This postprint was originally published by Elsevier as:
Zilker, V., & Pachur, T. (2023). Attribute attention and option attention in risky choice. *Cognition*, *236*, Article 105441.
https://doi.org/10.1016/j.cognition.2023.105441

Attribute attention and option attention in risky choice

Veronika Zilker[a,b,*], Thorsten Pachur[a,b]

[a] *Technical University of Munich, School of Management, Chair of Behavioral Research Methods, Arcisstraße 21, 80333 Munich, Germany*

[b] *Max Planck Institute for Human Development, Center for Adaptive Rationality, Lentzeallee 94, 14195 Berlin, Germany*

***** Corresponding author at: Technical University of Munich, School of Management, Chair of Behavioral Research Methods, Arcisstraße 21, 80333 Munich, Germany.

E-mail address: veronika.zilker@tum.de (V. Zilker).

Keywords: Attention, Attentional bias, Cumulative prospect theory, Probability weighting

---

A B S T R A C T

Probability weighting is one of the most powerful theoretical constructs in descriptive models of risky choice and constitutes a central component of cumulative prospect theory (CPT). Probability weighting has been shown to be related to two facets of attention allocation: one analysis showed that differences in the shape of CPT's probability-weighting function are linked to differences in how attention is allocated across attributes (i.e., probabilities vs. outcomes); another analysis (that used a different measure of attention) showed a link between probability weighting and differences in how attention is allocated across options. However, the relationship between these two links is unclear. We investigate to what extent attribute attention and option attention independently contribute to probability weighting. Reanalyzing data from a process-tracing study, we first demonstrate links between probability weighting and both attribute attention and option attention within the same data set and the same measure of attention. We then find that attribute attention and option attention are at best weakly related and have independent and distinct effects on probability weighting. Moreover, deviations from linear weighting mainly emerged when attribute attention or option attention were imbalanced. Our analyses enrich the understanding of the cognitive underpinnings of preferences and illustrate that similar probability-weighting patterns can be associated with very different attentional policies. This complicates an unambiguous psychological interpretation of psycho-economic functions. Our findings indicate that cognitive process models of decision making should aim to concurrently account for the effects of different facets of attention allocation on preference. In addition, we argue that the origins of biases in attribute attention and option attention need to be better understood.

---

1. Introduction

Decisions under risk involve options that offer outcomes with some probability. Think of choosing to invest an amount of money in stocks versus bonds, or between medications with potential side effects. Cumulative prospect theory (CPT; Tversky & Kahneman, 1992) is one of the most influential theories of how people make such decisions. It accommodates several classical phenomena in decision making under risk—including the Allais paradox, the certainty effect, and the fourfold pattern of risk attitudes (Allais, 1953; Kahneman & Tversky, 1979; Tversky & Kahneman, 1992)—that indicate how people deviate from rational standards (e.g., expected value and expected utility theory). CPT has also proven useful to describe and predict choice behavior (e.g., Camerer, 2000; Glöckner & Pachur, 2012). CPT consists of two psycho-economic functions: a value function and a probability-weighting function. The value function transforms outcomes in a nonlinear fashion. The transformed outcomes are multiplied by decision weights. The decision weight of each outcome is calculated based on the outcome's transformed probability, and the transformation of the probabilities is governed by a probability-weighting function. This function is conventionally assumed to be inverse S-shaped, such that outcomes with small probabilities are overweighted and outcomes with medium to large probabilities are underweighted relative to their objective probabilities. The probability-weighting function has received particular attention in the literature, because it provides an elegant account of empirical violations of expected utility theory and is useful for describing individual differences (e.g., Bruhin, Fehr-Duda, & Epper, 2010; Cavagnaro, Pitt, Gonzalez, & Myung, 2013; Gonzalez & Wu, 1999; Prelec, 1998).

Some of CPT's theoretical elements have been interpreted psychologically. Tversky and Kahneman (1992) viewed the shapes of CPT's psycho-economic functions as indicating diminishing sensitivity: The further away outcomes or probabilities are from a reference point (e.g., zero), the less sensitive the decision maker is to differences between them. For the probability-weighting function, there are two reference points—impossibility and certainty. The commonly assumed inverse S-shape implies that sensitivity to differences in probability is highest near the end points of the probability scale and lowest in the medium range. Gonzalez and Wu (1999) differentiated this probability sensitivity (or discriminability), represented by the curvature parameter of the probability-weighting function, from the elevation of the probability-weighting function (represented by a separate parameter), which they interpreted as expressing how attractive the decision maker regards gambling. The authors also observed individual differences in these two properties of the probability-weighting function across decision makers (see also Cavagnaro et al., 2013).

While interpretations such as "probability sensitivity" characterize the shapes of CPT's functions in a psychological manner, they have little explanatory value. After all, CPT is not a cognitive process model (but see Pachur, Hertwig, Gigerenzer, & Brandstätter, 2013); the algorithmic operations of transforming outcomes and probabilities into subjective utilities and decision weights are not meant to represent the cognitive processes underlying people's decisions (cf. Friedman & Savage, 1948). Furthermore, the model does not explain which cognitive processes might give rise to the shapes of the psycho-economic functions.

Process-tracing studies, in which people's predecisional information search (i.e., information acquisition during preference formation that occurs before a final response is given) is recorded, can provide more insight about how differences in the shapes of probability-weighting functions come about on the level of cognitive processing. Pachur, Schulte-Mecklenbeck, Murphy, and Hertwig (2018) showed that the probability-weighting functions of participants who paid less attention to probability information during information search tended to be more strongly curved—indicating a lower sensitivity to probability information (see also Pachur, Hertwig, & Wolkewitz, 2014). However, imbalances in attribute attention—which we define here as the amount of attention allocated to probabilities relative to outcomes—are not the only possible attentional explanation of the shape of the probability-weighting function. In another study, we (Zilker & Pachur, 2022c) showed that imbalances in option attention—which we define here as the allocation of attention across the options in a choice problem—were also related to the shape of the probability-weighting function estimated from people's choices. For instance, in choices between two risky options, the more attention the decision maker paid to the riskier option (riskier in the sense that its higher possible outcome was less likely) relative to the alternative option during predecisional information search, the more curved the probability-weighting function estimated from people's choices was.

In the quest towards a better understanding of the cognitive underpinnings of probability weighting, empirical investigations have thus revealed two facets of attention that are systematically associated with the shape of the probability-weighting function. However, these two links between attention and probability weighting were demonstrated in different paradigms, established using different analytic approaches, and based on different measures of attention (i.e., inspection time in an information board task vs. sampling behavior in decisions from experience; see the sections "Probability Weighting and Attribute Attention" and "Probability Weighting and Option Attention" for details). It is therefore unclear how robust the links between these two dimensions of attention allocation and probability weighting are. More importantly, it is also unclear how they relate to each other.

In this article, we first establish that both attribute attention and option attention can be linked to probability weighting within the same data set and when relying on the same measure of attention, as well as using the same analytic techniques implemented within a common computational framework. This underscores the robustness of the two links between attention and probability weighting. Building on this novel insight, we then probe the relationship between the two links: Do they reflect a common underlying mechanism, or do the two facets of attention contribute independently to nonlinearities in probability weighting? In doing so, we build connections between previously isolated process-level correlates of nonlinear probability weighting in the literature.

Investigating these issues is important for at least two more reasons. First, if attribute attention and option attention contribute independently to nonlinear probability weighting, then a probability-weighting function with a given shape might result from different attentional mechanisms. Depending on which attentional mechanism underlies the distortion, different approaches to improving decision making—in the sense of improving adherence to linear probability weighting—might be called for. Second, if probability weighting is shaped by attribute attention and option attention independently, this would have important implications for theory development. In order to provide a comprehensive theoretical account, cognitive process models of risky choice need to accommodate the impact of both facets of attention. Specifically, they should account for imbalances in both attribute attention and option attention and explain how they concurrently influence preference formation. Ideally, they should also point to possible reasons for different types of attentional bias (e.g., why probabilities might receive less attention than do outcomes, or why one option might receive more attention than the other). Similar questions have recently been discussed and driven theory development in multiattribute choice (Yang & Krajbich, 2022); the current research lays the foundation for similar theoretical progress in the domain of risky choice.[1]

2. The link between probability weighting and attention

In order for a decision maker to build up a mental representation of a choice option—which forms the basis of their subjective evaluation—they need to encode the option's properties, or attributes. The interface between this information in the external environment and its internal representation is attention. Attention can be driven by both bottom-up and top-down influences (Mullett & Stewart, 2016; Orquin, Lahm, & Stojić, 2021; Vanunu, Hotaling, Le Pelley, & Newell, 2021). Because attention is a scarce resource (Kahneman, 1973; Simon, 1978), it may not be equally distributed across all available information during predecisional information search (for a review, see Orquin, Perkovic, & Grunert, 2018). For instance, some pieces of information may require more attention than others because they are more difficult to encode (e. g., odd vs. round numbers) or because of redundancies (e.g., probabilities sum up to one, so the probability of one outcome can be inferred from the other). Moreover, decision makers may pay only little attention to an attribute if they consider it relatively unimportant (e.g., Wedell & Senter, 1997). In addition to imbalances in attribute attention, there may be imbalances in how attention is distributed across options. The importance of imbalances in option attention for decision making is emphasized by the attentional drift diffusion model (aDDM; Krajbich, Armel, & Rangel, 2010) and extensions thereof (e.g., Gluth, Kern, Kortmann, & Vitali, 2020; Krajbich & Rangel, 2011), which have been developed in the context of value-based choice (e.g., between food items). The aDDM highlights that attention not only serves as an interface between the environment and its internal representations, but is also intimately tied to preference formation itself. Specifically, it posits that evidence for an option accumulates at a higher rate when the option is attended to than when it is not attended to. Imbalances in option attention can therefore affect a decision maker's evaluations of options and thus also their choice. Other models rely on similar mechanisms to accommodate effects of attribute attention on preferences (e.g., Fisher,

[1] In the Discussion section, we examine why the multiattribute choice model by Yang and Krajbich (2022)—which has also been applied to risky choice—and similar accounts may not fully resolve the challenges for theory development posed by our research.

2017; Glickman et al., 2019).

We next describe in more detail how both attribute attention and option attention have previously been found to be reflected in specific shapes of the probability-weighting function. This lays the foundation for our analyses that bridge the gap between the two previously disconnected phenomena.

2.1. Probability weighting and attribute attention

In the experiments by Pachur et al. (2018), participants' predecisional information search in a risky-choice task was monitored using the process-tracing tool Mouselab (e.g., Payne, Bettman, & Johnson, 1993; Willemsen & Johnson, 2011). In Mouselab, the information about options' outcomes and probabilities is hidden behind boxes. Before making a choice between options, participants can inspect the information about each option by moving the cursor to the boxes onscreen, making the underlying information visible for as long as the cursor is positioned over a box. The authors measured the amount of time participants spent attending to outcome and probability information by recording for how long the cursor was positioned over the respective box. Participants' choices were modeled using CPT, yielding parameter estimates for each participant. As shown in Fig. 1A, for participants who tended to spend less time attending to probabilities before making a choice, the estimated probability-weighting functions tended to be more strongly curved (i.e., lower values on the γ parameter; for details on the CPT parameters see the section “Computational Modeling”) than they were for participants who spent more time inspecting probability information. In addition, less attention to probabilities was linked to a higher elevation of the probability-weighting function (captured in the δ parameter) in the gain domain; there was no association between attribute attention and the elevation in the loss domain. The results of Pachur et al. (2018) indicate that individual differences in attribute-specific attention allocation, measured using Mouselab, are associated with individual differences in the shape of the probability-weighting function, estimated from people's choices (see also Harrison & Swarthout, 2019).

2.2. Probability weighting and option attention

In a previous study, we (Zilker & Pachur, 2022c) analyzed the link between option attention and probability weighting in two empirical data sets, including an analysis of information search and choice behavior in decisions from experience (using data on the sampling paradigm compiled by Wulff, Mergenthaler-Canseco, & Hertwig, 2018). In decisions from experience, no summary information about the possible outcomes and their probabilities is provided; people can learn about the available options by sampling from the options' payoff distributions as many times as they want before making a choice. In our analyses, the parameters of CPT's probability-weighting function, estimated from people's choices, were systematically associated with the direction and strength of biases in option attention during sampling. Specifically, in choice problems offering a safe and a risky option, a stronger attentional bias to the safe option—that is, taking more samples from the safe than from the risky option (irrespective of the probability of encountering each option's individual outcomes)—was associated with a more convex and less elevated probability-weighting function. In choice problems offering two risky options, paying less attention to the option that offered the more extreme outcome with a higher probability—the “more likely” option (in short, the ML option)—was associated with a more strongly inverse S-shaped probability-weighting function (i.e., lower values on the γ parameter). The latter results are displayed in Fig. 1B. Similar associations between option attention and the shape of the probability-weighting function were found in data from eye-tracking studies, in which option attention was measured based on fixation duration (Zilker & Pachur, 2022c). Overall, these analyses provide evidence that biases in attention allocation across options can be associated with systematic deviations from an objective, linear weighting of probabilities—such as the classical inverse S-shaped probability-weighting function.

