)	15 (0.17)
	\$350 in 4 months with $p = .5$	\$350 in 8 months with $p = .85$	9 (0.1)	7 (0.08)	8 (0.09)	20 (0.22)
	\$150 in 5 months with $p = .75$	\$150 in 9 months with $p = .95$	10 (0.11)	10 (0.11)	6 (0.07)	24 (0.27)
	\$225 in 2 months with $p = .6$	\$225 in 6 months with $p = .9$	10 (0.11)	9 (0.1)	8 (0.09)	22 (0.24)
	\$150 in 6 months with $p = .2$	\$150 in 36 months with $p = .95$	17 (0.19)	9 (0.1)	15 (0.17)	28 (0.31)
	\$50 in 1 months with $p = .55$	\$50 in 5 months with $p = .9$	20 (0.22)	12 (0.13)	9 (0.1)	37 (0.41)
	\$100 in 0 months with $p = .7$	\$100 in 4 months with $p = 1$	22 (0.24)	12 (0.13)	10 (0.11)	28 (0.31)
	\$400 in 0 months with $p = .8$	\$400 in 8 months with $p = 1$	25 (0.28)	14 (0.16)	14 (0.16)	42 (0.47)
	\$300 in 9 months with $p = .4$	\$300 in 15 months with $p = .5$	25 (0.28)	19 (0.21)	34 (0.38)	36 (0.4)
	\$400 in 0 months with $p = .75$	\$400 in 15 months with $p = 1$	27 (0.3)	21 (0.23)	31 (0.34)	44 (0.49)
	\$75 in 0 months with $p = .1$	\$75 in 42 months with $p = 1$	29 (0.32)	15 (0.17)	24 (0.27)	24 (0.27)
	\$450 in 0 months with $p = .7$	\$450 in 21 months with $p = 1$	31 (0.34)	16 (0.18)	23 (0.26)	43 (0.48)
	\$225 in 0 months with $p = .85$	\$225 in 12 months with $p = 1$	37 (0.41)	23 (0.26)	35 (0.39)	46 (0.51)
	\$250 in 9 months with $p = .35$	\$250 in 18 months with $p = .4$	37 (0.41)	31 (0.34)	39 (0.43)	59 (0.66)
	\$50 in 0 months with $p = .85$	\$50 in 9 months with $p = 1$	39 (0.43)	29 (0.32)	32 (0.36)	51 (0.57)
	\$125 in 0 months with $p = .95$	\$125 in 5 months with $p = 1$	40 (0.44)	39 (0.43)	36 (0.4)	58 (0.64)
Risk vs. Amount & Delay	\$225 in 4 months with $p = 1$	\$450 in 0 months with $p = .7$	46 (0.51)	57 (0.63)	55 (0.61)	30 (0.33)
	\$125 in 12 months with $p = .9$	\$150 in 2 months with $p = .7$	46 (0.51)	63 (0.7)	65 (0.72)	35 (0.39)
	\$150 in 24 months with $p = .75$	\$325 in 21 months with $p = .55$	47 (0.52)	50 (0.56)	51 (0.57)	30 (0.33)
	\$50 in 3 months with $p = 1$	\$325 in 0 months with $p = .25$	49 (0.54)	65 (0.72)	55 (0.61)	44 (0.49)
	\$150 in 9 months with $p = 1$	\$325 in 0 months with $p = .5$	50 (0.56)	65 (0.72)	61 (0.68)	34 (0.38)
	\$50 in 6 months with $p = 1$	\$225 in 0 months with $p = .25$	54 (0.6)	65 (0.72)	58 (0.64)	52 (0.58)
	\$375 in 4 months with $p = 1$	\$475 in 0 months with $p = .75$	55 (0.61)	63 (0.7)	70 (0.78)	38 (0.42)
	\$175 in 1 months with $p = 1$	\$400 in 0 months with $p = .5$	57 (0.63)	67 (0.74)	66 (0.73)	29 (0.32)
	\$175 in 8 months with $p = .9$	\$400 in 3 months with $p = .5$	58 (0.64)	59 (0.66)	57 (0.63)	36 (0.4)
	\$175 in 2 months with $p = 1$	\$450 in 0 months with $p = .55$	63 (0.7)	70 (0.78)	64 (0.71)	44 (0.49)
	\$175 in 54 months with $p = .55$	\$450 in 24 months with $p = .05$	63 (0.7)	73 (0.81)	66 (0.73)	63 (0.7)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix D (*continued*)

Type	Option 1	Option 2	Base.	Mag.	Imm.	Cert.