3. Overview of the current analyses

In the following, we examine the relationship between the two facets of attention and their links to probability weighting. First, we investigate whether the link between attribute attention and probability weighting and the link between option attention and probability weighting can be demonstrated within the same data set, the same measure of attention, and using the same analytic procedures. This is important because the analyses discussed in the previous section (Pachur et al., 2018; Zilker & Pachur, 2022c) involved different paradigms and different measures of attention: One study measured attention with Mouselab; the other measured attention during information sampling in decisions from experience and fixation patterns using eye tracking in decisions from description. In addition, the conclusions of these two studies may depend on the analytic framework. Pachur et al. (2018) examined between-person variability in attention and probability weighting, whereas our (Zilker & Pachur, 2022c) previous analysis used an approach that accommodated both between- and within-person variability in attention and probability weighting. This approach is richer because it accounts not only for the effects of differences in attention allocation between individuals, but also for variability in attention across choice problems within the same individual; we therefore use the latter approach in the current analyses.

Second—and even more crucially—we investigate how the effects of the two attentional variables on probability weighting are related. Although the amount of attention allocated to probabilities relative to outcomes and the amount of attention allocated to one option relative to the other are conceptually distinct, they may be empirically related. In particular, in choice problems in which decision makers pay less attention to probabilities, they might also tend to pay less attention to the ML option. If probability attention and option attention are correlated, the findings in Pachur et al. (2018) and our (Zilker & Pachur, 2022c) previous study may not reflect distinct phenomena, but rather stem from a shared source of variance in attention.

4. Methods

4.1. Data

We re-analyzed data from Pachur et al.'s (2018) Experiment 1, a process-tracing study using Mouselab (Willemsen & Johnson, 2011). In this study, participants made choices between risky options.[2] In Mouselab, the attributes of an option—that is, the possible outcomes and their probabilities—are revealed only when the cursor is positioned over the box. The amount of time that each participant spent inspecting the individual boxes can be used as a measure of how attention was allocated to the different pieces of information before a choice was made. Fig. 2 displays an example of the screen setup on a given trial from Pachur et al. (2018). Whether the choice problems were shown in a horizontal or a vertical setup was counterbalanced across problems within participants, as was the order of outcome and probability information on the screen.

Pachur et al.'s (2018) experiment involved 60 pure-domain (35 pure gain, 25 pure loss) and 31 mixed-domain (in which options offered both

[2] We used Pachur et al.'s (2018) Mouselab data for our analyses rather than the sampling data from decisions from experience used in our (Zilker & Pachur, 2022c) previous analysis because it allowed us to measure both attribute attention and option attention. In sampling data from decisions from experience, probabilities are never explicitly displayed, thus making it impossible to measure attribute attention.

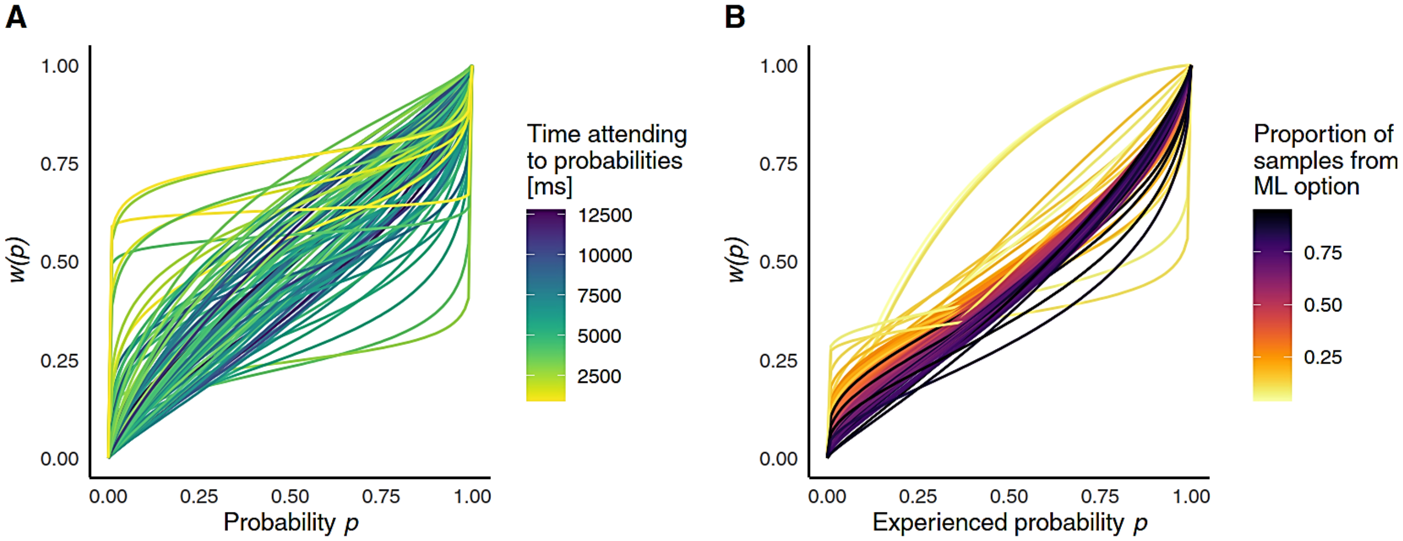

Fig. 1. Association of probability weighting with attribute attention (Panel A) and option attention (Panel B), as observed in Pachur et al. (2018) and Zilker and Pachur (2022c), respectively. ML option: option offering the more extreme outcome with a higher probability. Copyright © 2018 and © 2022 by the American Psychological Association. Adapted with permission. The official citations that should be used in referencing this material are Pachur et al. (2018) and Zilker and Pachur (2022c). The use of APA information does not imply endorsement by APA.

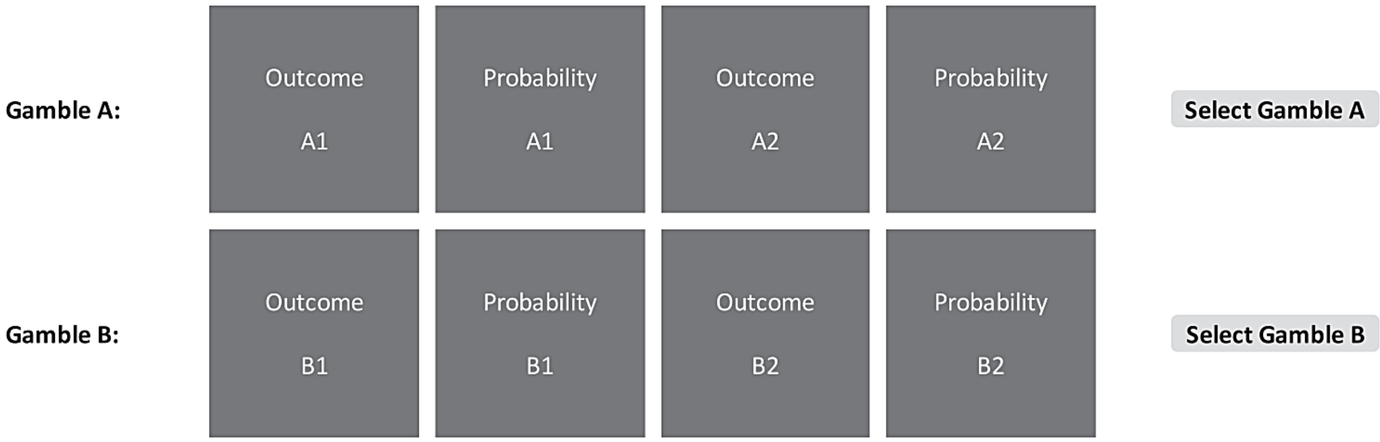

Fig. 2. Illustration of the horizontal setup of a choice problem in MouselabWEB in the experiment by Pachur et al. (2018). Each option consists of two outcomes and their probabilities. Copyright © 2018 by the American Psychological Association. Adapted with permission. The official citation that should be used in referencing this material is Pachur, Schulte-Mecklenbeck, Murphy, and Hertwig (2018). The use of APA information does not imply endorsement by APA.

positive and negative outcomes) choice problems. Our current analysis focuses on the pure-domain problems, because when deriving decision weights for problems in the mixed domain, CPT draws on a combination of the probability-weighting functions estimated for pure-gain domain and pure-loss domain problems. This may be problematic for the following reason. If the association between probability weighting and attention differs between problem types (as is suggested by the results from Pachur et al., 2018), the parameter estimates obtained when concurrently estimating CPT parameters for choices in pure-domain and mixed-domain problems would be difficult to interpret. All but one pure-domain choice problems involved two risky options, each consisting of two distinct outcomes with associated probabilities. The single pure-domain problem that offered a choice between two safe options was not included in the analysis.

In addition, our analysis included only those choice problems in which the probability of the more extreme (i.e., higher absolute) outcome differed between both options. Here is why. In our (Zilker & Pachur, 2022c) analysis of the link between option attention and probability weighting in choices between two risky options, we defined option attention in a given choice problem as the amount of attention paid to that option whose more extreme outcome was more likely (i.e., the ML option) relative to the other, less likely, option (i.e., the LL option). We relied on this distinction because CPT's probability-weighting function can modulate the relative attractiveness of the ML versus the LL option in a choice problem, allowing it to accommodate the impact of attentional biases on the relative attractiveness of options in risky choice. To illustrate the difference between ML and LL options, consider a choice between two risky options, A and B, where option A offers a 20% chance to win \$100 and an 80% chance to win \$50, and where option B offers a 40% chance to win \$80 and a 60% chance to win \$70. Here option B is the ML option, since its more extreme outcome (\$80) occurs with a higher probability (40%) than the more extreme outcome of option A (\$100), which occurs with a lower probability (20%). The distinction between ML and LL options cannot be applied in problems in which the probability of the more extreme outcome is exactly the same in both options. In the analyzed data, ten of the pure-gain choice problems did not have an ML option and were therefore not included in the analysis. Overall, our analysis included 50 distinct choice problems for each participant.

Each participant took part in two sessions, conducted three weeks apart. In both sessions, participants responded to the same choice problems, presented in a different random order. Our main analysis combines the data from both sessions. Appendix E presents additional analyses addressing the stability of results across the two sessions. For a typical participant, we therefore had data on 100 trials (two responses to each of the 50 choice problems, one from each session). For one participant the data only contained 99 responses due to a technical error. Seven participants in a total of 38 trials provided a response but did not inspect any information before making a choice. Since no attention data was available for these trials, they were not included in the analysis. In total, the analysis included data on 8,961 trials from 90 participants.

For each participant s and trial i, we calculated the total time that the participant spent inspecting information on probabilities, $t_{s,i}^{prob}$, the total time they inspected information on the ML option, $t_{s,i}^{ML}$, and the total time they spent inspecting any type of information, $t_{s,i}$. Attribute attention was then defined as the proportion of time spent inspecting probability information, relative to the total inspection time, $a_{s,i}^{attribute} = t_{s,i}^{prob} / t_{s,i}$. Option attention was defined as the proportion of time spent inspecting information about the ML option, relative to the total inspection time, $a_{s,i}^{option} = t_{s,i}^{ML} / t_{s,i}$. For readability, in the Results section we will refer to the indices without the subscripts (i.e., $a^{attribute}$ instead of $a_{s,i}^{attribute}$ and a^{option} instead of $a_{s,i}^{option}$). Note that Pachur et al. (2018) used the absolute time spent attending to probabilities as a measure of attribute attention. Measures of attention based on the absolute time spent attending to specific pieces of information, however, can be confounded by the total time spent on the problem. For instance, if more absolute time attending to probabilities is associated with more linear probability weighting, then this might reflect that more diligent people, who overall take more time on each problem (and therefore also spend more time attending to probabilities), behave more consistently with an objective treatment of probabilities (see Pachur et al., 2018). To avoid this potential confound, the current analyses use relative attention scores, for both attribute attention and option attention. In Appendix C we present additional analyses testing whether people tended to allocate their attention more or less evenly across options and attributes when they spent overall more time searching for information preceding a choice.

We rely on Bayesian methods for data analysis. For the analyses of participants' choice behavior and attention allocation during information search, we used general linear models (GLMs), estimated with the rstanarm package in R (Goodrich, Gabry, Ali, Brilleman, & Buros Novik, 2018). The effect of a given predictor variable is considered credible if the corresponding 95% posterior interval (PI) excluded zero.

4.2. Computational modeling

We modeled participants' choices using CPT. Recall that CPT assumes a value function that transforms an option's outcomes into subjective values, and a probability-weighting function that yields decision weights for the outcomes based on their cumulative probabilities. According to the value function, v, the subjective value of an outcome x is given by

$$v(x) = \begin{cases} x^{\alpha^+}, & if\ x \geq 0 \\ -|x|^{\alpha^-}, & if\ x < 0. \end{cases} \tag{1}$$

When $0 < \alpha < 1$, the value function is concave for gains and convex for losses (implying risk aversion for gains and risk seeking for losses); the opposite is the case when $\alpha > 1$. When $\alpha = 1$, the value function is linear, implying an objective treatment of the outcomes' values. Note that because our analyses did not include mixed gambles, we did not estimate a loss aversion parameter.

The subjective values of the outcomes are weighted by decision weights π. CPT assumes rank dependency, meaning that the weight attached to an outcome depends on its rank among the possible outcomes of an option. Specifically, π is derived from the outcomes' cumulative probabilities. In a cumulative distribution, the N positive (M negative) outcomes are ordered from best to worst (worst to best), $x_1^+ > x_2^+ > ... > x_N^+$ ($x_1^- < x_2^- < ... < x_M^+$), and the cumulative probability of an outcome x_i represents the probability of obtaining an outcome at least as good (bad) as x_i. The decision weights are defined as the difference between the probability of obtaining an outcome at least as good (bad) as x_i and the probability of obtaining a strictly better (worse) outcome, transformed by a nonlinear probability-weighting function w (Tversky & Kahneman, 1992):

$$\begin{aligned} \pi_1^+ &= w^+(p_1) \\ \pi_1^- &= w^-(p_1) \\ \pi_n^+ &= w^+(p_1 + \ldots + p_n) - w^+(p_1 + \ldots + p_{n-1})\ for\ 1 < n \leq N \\ \pi_m^- &= w^-(p_1 + \ldots + p_m) - w^-(p_1 + \ldots + p_{m-1})\ for\ 1 < m \leq M. \end{aligned} \tag{2}$$

w^+ and w^- are probability-weighting functions that transform the cumulative probabilities of gain and loss outcomes, respectively.

Of the various functional forms of w that have been proposed, we rely on Prelec's (1998) two-parameter probability-weighting function:

$$w(p) = e^{-\delta(-log(p))^\gamma}. \tag{3}$$

The parameter γ (> 0) governs the curvature of the probability-weighting function. When $\gamma < 1$, the probability-weighting function is inverse S-shaped (or convex), implying an overweighting of rare events and an underweighting of events with medium to large probabilities. Moreover, when $\gamma < 1$ the probability-weighting function discriminates less clearly between differences in probabilities in the medium range (i.e., reduced probability sensitivity; Gonzalez & Wu, 1999). When $\gamma > 1$, the probability-weighting function is S-shaped (or concave), implying an underweighting of small-probability events and an overweighting of events with medium to large probabilities. The parameter δ (> 0) governs the elevation of the probability-weighting function. The further δ is below 1, the higher the elevation of the probability-weighting function, such that probabilistic events (depending on their rank) receive more weight than they objectively deserve. This implies risk seeking in the gain domain and risk aversion in the loss domain. The more δ exceeds 1, the lower the elevation of the probability-weighting function, such that probabilistic events (depending on their rank) receive less weight than they objectively deserve. This implies risk aversion in the gain domain and risk seeking in the loss domain. Elevation has therefore been interpreted as reflecting how attractive the decision maker finds gambling (Gonzalez & Wu, 1999). When $\gamma = 1$ and $\delta = 1$, the probability-weighting function is linear, implying a weighting of events based on their objective probabilities. In Appendix F we show that the main conclusions of our analyses also hold when using another commonly applied probability-weighting function, namely the one by Goldstein and Einhorn (1987).

The overall evaluation of an option A with a total of J ($J = N$ in the gain domain and $J = M$ in the loss domain) outcomes is given by summing up the subjective values of the option's outcomes, each weighted by its decision weight:

$$V(A) = \sum_{j=1}^{J} v(x_j) \times \pi_j. \tag{4}$$

Based on the valuations of two options A and B, in our analyses we derived the predicted probability of choosing option A using the logit choice rule (also known as softmax),

$$p(A, B) = \frac{1}{1+e^{-\rho[V(A)-V(B)]}}. \tag{5}$$

The parameter ρ (> 0; which we estimated for each participant) is a choice-sensitivity parameter. The higher ρ, the more closely the probability of choosing the option with the higher valuation V approaches 1; given $\rho = 0$, choices are random with respect to the evaluations of the options. In models that combine a nonlinear value function with a stochastic choice rule such as softmax, the parameters ρ and α are often highly correlated. This can distort parameter estimation (Krefeld-Schwalb, Pachur, & Scheibehenne, 2022; Vincent & Stewart, 2020). To counter this problem, we applied the reformulation suggested by Stewart, Scheibehenne, and Pachur (2018) and Krefeld-Schwalb et al. (2022, see their Eq. 18). Due to the hierarchical model structure, the individual-level estimates of ρ were informed by a group-level distribution.

Based on the findings by Pachur et al. (2018), we expected that higher attribute attention (i.e., higher relative attention to probabilities) would be reflected in higher values of γ, irrespective of domain. Note that Pachur et al. (2018) did not estimate γ separately for gains and losses; therefore, it is unclear whether the same type of association would emerge in both domains when estimating γ separately for each (as in the current analyses; see details below). Moreover, based on the findings by Pachur et al. (2018), we expected that higher values of $a_{s,i}^{attribute}$ would be linked to a lower elevation (i.e., higher δ) in the gain domain, but not in the loss domain. What pattern could we expect regarding a link between option attention and probability weighting? In previous analyses, we (Zilker & Pachur, 2022c) showed that the predicted association between option attention and probability weighting depends on whether attending more to an option has a positive or a negative effect on choosing that option. Specifically, if more attention paid to the ML option increases the probability of choosing that option in the gain (loss) domain, then higher values on $a_{s,i}^{option}$ are associated with higher (lower) values on γ and higher (lower) values on δ in the gain (loss) domain. If, instead, more attention paid to the ML option decreases the probability of choosing that option in the gain (loss) domain, then higher values on $a_{s,i}^{option}$ are associated with lower (higher) values on γ and lower (higher) values on δ in the gain (loss) domain. As reported in the "Results" section, in our empirical analyses we found a positive link between option attention and choice in both domains. Therefore, we expect that higher values of $a_{s,i}^{option}$ would be associated with higher (lower) values on γ and with higher (lower) values on δ in the gain (loss) domain.

4.2.1. Regression framework for estimating CPT parameters

We used a hierarchical Bayesian implementation of CPT (see also Nilsson, Rieskamp, & Wagenmakers, 2011; Pachur, Mata, & Hertwig, 2017; Scheibehenne & Pachur, 2015; Zilker, Hertwig, & Pachur, 2020). To test how attribute attention and option attention are associated with the CPT parameters γ, δ, and α, we applied a regression framework for parameter estimation (see Boehm, Steingroever, & Wagenmakers, 2018; Vandekerckhove, Tuerlinckx, & Lee, 2011). This framework allowed us to estimate the CPT parameters while concurrently testing, using the parameters of a regression model, the extent to which the CPT parameters are associated with other variables—in this case, the attentional variables collected with Mouselab. This one-step procedure avoids biases in inference that may arise when first estimating the CPT parameters and subsequently submitting them to statistical tests (for details see Boehm, Marsman, Matzke, & Wagenmakers, 2018).

The CPT parameters γ, δ, and α were expressed as a linear combination of a fixed intercept and fixed slopes for each attentional variable, as well as person-specific random intercepts and person-specific random slopes for each attentional variable. A given CPT parameter θ for participant s on trial i was thus expressed as[3]

$$\begin{aligned}\theta'_{s,i} = {} & \beta_0^\theta + \beta_{0,s}^\theta \\ & + \left(\beta_{a^{\text{attribute}}}^\theta + \beta_{a^{\text{attribute}},s}^\theta\right) \times a_{s,i}^{\text{attribute}} + \left(\beta_{a^{\text{option}}}^\theta + \beta_{a^{\text{option}},s}^\theta\right) \times a_{s,i}^{\text{option}} \\ & + \left(\beta_{a^{\text{option}\times\text{attribute}}}^\theta + \beta_{a^{\text{option}\times\text{attribute}},s}^\theta\right) \times a_{s,i}^{\text{option}} \times a_{s,i}^{\text{attribute}}\end{aligned} \tag{6}$$

where β_0^θ and $\beta_{0,s}^\theta$ are the fixed intercept and the person-specific random intercepts, respectively. $\beta_{a^{attribute}}^\theta$ and $\beta_{a^{attribute},s}^\theta$ are the fixed slope and the person-specific random slopes, respectively, for the effect of attribute attention on $\theta'_{s,i}$. $\beta_{a^{option}}^\theta$ and $\beta_{a^{option},s}^\theta$ are the fixed slope and the person-specific random slopes, respectively, for the effect of option attention on $\theta'_{s,i}$. $\beta_{a^{option\times attribute}}^\theta$ and $\beta_{a^{option\times attribute},s}^\theta$ are the fixed slope and the person-specific random slopes, respectively, for the interactive effect of option attention and attribute attention on $\theta'_{s,i}$. The predictor variables $a_{s,i}^{attribute}$ and $a_{s,i}^{option}$ were z-standardized before they were supplied to the model; their estimated regression coefficients are therefore standardized. Note that when estimating CPT using such a hierarchical estimation approach, parameters on the trial-level are informed and constrained by parameters on higher levels of the hierarchy (i.e., individual- and group-level parameters). This allows us to capture the effects of attention on different levels of the hierarchy.

$\theta'_{s,i}$ can, in principle, take on values on the entire real line. In order to serve as inputs to CPT's functions, the values need to be rescaled. The curvature parameter γ was scaled to a range between 0 and 2, the elevation parameter δ was scaled to a range between 0 and 5, and the curvature of the value function α was scaled to a range between 0 and 2 (cf. Pachur et al., 2018). For the rescaling, a given parameter θ' was phi-transformed (yielding values between 0 and 1; Rouder & Lu, 2005), and multiplied by the value representing the upper bound of the respective parameter range, max_θ:

$$\theta = \Phi(\theta') \times max_\theta. \tag{7}$$

All parameters were estimated separately for the gain and loss domains.

Although our focus in this article is on the association between attention and probability weighting, we also estimated the parameter α of CPT's value function (Eq. 1) using the regression approach. The analyses by Pachur et al. (2018) demonstrated that attribute attention may not only modulate the shape of the probability-weighting function, but also that of the value function. Our modeling framework allows us to also accommodate attentional effects on the value function—even if they are not at the heart of our current research question. The results regarding α are reported in Appendix A.

4.2.2. Model variants

In addition to the *full model* described above—which includes attribute attention, option attention, and their interaction as predictors of the CPT parameters—we also considered two reduced models. In the *attribute-attention model*, each CPT parameter was estimated based on a linear combination of an intercept and the effect of attribute attention (i.e., $a_{s,i}^{attribute}$) only. In the *option-attention model*, each CPT parameter was estimated based on a linear combination of an intercept and the effect of option attention (i.e., $a_{s,i}^{option}$) only. We use the reduced models to gauge, in a first step, whether and how attribute attention is associated with the CPT parameters when not taking option attention into account, and vice versa. Further, the estimated regression coefficients for attribute and option attention obtained with the reduced models provide reference values for the regression coefficients obtained with the full model; this comparison allows us to assess in another way whether the effects of each attentional variable are interdependent.

4.2.3. Parameter estimation procedure

We estimated the posterior parameter distributions using Markov Chain Monte Carlo methods implemented in JAGS and using the R2jags package (Su & Yajima, 2015). We ran 30 chains of 25,000 samples each, including 5,000 burn-in samples. The chains were thinned such that only every second sample was recorded. The potential scale reduction factor (Gelman & Rubin, 1992) was $\hat{R} \leq 1.01$ for all estimated parameters, indicating good convergence. In the "Results" section we rely on the group-level estimates of the fixed effects of the attentional variables (e.g., $\beta_{a^{attribute}}^\gamma$), and on the trial-level estimates of the CPT parameters (e.g., $\gamma_{s,i}$). A parameter recovery analysis reported in Appendix B demonstrates that with the present estimation approach these parameters can be reliably identified.

[3] The parameter θ serves as a placeholder for γ, δ, or α, since all of these parameters were defined in the described manner. θ' denotes the respective parameter before transformation to the target range (see Eq. (7) for details).

5. Results

5.1. Attention allocation

We first provide an overview of patterns in attribute and option attention in the analyzed data. Across trials and participants, attribute attention—that is, the relative amount of attention to probabilities vs. outcomes—was, on average, $M = 0.452$ $(SD = 0.141)$ in the gain domain and $M = 0.442$ $(SD = 0.141)$ in the loss domain. This indicates that, on average, participants tended to spend slightly more time attending to outcomes than to probabilities. Option attention—that is, the relative amount of attention to the ML vs. the LL option—was, on average, $M = 0.485$ $(SD = 0.151)$ in the gain domain and $M = 0.503$ $(SD = 0.136)$ in the loss domain. This indicates that, on average, participants tended to allocate their attention relatively evenly across the two options in a problem. Additional analyses (reported in Appendix C) show that overall longer inspection time of a problem was associated with more balanced attribute attention, but with more extreme biases in option attention.