	\$250 in 13 months with $p = .35$	\$325 in 12 months with $p = .2$	64 (0.71)	66 (0.73)	57 (0.63)	42 (0.47)
	\$200 in 9 months with $p = 1$	\$450 in 0 months with $p = .5$	65 (0.72)	63 (0.7)	58 (0.64)	32 (0.36)
	\$100 in 12 months with $p = .45$	\$400 in 5 months with $p = .1$	68 (0.76)	74 (0.82)	62 (0.69)	40 (0.44)
	\$100 in 21 months with $p = .6$	\$125 in 2 months with $p = .2$	72 (0.8)	79 (0.88)	63 (0.7)	68 (0.76)
	\$375 in 42 months with $p = .7$	\$450 in 36 months with $p = .45$	78 (0.87)	83 (0.92)	69 (0.77)	50 (0.56)
Delay vs. Amount & Risk	\$125 in 13 months with $p = .7$	\$150 in 15 months with $p = .75$	9 (0.1)	8 (0.09)	12 (0.13)	15 (0.17)
	\$325 in 1 months with $p = .5$	\$425 in 2 months with $p = .6$	10 (0.11)	7 (0.08)	3 (0.03)	20 (0.22)
	\$425 in 5 months with $p = .2$	\$450 in 24 months with $p = .65$	18 (0.2)	11 (0.12)	14 (0.16)	30 (0.33)
	\$50 in 30 months with $p = .4$	\$275 in 54 months with $p = .5$	18 (0.2)	8 (0.09)	18 (0.2)	23 (0.26)
	\$300 in 0 months with $p = .95$	\$450 in 2 months with $p = 1$	20 (0.22)	10 (0.11)	11 (0.12)	30 (0.33)
	\$100 in 5 months with $p = .35$	\$475 in 30 months with $p = .5$	20 (0.22)	9 (0.1)	18 (0.2)	19 (0.21)
	\$425 in 1 months with $p = .2$	\$450 in 15 months with $p = .4$	20 (0.22)	17 (0.19)	28 (0.31)	36 (0.4)
	\$75 in 0 months with $p = .9$	\$150 in 4 months with $p = 1$	21 (0.23)	10 (0.11)	11 (0.12)	25 (0.28)
	\$175 in 0 months with $p = .75$	\$250 in 21 months with $p = 1$	22 (0.24)	19 (0.21)	21 (0.23)	42 (0.47)
	\$375 in 0 months with $p = .75$	\$475 in 15 months with $p = 1$	22 (0.24)	17 (0.19)	14 (0.16)	32 (0.36)
	\$225 in 12 months with $p = .2$	\$250 in 30 months with $p = .5$	22 (0.24)	15 (0.17)	16 (0.18)	25 (0.28)
	\$275 in 0 months with $p = .85$	\$325 in 12 months with $p = 1$	25 (0.28)	17 (0.19)	22 (0.24)	39 (0.43)
	\$225 in 0 months with $p = .95$	\$450 in 15 months with $p = 1$	33 (0.37)	18 (0.2)	18 (0.2)	37 (0.41)
	\$425 in 0 months with $p = .8$	\$450 in 13 months with $p = 1$	34 (0.38)	19 (0.21)	29 (0.32)	41 (0.46)
	\$150 in 5 months with $p = .4$	\$350 in 36 months with $p = .45$	34 (0.38)	30 (0.33)	41 (0.46)	39 (0.43)
	\$75 in 0 months with $p = .95$	\$125 in 24 months with $p = 1$	35 (0.39)	24 (0.27)	36 (0.4)	50 (0.56)
Delay & Risk vs. Amount	\$175 in 3 months with $p = .9$	\$300 in 4 months with $p = .85$	22 (0.24)	26 (0.29)	33 (0.37)	22 (0.24)
	\$275 in 0 months with $p = 1$	\$475 in 4 months with $p = .95$	40 (0.44)	39 (0.43)	33 (0.37)	54 (0.6)