We also examined to what extent variability in attribute and option attention was driven by individual differences between participants and by differences between choice problems. To explore the contribution of person-specific and item-specific factors, we conducted ANOVAs in which either of the attentional variables (i.e., $a^{attribute}$ or a^{option}) was used as dependent variable and unique identifiers for participants and choice problems were used as predictors. We estimated the models separately for the gain and loss domains and computed the proportion of variance accounted for by each of the two factors. For attribute attention in the gain (loss) domain, differences between participants accounted for 28.5% (30.4%) of the variance, and differences between choice problems for 2.2% (0.8%). That is, variability in attribute attention was to a considerably stronger degree due to differences between decision makers than to differences between choice problems. For option attention in the gain (loss) domain, differences between participants accounted for 2.6% (1.6%) of the variance, and differences between choice problems for 3.5% (5.2%) (see Appendix D for additional analyses).

Next we turn to the results of the computational modeling and to the question of whether variation in attention allocation across options and attributes was linked to differences in probability weighting.

5.2. Model performance

We first examined how well CPT with the parameters estimated with the different models captured participants' choices. To that end, we determined the choices generated by each model based on the posterior trial-level distributions of the parameters. Then we compared these choices to the empirical ones and calculated the proportion of overlapping choices. The full model, which incorporated the potential effects of both attribute attention and option attention and their interaction on CPT parameters and is thus the most complex model, matched 74.7% of participants' choices. This was followed by the option-attention model (72.4%) and the attribute-attention model (71.9%). Overall, the models thus captured the empirical choices quite well. The attribute-attention model had the lowest deviance information criterion (DIC; Spiegelhalter, Best, Carlin, & Van Der Linde, 2002, DIC = $8{,}934$), closely followed by the option-attention model (DIC = $8{,}942$) and the full model (DIC = $9{,}859$). Note that the DIC (with a lower DIC indicating better model performance) does not quantify mere model fit but penalizes for model complexity. The full model is considerably more complex than the reduced models (because it includes an additional fixed effect as well as additional random effects); the full model's higher DIC thus indicates that its higher complexity is not matched by a proportional increase in fit. Nevertheless, to statistically test our research question—that is, whether the effects of attribute attention and option attention on probability weighting are independent from each other—it is necessary to include an interaction term, and it thus seems warranted that we use the full model for our inferences.

Finally, we compared the performance of all models to an intercept-only model, that represents the standard implementation of CPT and does not allow for trial-by-trial variability in parameters depending on attention. Instead, in the intercept-only model the CPT parameters $\alpha, \gamma,$ and δ are estimated only with person-specific random intercepts (thus allowing for differences between participants). This model matched 69.3% of the empirical choices, consistent with existing studies that have found that standard CPT fits choice data quite well (e.g., Glöckner & Pachur, 2012; Kellen, Pachur, & Hertwig, 2016). When taking into account differences in model complexity, with a DIC of 8,748 the intercept-only model outperformed both the reduced and the full models (but note that the while the intercept-only model is the least complex, it does not allow us to test for any effects of attention). Overall, the results underscore that allowing CPT parameters to covary with measures of attention allocation improves the models' ability to accommodate the data, although this improvement comes at the cost of increased model complexity.

5.3. Is probability weighting linked to attention?

We next examined whether attribute attention and option attention are associated with probability weighting when, as in Pachur et al. (2018) and Zilker & Pachur (2022c), the two links are analyzed independently. For these analyses, we relied on the reduced models. Fig. 3 displays the posterior group-level means and 95% posterior intervals (PIs) for the regression coefficients obtained with the attribute-attention model (Fig. 3A) and the option-attention model (Fig. 3B). The coefficients indicate the extent to which attribute attention and option attention, respectively, are related to the probability-weighting parameters γ and δ. Fig. 4 displays the trial-level estimates of γ and δ, conditional on $a^{attribute}$ and on a^{option}.

5.3.1. Probability weighting and attribute attention

How was attribute attention associated with the curvature of the probability-weighting function—that is, the γ parameter? Focusing on the attribute-attention model, Fig. 3A shows that the posterior mean of the coefficient $\beta^{\gamma}_{a^{attribute}}$ is positive for the domains of both gains and losses. These effects are credible, indicated by the 95% posterior intervals not including zero. Paying more attention to probabilities (relative to out-comes) was thus associated with higher values of the γ parameter. This pattern is also evident when considering the trial-level estimates of γ in Fig. 4. Regarding the elevation δ of the probability-weighting function, the regression coefficient $\beta^{\delta}_{a^{attribute}}$ indicates that there were no credible effects of attribute attention in either the gain or the loss domain. Overall, like Pachur et al. (2018) we conclude that attention to probability information was linked to probability weighting. Importantly, however, the current analyses go beyond those of Pachur et al. (2018), who had relied on measures of absolute (rather than relative) attention, accommodated only individual variability between participants (and not between trials), and did not estimate the CPT parameters and the effects of the attentional measures in a single step.

5.3.2. Probability weighting and option attention

We next examined the extent to which option attention was associated with probability weighting in the Pachur et al. (2018) data set. The association between option attention and probability weighting depends on how option attention is related to choice (for details see Zilker & Pachur, 2022c). In Zilker & Pachur (2022c), we found based on sampling behavior in decisions from experience as well as on fixation behavior measured with eye tracking that attending more to the ML option was linked to a higher probability of choosing it. We therefore first analyzed whether more attention paid to the ML option was also related to a stronger preference for that option when attention is measured with Mouselab. To this end, we estimated logistic GLMs, using choice of the ML option as the dependent variable and a^{option} as the

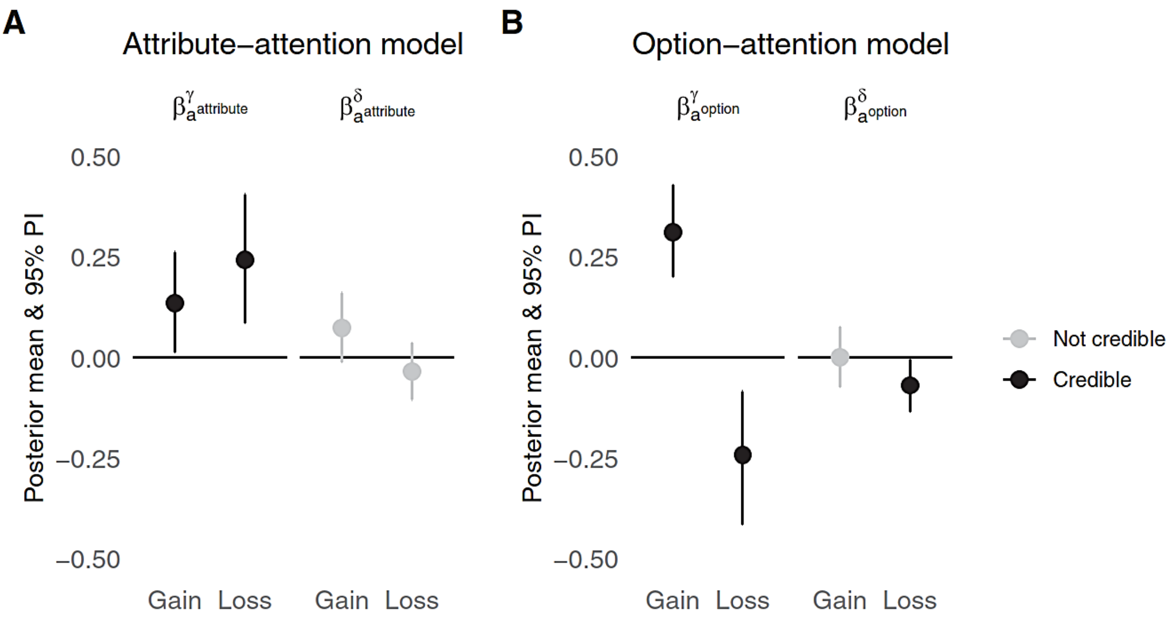

Fig. 3. Results of the attribute-attention model (Panel A) and the option-attention model (Panel B) regarding the association between probability weighting and attention. Shown are the group-level posterior means and 95% posterior intervals (PI) of the regression coefficients of the attentional variables for predicting the parameters γ and δ of cumulative prospect theory's probability-weighting function, separately for each model. Attribute attention is defined as the relative amount of attention to probability information; option attention is defined as the relative amount of attention to the ML option.

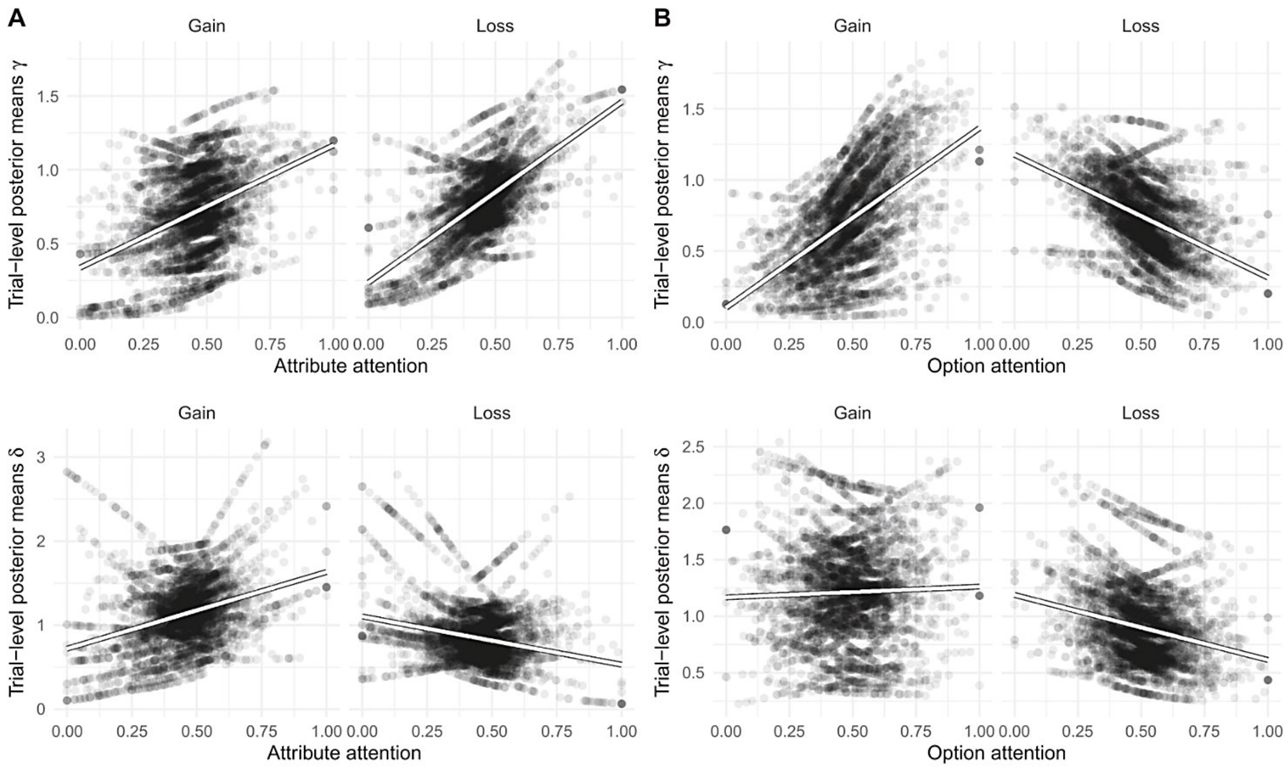

Fig. 4. Association between trial-level parameter estimates of cumulative prospect theory's probability-weighting function and attention. Panel A: Results regarding attribute attention from the attribute-attention model. Panel B: Results regarding option attention from the option-attention model.

predictor variable, separately for the gain and loss domains. Option attention had a positive and credible association with the tendency to choose the ML option, both in the gain domain ($\beta_{a^{option}} = 0.605, PI_{95\%}$ [0.512; 0.697]) and in the loss domain ($\beta_{a^{option}} = 0.250, PI_{95\%}$ [0.144; 0.355]).

Given the positive link between option attention and choice, one would expect that higher values on a^{option} are associated with higher (lower) values on γ and higher (lower) values on δ in the gain (loss) domain. Is there evidence for such an association? Focusing on the option-attention model, Fig. 3B shows that the posterior mean of the coefficient $\beta^{\gamma}_{a^{option}}$ is positive and credible in the domain of gains and negative and credible in the domain of losses. That is, paying more attention to the ML option was linked to higher values of γ in the gain domain and lower values of γ in the loss domain. This pattern is also evident when considering the trial-level estimates in Fig. 4.

What about a possible link between option attention and the elevation of the probability-weighting function, that is, the δ parameter? Based on the posterior mean of $\beta^{\delta}_{a^{option}}$ in the option-attention model, option attention was not credibly associated with δ in the gain domain, and credibly but only very weakly and negatively related to δ in the loss domain. Consistent with analyses in our (Zilker & Pachur, 2022c) previous study, option attention was thus considerably more strongly associated with γ than with δ in choice problems offering two risky options. Although small effects of option attention on δ can be identified in simulations (see Zilker & Pachur, 2022c), they seem to be elusive in empirical data.

5.3.3. Are attribute attention and option attention independently associated with probability weighting?

The previous section established, in separate analyses, that attribute attention and option attention are associated with probability weighting within the same data set and when relying on the same measure of attention and the same analytic strategy. Next, we address whether these two links between attention and probability weighting represent independent phenomena. In a first step, we examined the extent to which attribute attention and option attention are correlated. For instance, trials in which decision makers pay less attention to probabilities might also be those in which they attend less to the ML option. In a second step, we tested whether the observed associations of attribute attention and of option attention with probability weighting also emerge when the two facets of attention allocation as well as their interaction were simultaneously included in the statistical model.

5.3.4. Are attribute attention and option attention related?

Fig. 5 plots trial-wise attribute attention $a^{attribute}$ against trial-wise option attention a^{option} and shows that the two attentional variables were barely correlated. To examine the association statistically, we estimated GLMs with option attention as the dependent variable and attribute attention as the predictor variable, separately for the gain and the loss domain. In the gain domain attribute attention was credibly but weakly related to option attention ($\beta^{a^{option}}_{a^{attribute}} = 0.038, PI_{95\%}\ [0.007;\ 0.070]$). That is, in trials in which participants spent relatively more time attending to probabilities, they also spent slightly more time attending to the ML option. We also calculated a Bayesian analogue to R^2 (Gelman, Goodrich, Gabry, & Vehtari, 2019). The posterior mean Bayesian R^2 for the domain of gains equaled 0.0015, indicating that attribute attention accounted for little variance in option attention, even if the effect was credible. In the domain of losses, attribute attention was not credibly related to option attention ($\beta^{a^{option}}_{a^{attribute}} = -0.013, PI_{95\%}\ [-0.042;\ 0.016]$). The posterior mean Bayesian R^2 for the domain of losses equaled 0.0004. Overall, these analyses indicate that imbalances in the allocation of attention to probabilities versus outcomes and imbalances in the allocation of attention across options were largely independent from each other.

5.3.5. Do attribute attention and option attention have separate effects on probability weighting?

Next, we addressed whether the effects of the two facets of attention allocation on probability weighting were also independent from each other. To that end, we consider the full model (see Eq. (6)), in which both attribute attention and option attention as well as their interaction are included as predictors of the parameters of CPT's probability-weighting function.

Fig. 6A shows that the effects of attribute attention on γ were credible in both the gain and the loss domain, and also pointed in the same direction as the corresponding effects in the reduced model (i.e., the attribute-attention model). The positive association between attribute attention and δ in the domain of gains was credible in the full model (it was not credible in the attribute-attention model). Attribute attention had no credible effect on δ in the domain of losses in the full model, as was the case in the attribute-attention model.

Fig. 6B shows that the effects of option attention on γ were credible in both the gain and the loss domain, and also pointed in the same direction as the corresponding effects in the option-attention model. Option attention was not credibly associated with δ in the full model in the domain of gains, analogous to the finding from the option-attention model. In the domain of losses, option attention showed a credible negative association with δ in the full model, also echoing the finding obtained using the option-attention model.

Overall, the estimated effects of attribute attention and option attention in the full model were thus almost identical to those obtained with the reduced models (see Fig. 3). The $PI_{95\%}$s of the effects obtained with the full versus the reduced models also largely overlapped. This indicates that both facets of attention allocation influenced the parameters of the probability-weighting function even when the effect of the other attentional variable and their potential interaction was controlled for. Finally, Fig. 6C shows that the interaction between attribute attention and option attention on γ and δ was not credibly different from zero in either domain. The two facets of attention allocation thus independently influenced the shape of the probability-weighting function.

5.3.6. Distinct patterns due to biases in attribute attention and option attention

Building on the results from the previous sections, we finally elaborate to what extent attribute attention and option attention can produce distinguishable patterns in the shape of the probability-weighting function. Given that both facets of attention are linked with probability weighting, does a change in one facet have a similar effect as a corresponding change in the other facet, even if the two effects are independent? Or do attribute attention and option attention give rise to distinct patterns in the overall shape of the weighting function, that can be produced by one but not the other?

As a starting point, assume that attribute attention and option attention are both balanced (i.e., $a^{attribute} = 0.5$ and $a^{option} = 0.5$). According to our full model, such an attentional pattern leads to predicted values of $\gamma = 0.73$ ($\gamma = 0.82$) and $\delta = 1.25$ ($\delta = 0.91$) in the gain (loss) domain. The resulting probability-weighting functions are illustrated in Fig. 7, marked by dashed lines. Note that given such balanced attention, probability weighting comes relatively close to linear weighting—that is, weighting of events based on their objective probabilities. When attention becomes biased, either to a particular attribute or option, also probability weighting becomes more nonlinear. But are biases in attribute attention and option attention and option attention linked to distinct distortions of the probability-weighting function? This question is addressed next.

If there is a bias in option attention towards the ML option in the gain domain, this will result in a more strongly S-shaped probability-weighting function but it will not affect the elevation (as indicated by the regression coefficients in Fig. 6). Would a bias in attribute attention have a similar effect? If attention were increasingly biased to probabilities in the gain domain, this would also result in a more strongly S-shaped probability-weighting function; importantly, however, this would be accompanied by a lower elevation (higher values on δ; see Fig. 6). This example can also be substantiated numerically. For instance, when attribute attention is balanced (i.e., $a^{attribute} = 0.5$) while the decision maker exclusively attends to the ML option (i.e., $a^{option} = 1.0$), the model predicts values of $\gamma = 1.6$ and $\delta = 1.1$. In contrast, when option attention is balanced (i.e., $a^{option} = 0.5$) while the decision maker exclusively attends to probabilities (i.e., $a^{attribute} = 1.0$), the model predicts values of $\gamma = 1.2$ and $\delta = 1.9$. This highlights the distinct consequences of biases in attribute attention and options attention on the elevation parameter.

Similar differences between the effects of attribute attention and

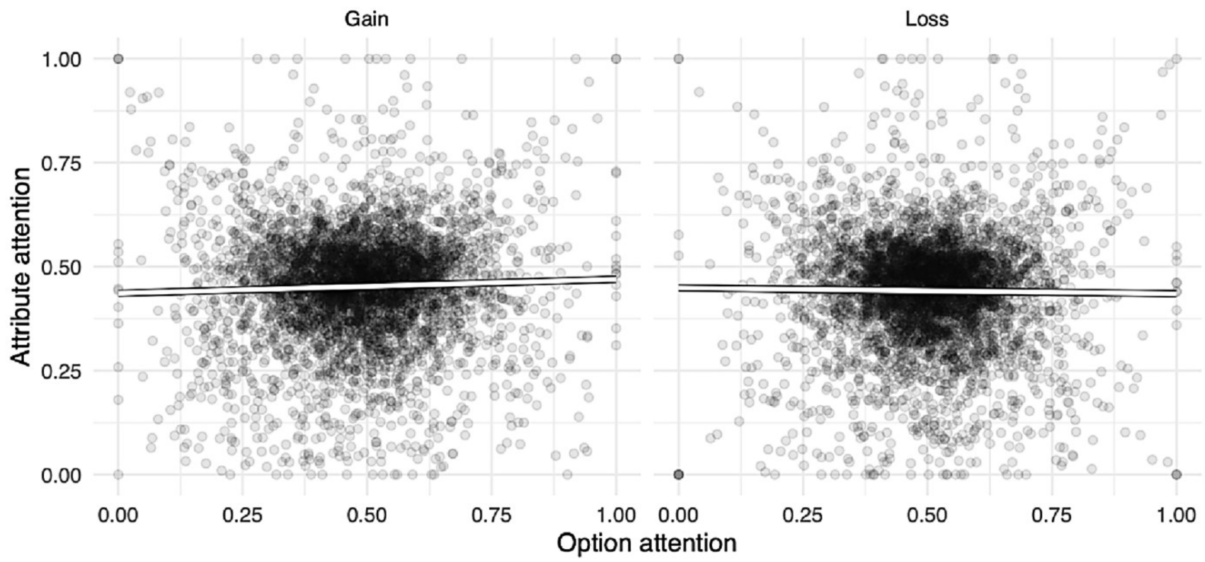

Fig. 5. Association between trial-wise attribute attention (defined as the proportion of time attending to probabilities, $a^{attribute}$) and trial-wise option attention (defined as the proportion of time attending to the ML option, a^{option}), separately for the domains of gains and losses.

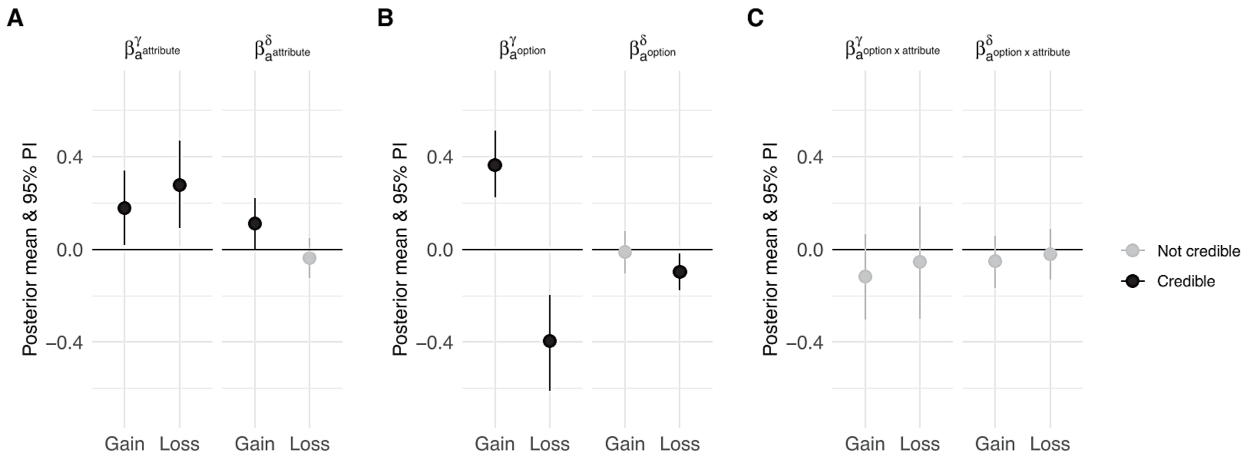

Fig. 6. Results of the full model regarding the association between probability weighting and attention. Shown are the group-level posterior means and 95% posterior intervals (PI) of the regression coefficients of (A) attribute attention, (B) option attention, and (C) their interaction for predicting the parameters γ and δ of cumulative prospect theory's probability-weighting function. Attribute attention is defined as the relative amount of attention to probability information; option attention is defined as the relative amount of attention to the ML option.

option attention on probability weighting exist in the domain of losses. Here, a bias in option attention towards the ML option results in a more strongly inverse S-shaped probability-weighting function, accompanied by a higher elevation (lower values on δ). While a bias in attribute attention (towards outcomes rather than probabilities) could also produce a stronger inverse S-shape, there would be no change in the elevation (see Fig. 6). Also the example from the loss domain can be substantiated numerically. For instance, when attribute attention is balanced (i.e., $a^{attribute} = 0.5$) while the decision maker exclusively attends to the ML option (i.e., $a^{option} = 1.0$), the model predicts values of $\gamma = 0.1$ and $\delta = 0.5$. In contrast, when option attention is balanced (i.e., $a^{option} = 0.5$) while the decision maker exclusively attends to outcomes (i.e., $a^{attribute} = 0$), the model predicts values of $\gamma = 0.3$ and $\delta = 1.1$.

To visually illustrate the distinct effects of attribute attention and option attention on probability weighting, Fig. 7 displays the shapes of the probability-weighting functions predicted by the full model, with the color of each curve indicating different levels of attribute attention (Fig. 7A) and option attention (Fig. 7B). These probability-weighting functions were obtained by using the estimates for the group-level regression coefficients from the full model to predict the values of the parameters γ and δ while systematically varying either option attention or attribute attention but keeping the other facet of attention balanced (i.e., assuming a value of 0.5). As can be seen, there is a clear association between the shape of the resulting probability-weighting functions and the color in which they are displayed—a powerful illustration of the effects of both facets of attention and of how these effects differ between the gain and loss domains.

In sum, differences in attribute attention and option attention seem to entail distinct patterns in probability weighting—further underscoring that the links of attribute attention and option attention with probability weighting represent separate phenomena.