	\$125 in 0 months with $p = 1$	\$375 in 21 months with $p = .9$	57 (0.63)	55 (0.61)	54 (0.6)	55 (0.61)
	\$175 in 0 months with $p = 1$	\$300 in 5 months with $p = .8$	58 (0.64)	60 (0.67)	54 (0.6)	58 (0.64)
	\$175 in 8 months with $p = .55$	\$350 in 18 months with $p = .45$	59 (0.66)	48 (0.53)	64 (0.71)	48 (0.53)
	\$175 in 0 months with $p = 1$	\$375 in 2 months with $p = .75$	63 (0.7)	62 (0.69)	51 (0.57)	56 (0.62)
	\$125 in 9 months with $p = .6$	\$350 in 54 months with $p = .5$	63 (0.7)	64 (0.71)	61 (0.68)	59 (0.66)
	\$250 in 0 months with $p = 1$	\$375 in 5 months with $p = .85$	67 (0.74)	65 (0.72)	66 (0.73)	57 (0.63)
	\$150 in 1 months with $p = .95$	\$475 in 4 months with $p = .6$	68 (0.76)	67 (0.74)	67 (0.74)	46 (0.51)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix D (*continued*)

Type	Option 1	Option 2	Base.	Mag.	Imm.	Cert.
	\$75 in 18 months with $p = .95$	\$450 in 21 months with $p = .1$	74 (0.82)	82 (0.91)	68 (0.76)	58 (0.64)
	\$50 in 0 months with $p = 1$	\$225 in 21 months with $p = .45$	76 (0.84)	79 (0.88)	69 (0.77)	65 (0.72)
	\$50 in 0 months with $p = 1$	\$75 in 3 months with $p = .85$	76 (0.84)	77 (0.86)	58 (0.64)	69 (0.77)
	\$175 in 5 months with $p = .55$	\$225 in 8 months with $p = .3$	78 (0.87)	83 (0.92)	82 (0.91)	73 (0.81)
	\$100 in 0 months with $p = 1$	\$450 in 15 months with $p = .45$	80 (0.89)	81 (0.9)	83 (0.92)	65 (0.72)
	\$325 in 5 months with $p = .45$	\$375 in 8 months with $p = .35$	80 (0.89)	77 (0.86)	72 (0.8)	65 (0.72)
	\$200 in 3 months with $p = .6$	\$300 in 24 months with $p = .25$	86 (0.96)	87 (0.97)	85 (0.94)	80 (0.89)
Dominated options	\$50 in 0 months with $p = .4$	\$50 in 0 months with $p = .3$	87 (0.97)			
	\$50 in 18 months, with $p = 1$	\$50 in 24 months, with $p = 1$	85 (0.94)			
	\$400 in 6 months with $p = .7$	\$400 in 9 months with $p = .6$	90 (1)			
	\$300 in 3 months with $p = .8$	\$275 in 6 months with $p = .8$	87 (0.97)			
	\$300 in 9 months with $p = .85$	\$275 in 9 months with $p = .7$	86 (0.96)			
	\$225 in 18 months with $p = .4$	\$175 in 18 months with $p = .4$	83 (0.92)			

Note. Number (proportion) of participants choosing Option 1 in the baseline, magnitude, immediacy, and certainty datasets also supplied. Within each choice type individual choice items have been ordered so that the proportion choosing Option 1 in the baseline dataset increases. This matches the plot order in Figure 1.

Received February 22, 2018
Revision received May 1, 2020
Accepted May 4, 2020 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.