6. Discussion

Nonlinear probability weighting is one of the most influential theoretical constructs in descriptive models of risky choice (e.g., Bhatia, 2014; Birnbaum, 2008; Lopes, 1987; Tversky & Kahneman, 1992). Recent analyses (Pachur et al., 2018; Zilker & Pachur, 2022c) showed

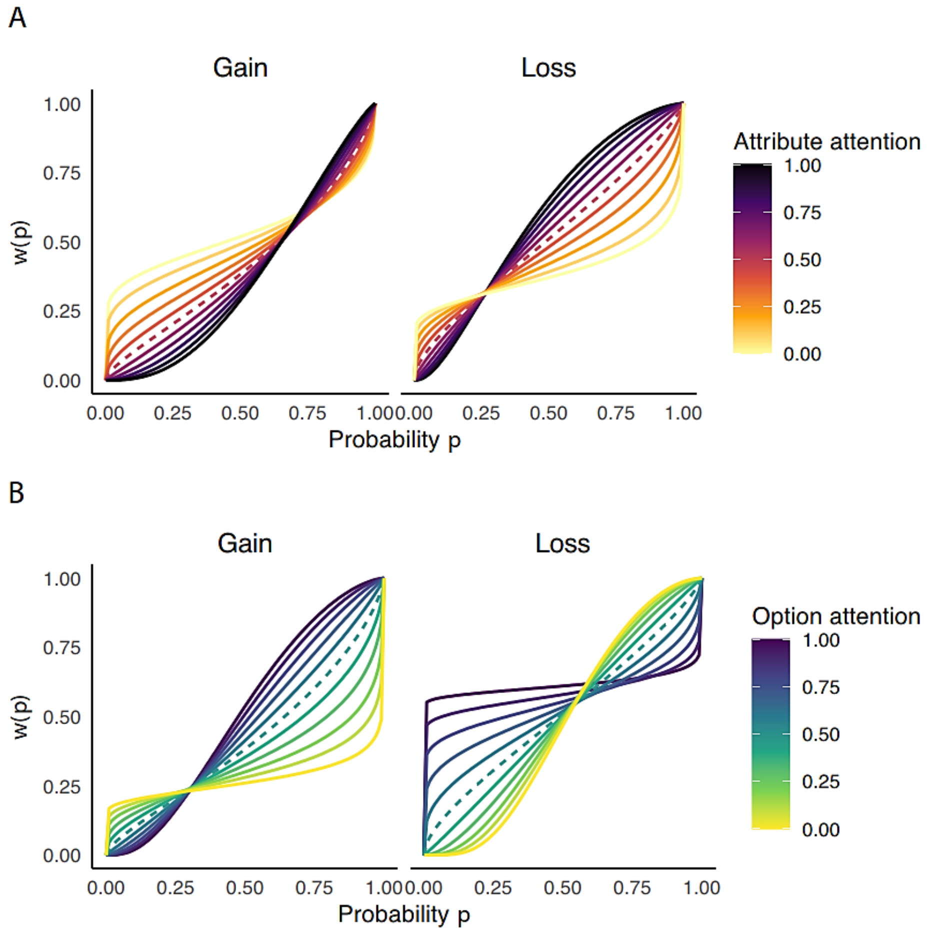

Fig. 7. Predicted effects of option attention and attribute attention on the shape of the probability-weighting function. Panel A: Probability-weighting functions for different levels of attribute attention (i.e., the relative amount of attention to probabilities vs. outcomes). Each probability-weighting function is based on the values for the γ and δ parameters predicted from the estimated regression coefficients. Panel B: Probability-weighting functions for different levels of option attention (i.e., the relative amount of attention to the ML vs. the LL option). Each probability-weighting function is based on the values for the γ and δ parameters predicted from the estimated regression oefficients. Dashed lines mark the weighting functions predicted when attribute attention and option attention are both balanced (i.e., $a^{attribute} = 0.5$ and $a^{option} = 0.5$).

that patterns in attention allocation across attributes (probabilities versus outcomes) and across options are systematically linked to the shape of CPT's probability-weighting function. These analyses point to possible cognitive underpinnings of choice phenomena captured in characteristic shapes of the probability-weighting function such as the certainty effect and the fourfold pattern of risk attitudes. It is currently unclear, however, how these links between attention and preference—obtained in different decision-making paradigms, with different measures of attention and data-analytic procedures—are related to each other. As a consequence, it is difficult to assess the implications of these findings for existing psychological interpretations of the shape of the probability-weighting function and the challenges they pose to the development of mechanistic (i.e., process-level) theories of decision making under risk.

The current work investigated whether attribute attention and option attention independently contribute to the shape of the probability-weighting function in the same data set, or whether their effect(s) can be ascribed to shared variance. Using data from a process-tracing study with Mouselab, our analyses demonstrated that both attribute attention and option attention are related to probability weighting within the same data set, when using the same measure of attention and analytic approach. More importantly, these two attentional effects on nonlinear probability weighting seem to be largely independent from each other. We showed that empirical patterns in attribute attention and option attention are uncorrelated. The effects of attribute attention and option attention on probability weighting are practically unchanged when both are considered simultaneously in the analysis, and their interaction effect was not credible. Our analyses also underlined that deviations from linear probability weighting are consistently tied to imbalances in attention allocation. When attribute attention and option attention are balanced, choices are approximately in line with linear probability weighting; when attention is biased on one or both of these dimensions, choices are in line with nonlinear probability weighting. In the following, we discuss implications of our findings.

6.1. Implications for cognitive process models of risky choice

Our analyses demonstrate and disentangle distinct attentional underpinnings of characteristic preference patterns in decisions under risk—patterns captured in the shape of the nonlinear probability-weighting functions. Our key finding—that attribute attention and option attention independently shape such preferences—poses a challenge for cognitive process models of risky choice. Many existing models tend to focus on one of these facets only; new models should strive to incorporate both and their respective effects on preference. We next illustrate this point by reference to some prominent existing theories that accommodate a link between attention and preference, in risky choice and other areas of decision making.

The link between option attention and preference has been formalized in models belonging to the sequential sampling tradition. For instance, the aDDM (Krajbich et al., 2010) posits that evidence accumulates faster for an option when it is attended to than when it is not attended to. Imbalances in the distribution of attention across options can thus lead to biases in choice behavior. Notably, the aDDM was originally developed in value-based choice tasks offering food items, where options are presented as bundles, without explicitly distinguishing between different attributes such as outcomes and probabilities. Therefore, the original aDDM cannot accommodate differential attention allocation across attributes, and its effects on preference; it also cannot accommodate signatures of attribute attention in nonlinear probability weighting (unless differences in attribute attention are linked to differences in option attention; our results speak against such an association). In another model belonging to the sequential sampling framework, Johnson and Busemeyer (2016) demonstrated with an extension of decision field theory (Busemeyer & Townsend, 1993) how attention allocation across different outcomes in risky choice can lead to different patterns of probability weighting. However, the assumed process of sampling from the outcome distribution does not allow the model to accommodate that attention might be explicitly directed to probability information (in choice tasks where such information is explicitly provided). Consequently, the model cannot capture that the relative amount of attention paid to probabilities versus outcomes might shape preferences.

Another class of models, heuristics, often predicts specific patterns in attribute attention in risky choice. For instance, according to the minimax and the maximax heuristics (Coombs, Dawes, & Tversky, 1970; Savage, 1951), the decision maker focuses on the outcomes of the options while neglecting probability information, and this can give rise to characteristic patterns in probability weighting (Pachur, Suter, & Hertwig, 2017). However, it is not clear how heuristics might give rise to option-specific biases in attention. None of the models discussed thus far can offer a comprehensive account of how attribute attention and option attention shape preference and probability weighting in risky choice.

In the context of multiattribute choice, extensions of the aDDM have been proposed that allow for biases in both attribute attention and option attention (e.g., Fisher, 2017; Yang & Krajbich, 2022). Notably, multiattribute choice differs in some critical aspects from risky choice. Perhaps most importantly, the attributes in risky choice tasks—that is, outcomes and probabilities—are not independent from each other. Instead, probabilities are associated with specific outcomes and they are often assumed to serve as weights for these outcomes. This is also reflected in people's information search (e.g., Rosen & Rosenkoetter, 1976). In Yang and Krajbich's model this issue is bypassed by conceptualizing attribute attention in risky choice as attention to bundles of probabilities and outcomes—that is, each outcome and its associated probability is considered an “attribute.”[4] While this approach makes the model applicable to both multiattribute choice and risky choice, it does not allow the model to accommodate our finding that attribute attention, defined as the distribution of attention across probabilities and outcomes, shapes preferences. It may be possible to further refine the model to accommodate these intricacies of the attention–preference link in risky choice, but this would probably come at the cost of losing generality for multiattribute choice scenarios.

Finally, Glickman et al. (2019) proposed a model that allows for a confluence of biases in attribute attention and option attention in risky choice (they referred to it as “hybrid” model). The model treats outcomes and probabilities as independent attributes, and evidence regarding these attributes is integrated in an additive fashion. Biases in attribute attention influence the relative weight given to outcomes and probabilities. In addition, the evidence for each option can be modulated by biases in option attention, similar to the implementation in the aDDM. While this is a promising approach, the model was tested using relatively simple choice problems, in which each option consisted of one outcome and one probability. Whether the results generalize to more complex problems with several outcomes and probabilities per option, and whether the model can account for the association between patterns in attention allocation and differences in probability weighting demonstrated here, needs to be tested in future research.

6.2. Generalizing the link between attention and probability weighting

The results of our current analyses expand the scope of our (Zilker & Pachur, 2022c) earlier insights regarding the impact of option attention on probability weighting in several ways. First, in our previous analyses, the effect of option attention on probability weighting was demonstrated in experience-based decision making, using sampling behavior as an indicator of attention, and in description-based decision making, with attention measured based on eye tracking (Zilker & Pachur, 2022c). Our current analysis shows that key conclusions from these analyses generalize to risky choice paradigms in which information search is measured using Mouselab, an information-board procedure.

Second, whereas our (Zilker & Pachur, 2022c) previous empirical analyses relied on choice problems from the domain of gains only, the current work extends the evidence for effects of option attention on probability weighting to the domain of losses. Importantly, we find that as in the domain of gains, also in the domain of losses more attention paid to an option increased the tendency to choose this option. Note that in order to capture that the effect of attention on preference points in the same direction in the gain and loss domains, the direction of the effect of option attention on probability weighting needs to reverse between the domains.

The finding that attending more to an option in the domain of losses increases the tendency to choose this option is also theoretically illuminating. Models such as the aDDM (Krajbich et al., 2010) posit that attending more to an aversive option implies the accumulation of more negative evidence, that is, evidence against this option, and should therefore lead to a decreased tendency to choose it. Our findings are at odds with this prediction, suggesting that a prominent model of the interplay between attention and preference may not generalize seamlessly across the gain and loss domains. Because most studies testing the aDDM (and variants thereof) have involved positively valued stimuli only (but see Armel, Beaumel, & Rangel, 2008; Fisher, 2021b), this issue may thus far have been overlooked.

Moreover, our current and prior (Zilker & Pachur, 2022c) analyses suggest that the effect of option attention in risky choice generalizes across decisions from experience and decisions from description. This seems remarkable: after all, attention plays, to some extent, different roles in experience-based and description-based settings. In descriptive settings, attention operates on fully characterized options and additional inspections of an option do not alter the payoff distribution of the option as perceived by the decision maker. In experiential settings—where attention is operationalized based on the number of samples drawn from the option—attending more to an option (i.e., sampling from it more) implies that the decision maker can gather additional, previously unavailable information about the option and thus reduce ambiguity about its payoff distribution. The finding that attention affects preferences in descriptive and experiential settings alike may suggest that—as is assumed but hardly tested in evidence accumulation models of decision making—gathering information from descriptive summaries of a risky option bears similarities to sequential sampling from a payoff distribution (e.g., Busemeyer & Townsend, 1993).

Finally, our analyses refine the results obtained by Pachur et al. (2018) regarding the effects of attribute attention on probability weighting. Recall that when attribute attention was balanced (i.e., given $a^{attribute} = 0.5$) probability weighting was approximately linear (see Fig. 7). That is, when participants allocated their attention evenly across probabilities and outcomes they tended to weight events more objectively, whereas the probability-weighting function became more strongly curved when attention was more biased to one of the attributes. This result diverts somewhat from that of the analyses by Pachur et al. (2018). Specifically, Pachur et al. (2018) found that more attention allocated to probabilities was linked to more linear probability weighting, such that for participants who paid most attention to probabilities γ was closest to 1. Our results confirm that probability weighting tended to be more inverse S-shaped ($\gamma < 1$) when people

[4] In the data analyzed by Yang and Krajbich (2022) based on a task by Smith and Krajbich (2018) information on probabilities of the outcomes was not explicitly displayed onscreen during the choice (all probabilities equaled 0.5) and there was therefore no way to measure “attention to probabilities.”

attended very little to probabilities, and that attending relatively more to probabilities was linked to more linear probability weighting (values of γ closer to 1) in such cases. Crucially, our analyses add the insight that probability weighting tended to become more biased in the opposite direction ($\gamma > 1$) when participants excessively focused on probabilities. That is, the current results suggest that probability weighting was more unbiased when attention allocation was more unbiased. One reason for why this was not evident in the analyses by Pachur et al. (2018) is that they constrained the range of γ between 0 and 1— making it impossible to observe biases in probability weighting in the range $\gamma > 1$ — whereas our current analyses allowed γ to vary between 0 and 2. Notably, balanced attention is associated with (approximately) linear probability weighting both when considering attribute attention (where probability weighting tends to be more linear when attention is neither strongly biased towards outcomes or probabilities) and when considering option attention (where probability weighting tends to be more linear when attention is not strongly biased towards either of the options).

6.3. Implications for attentional interpretations of the probability-weighting function

As mentioned in the introduction of this article, there have been several suggestions how to interpret properties of the probability-weighting function from a psychological point of view. For instance, differences in the curvature of the probability weighting are commonly viewed as reflecting differences in probability sensitivity (Gonzalez & Wu, 1999; Tversky & Kahneman, 1992). Our current analyses complicate this interpretation, since they highlight that parameters of the probability-weighting function can reflect different facets of attention allocation. For instance, when the γ parameter is estimated to be smaller than 1 (yielding an inverse-S-shaped probability-weighting function), this might indicate that the decision maker is insensitive to probability information, or that they attended more to outcomes than to probabilities; it could also indicate that they attended more to the LL option than to the ML option. A combination of these scenarios is also conceivable. Therefore, it is difficult to unequivocally ascribe a given shape of the probability-weighting function to a specific cognitive underpinning. In light of the results of our current analyses, attentional interpretations of estimated probability-weighting parameters should be accompanied and substantiated by external measures of attention, such as Mouselab data, gaze data from eye tracking, or sampling data (see also Zilker & Pachur, 2022c).

In addition to refining insights regarding the psychological interpretation of specific shapes of probability-weighting functions, the current results also have practical implications. Nonlinear probability weighting indicates deviations from the normative standard of objective weighting, formalized in EU theory. To rectify such deviations and bring decisions more in line with this normative standard, knowledge about the psychological processes underlying preferences is essential (see also Payne & Venkatraman, 2011; Weber & Johnson, 2009). Our investigation highlights that interventions targeting nonlinear probability weighting might have to rely on distinct approaches, depending on which attentional process is mainly responsible for the observed deviation from objective weighting in a given case. For instance, if a decision maker displays a bias in option attention and their preferences deviate from objective probability weighting accordingly, then trying to reduce such deviations based on an intervention that targets attribute attention may be relatively inefficient—and vice versa.

6.4. Does attention causally shape probability weighting?

One may wonder about the causal nature of the association between attention allocation and probability weighting. Does attention allocation causally shape probability weighting, or do specific patterns in probability weighting co-occur with specific patterns in attention allocation for other reasons—for instance, due to a common dependence on the decision maker's risk propensity? Since in Experiment 1 by Pachur et al. (2018) attention allocation was not actively manipulated, this question cannot be conclusively answered with the current data. However, other investigations have addressed the question of whether attention can causally shape preferences. For instance, in Experiment 2 by Pachur et al. (2018) the relative presentation time of gains vs. losses was manipulated and this manipulation modulated CPT's loss aversion parameter—speaking towards a causal effect of attention on choice (for a manipulation of attention to outcome information to increase sensitivity to outcomes, see Hirmas & Engelmann, 2022). Also beyond risky choice, various studies have demonstrated that increasing the exposure duration of a choice option (i.e., option attention) can increase the probability that the option is chosen (Armel et al., 2008; Bird, Lauwereyns, & Crawford, 2012; Glaholt & Reingold, 2011; Lim, O'Doherty, & Rangel, 2011; Nittono & Wada, 2009; Shimojo, Simion, Shimojo, & Scheier, 2003).

Not only exposure duration, but also other facets of attention have been manipulated in choice tasks (for an overview see Bhatnagar & Orquin, 2022), with the conclusion that attention can causally shape preference formation. For instance, Molter and Mohr (2021) found that options were more likely to be chosen when they were presented last in a sequence. Since our current analyses relied on aggregate attention scores across entire trials, they do not provide insights as to whether the order with which attributes and options were inspected within each trial modulate the effects of attention on choice. However, it seems conceivable that order effects also contribute to the current findings. For instance, option-specific biases in attention allocation in eye-tracking studies often seem to evolve towards the end of the trial (a phenomenon referred to as gaze cascade; Krajbich et al., 2010; Shimojo et al., 2003), and Pachur et al. (2018) reported a link between final inspection and choice also for their data collected with Mouselab.

Finally, note that there is a perspective on the link between attention allocation and probability weighting under which it would be futile to ask which of the two is cause and which effect. Specifically, attentional biases and probability weighting might be constructs located on different levels of explanation (Marr, 1982). Probability weighting is located on the computational level, since it abstractly characterizes the decision maker's goals in terms of the expectation to be maximized within the framework of expected utility. Attention, by contrast, is located on the algorithmic level, which describes the cognitive processes decision makers apply to to maximize a given expectation. From this perspective, both attention and probability weighting might characterize the same phenomenon, but by invoking different explanatory stances.

6.5. What might drive imbalances in attention allocation across attributes and options?

As shown in Fig. 5, both attribute attention and option attention varied considerably. Although attention was approximately balanced on average, in some trials there was a strong bias towards probability information (i.e., very high values of $a^{attribute}$), whereas in other trials the focus was almost exclusively on outcome information (i.e., very low values of $a^{attribute}$). Similarly, in some trials attention tended to be focused on the ML option (i.e., very high values of a^{option}), whereas in others it tended to be focused on the LL option (i.e., very low values of a^{option}). Given that variability in attribute attention and option attention is linked to differences in preference, an important question is: What factors might drive the variability in attention?

The results reported in subsection "Attention Allocation" indicated that variability in attribute attention was to a considerable extent due to individual differences between decision makers. Which characteristics of decision makers might drive attribute attention? One possible person-specific factor contributing to individual differences in attribute attention is numeracy, a person's ability to comprehend probabilistic information (e.g., Peters, 2020). People with higher numeracy have a better understanding of and ability to operate with information on probabilities, which might lead them to spend more time attending to probabilities. Evidence consistent with this pattern was obtained by Jasper, Bhattacharya, and Corser (2017) and Keller, Kreuzmair, Leins-Hess, and Siegrist (2014).

While our analyses of the Mouselab data found that problem-specific features barely explained variability in attention allocation (Appendix D), other studies using different paradigms and measures of attention have provided evidence that attention can depend on features of a given choice problem. For instance, analyses using eye-tracking (in which all information on options is visible from the trial onset and does not have to be actively revealed as in Mouselab) have shown that higher outcomes and probabilities seem to attract more attention than lower ones do (Fiedler & Glöckner, 2012; Stewart, Hermens, & Matthews, 2016); this may lead to biases in both attribute attention and option attention. The amount of attention that outcomes and probabilities receive also depends on the similarity of the options on a given attribute (Pachur et al., 2013). Moreover, experiments allowing participants to explore options by sampling from their outcome distributions, thus offering them the opportunity to experience the variability of these distributions first hand, have shown that differences between gambles in variance may drive option attention: Gambles with higher variance tend to attract more attention during sampling (Lejarraga, Hertwig, & Gonzalez, 2012; Pachur & Scheibehenne, 2012). This indicates that, depending on the paradigm and measure of attention used, attention allocation may also be driven by problem-specific factors.

7. Conclusion

For a long time, research on people's preferences in risky choice has been primarily concerned with cataloguing the myriad ways in which choice behavior deviates from classical standards of rationality (e.g., expected value and expected utility theory), and with describing such preference patterns using elegant formal constructs such as nonlinear probability weighting. However, it is not clear what the cognitive underpinnings of these phenomena are (see Zilker & Pachur, 2022b). In theoretical and empirical work on value-based choice (e.g., between food items) one of the key breakthroughs has been the insight that preferences are associated with specific patterns in attention allocation during predecisional information search (e.g., Krajbich et al., 2010). Our study contributes to a better understanding of the cognitive processes—specifically, attention—underlying preference patterns in risky choice. In particular, we highlighted the conceptual and empirical distinction between attribute attention and option attention and clarify that these two facets of attention have systematic and largely independent effects on probability weighting. This distinction may be similarly important in other choice tasks in which attention can vary across options and their attributes, such as intertemporal choice (where rewards are associated with delays; Fisher, 2021a) or multiattribute choice. Our analyses may inspire further fine-grained analyses of attentional patterns also in these domains.

CRediT authorship contribution statement

Veronika Zilker: Conceptualization, Data curation, Formal analysis, Visualization, Writing - original draft, Writing - review & editing. **Thorsten Pachur:** Conceptualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

None.

Data availability

Data & code are available at https://doi.org/10.17605/OSF.IO/QHMTW (Zilker and Pachur, 2022a).

Acknowledgements

The authors thank Deb Ain for editing the manuscript. The research was supported by grants of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation): grant PA 1925/2-1 and under Germany's Excellence Strategy – EXC 2002/1 “Science of Intelligence” – project number 390523135.

Appendix A. The link between attention and CPT's value function

Pachur et al. (2018) found links between attention allocation during predecisional information search and cumulative prospect theory's value function: More attention to outcomes was associated with higher values of the α parameter (see Eq. 1).

Fig. A1 displays the parameter estimates capturing the effects of our measures of attribute attention and option attention on α. The parameter estimates of $\beta^{\alpha}_{a^{attribute}}$ show that there was a credible and negative effect of relative attention to probabilities on α in both the gain and the loss domains. Moreover, the trial-specific estimates, displayed in Fig. A2, show that given a balanced allocation of attention between probabilities and outcomes ($a^{attribute} = 0.5$) α approximates 1, that is, an objective treatment of outcomes. The more biased attribute attention is—whether towards outcomes or probabilities—the more α deviates from 1, indicating a more curved value function. Similar to the results for the γ parameter, this conclusion slightly differs from the prior analyses by Pachur et al. (2018), who found that those participants who paid most attention to outcomes had the most linear value functions, that is, values of α closest to 1. The likely reason for this divergence in conclusions is that Pachur et al. constrained the range of α between 0 and 1, whereas in our analyses α could vary between 0 and 2. Allowing for this extended range reveals that when participants paid disproportionally more attention to outcomes ($a^{attribute} < 0.5$), their parameter α tended to exceed 1, indicating an increasingly curved rather than a linear value function.

Regarding a possible link between option attention and the curvature of the value function, α, the parameter estimates of $\beta^{\alpha}_{a^{option}}$ indicated no credible effect of option a^{option} attention on α in either the gain or loss domain.

Considering the full model, the estimated main effects of attribute attention and option attention replicate the findings obtained in the reduced models. Moreover, there was no credible interaction between attribute and option attention on α, resembling the results regarding probability weighting.

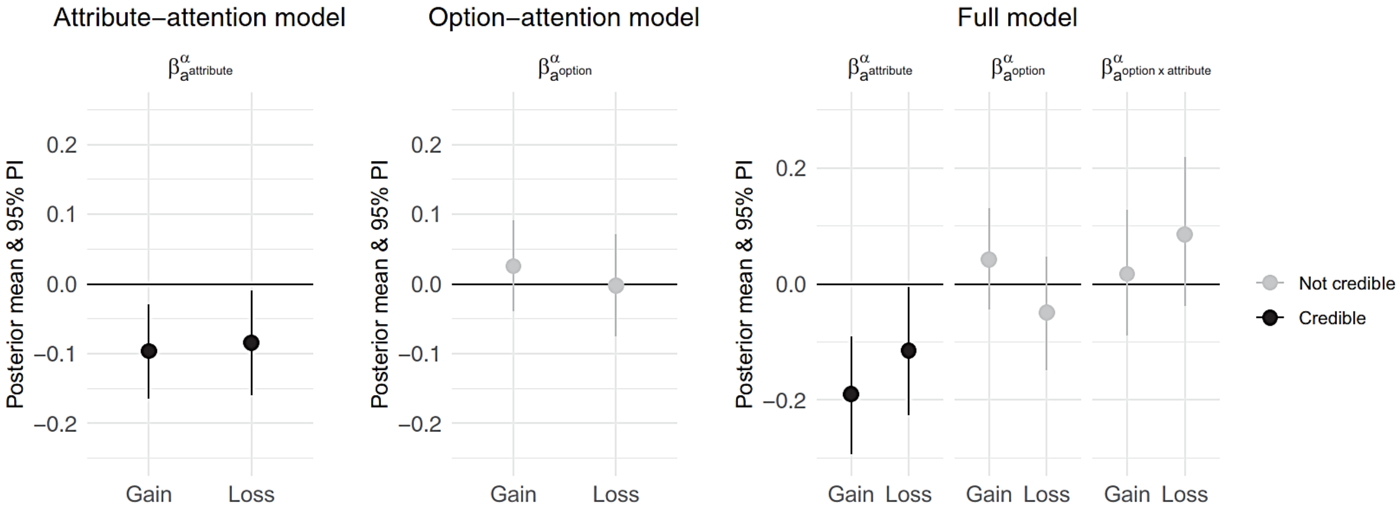

Fig. A1. Results regarding the association between attention and the α parameter of cumulative prospect theory's value function. Shown are the group-level posterior means and 95% posterior intervals (PI) of the regression coefficients.

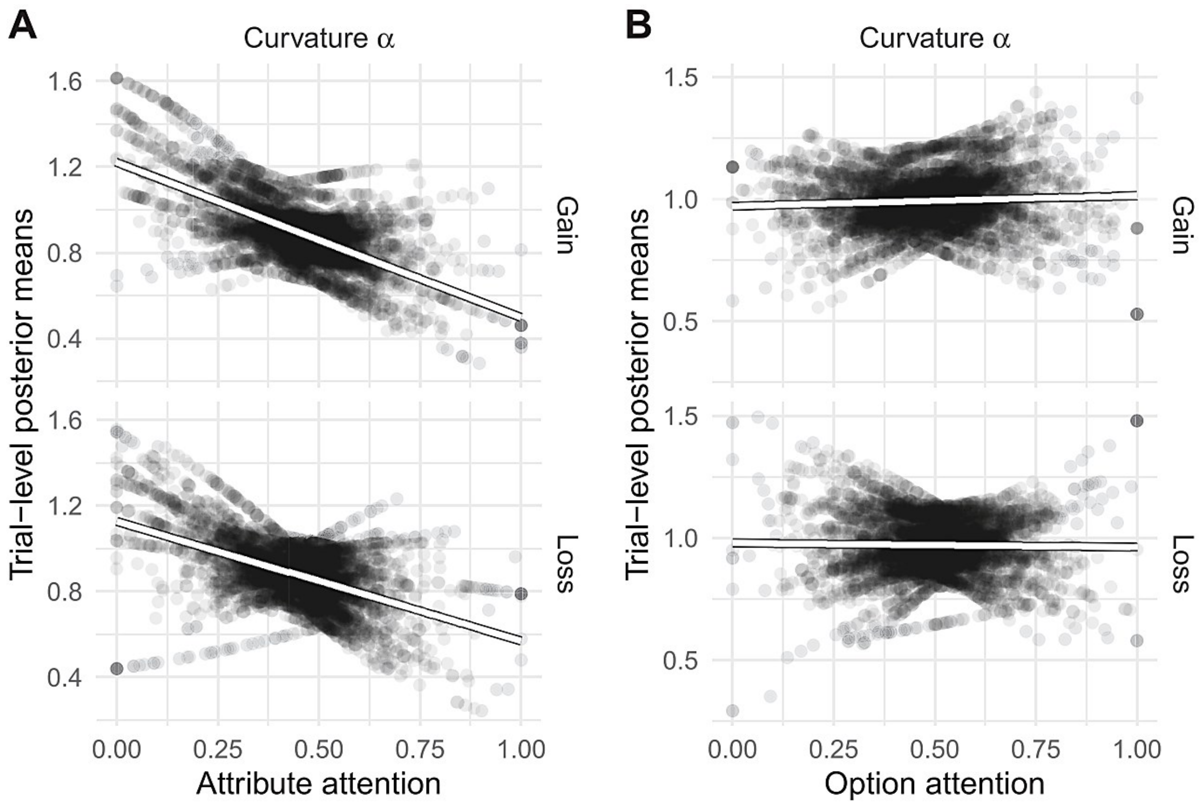

Fig. A2. Association between trial-level parameter estimates of the α parameter of cumulative prospect theory's value function and attention. Panel A: Results regarding attribute attention from the attribute-attention model. Panel B: Results regarding option attention from the option-attention model.

Appendix B. Parameter recovery analysis

To demonstrate that the applied modeling approach allows us to reliably identify the parameters of the regression model capturing the effects of attribute attention and option attention on CPT parameters, a parameter recovery was conducted for the full model used. The recovery relied on the same choice problems and on the same number of participants and trials as the empirical analyses. The empirically observed values of the attentional variables ($a^{attribute}$ and a^{option}) were used as predictors in the model. This ensured that the recovery results are informative regarding the inferences drawn from the specific experiment and analytic strategy applied in the main text.

First, the values of each group-level parameter capturing the effects of attribute and option attention on CPT parameters (e.g., $\beta^{\gamma}_{a_{option}}$, $\beta^{\gamma}_{a_{attribute}}$, $\beta^{\gamma}_{a_{option \times attribute}}$) were systematically varied by drawing 10 evenly spaced values between -0.5 and 0.5. The resulting 10 values for each group-level parameter were randomized and randomly combined with the values of the other group-level parameters. This resulted in 10 combinations of parameter values, for both the domain of gains and the domain of losses. The person-level random effects were sampled from a zero-centered Gaussian with a standard deviation of 0.01. Based on each of these group-level parameter settings and person-level random effects, and based on the values of the attentional covariates on each trial, trial-level parameter settings for each choice problem and participant were derived. These trial-level settings were used to simulate CPT-based choices. The full model was then fitted to the data generated in this way, separately for each parameter setting, and using the same methods applied when modeling the empirical data. Fig. B1 displays the resulting parameter estimates for the group-level parameters capturing the effects of attribute and option attention, plotted against the generative parameter values that were used in the simulation. As can be seen, differences in the true parameter values were recovered very well. The recovered posterior means are overall very close to the true parameter values, and the 95% posterior intervals include the true value in almost all cases, especially in the ranges of the empirically observed estimates. Only in very few and extreme cases did the recovered parameters over- or underestimate the true values (e.g., for the lowest setting of $\beta^{\delta}_{a^{option}}$ in the gain domain). Likewise, only in very few cases was the recovery of the parameters rather imprecise, indicated by wide 95% posterior intervals. Overall, these results suggest that our modeling approach is well suited to capture the effects of attention allocation on CPT parameters.

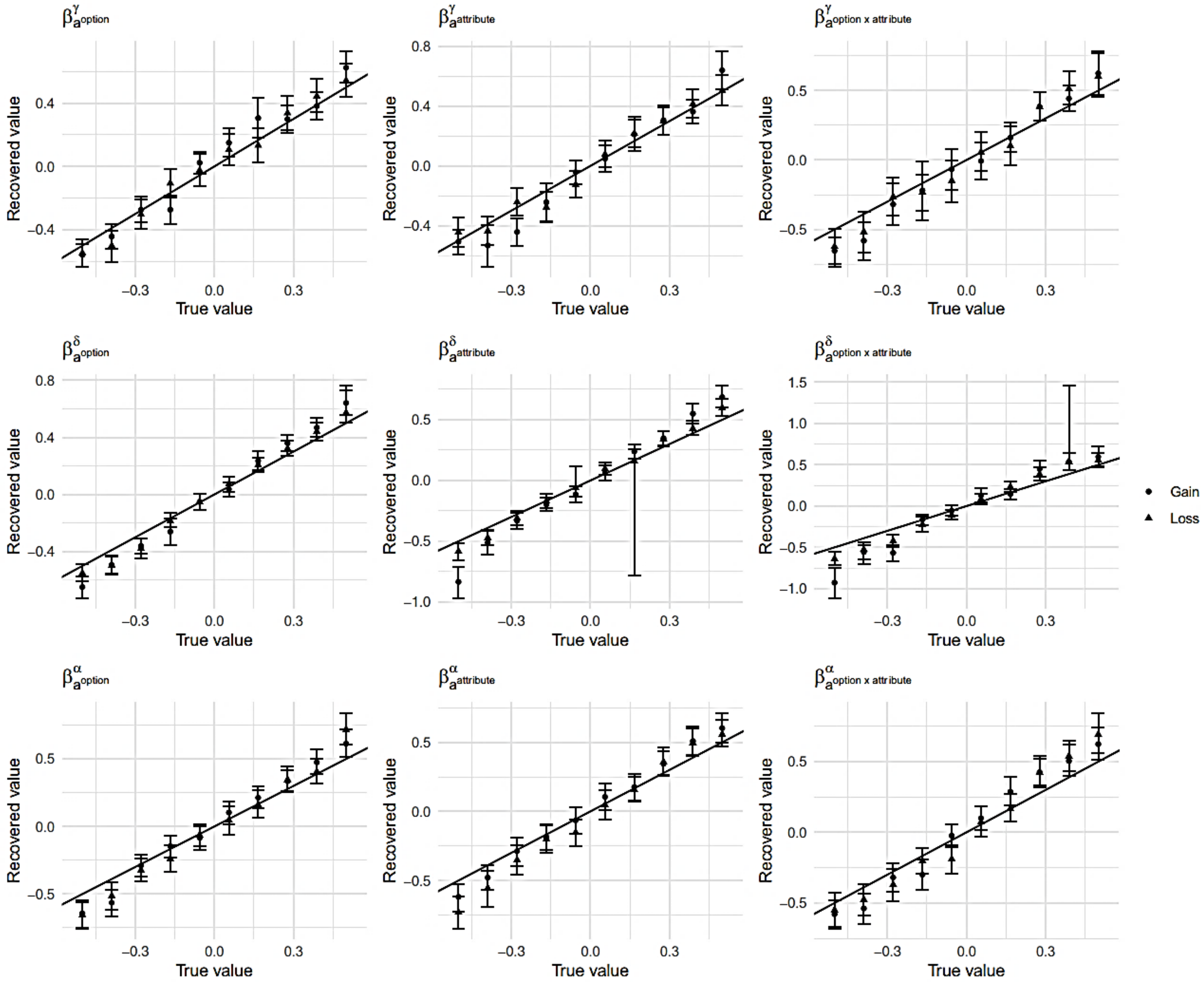

Fig. B1. Results of the parameter recovery analysis.

Appendix C. Association between total inspection time and strength of attentional bias

In the analyses reported in the main text, we used relative measures of attention allocation, that normalized the time people spent attending to individual attributes or options by the total inspection time attending to information on each trial. Here we examine to what extent relative measures, expressing the degree to which attention is imbalanced across attributes or options, covary with the overall inspection time. For instance, people who spend more time on the problem may allocate their attention in a more balanced manner (captured in relative attention scores) across the available attributes and options. To address this issue, we calculated an index for the strength of attribute-specific attentional bias, $b^{attribute} = | \ a^{attribute} - 0.5|$, and an index for the strength of option-specific attentional bias, $b^{option} = | \ a^{option} - 0.5|$. Both bias indices range between 0 and 0.5; they equal zero when both attributes (options) are attended to for the exact same amount of time (indicating no attentional bias), and they equal 0.5 when one attribute (option), regardless which one, is attended to exclusively (indicating an extreme attentional bias). Fig. C1 plots the trial-wise values of $b^{attribute}$ and b^{option} against the log-transformed total time people spent inspecting information before they made a choice, separately for the gain and loss domains.

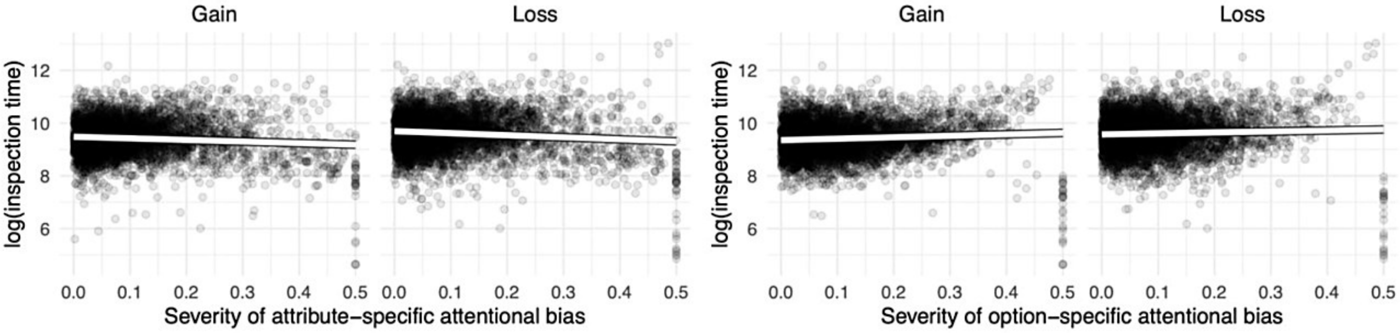

Fig. C1. Association between how balanced attention was between attributes and options and the log-transformed total inspection time.

To examine these associations statistically, we estimated GLMs with the log-transformed total inspection time as the dependent variable and $b^{attribute}$ as the predictor variable, separately for the gain and the loss domains. Analogous models were estimated using b^{option} as the predictor. There was a slight negative association between $b^{attribute}$ and total inspection time in both domains (gains: $\beta_{b^{attribute}} = -0.606, PI_{95\%}$ $[-0.813, -0.395]$; losses: $\beta_{b^{attribute}} = -0.779, PI_{95\%}$ $[-0.987, -0.572]$), indicating that people allocated their attention more evenly across outcomes and probabilities when they inspected information for a longer time on a problem. Moreover, there was a slight positive association between b^{option} and total inspection time in both domains (gains: $\beta_{b^{option}} = 0.549, PI_{95\%}$ $[0.337, 0.765]$; losses: $\beta_{b^{option}} = 0.372, PI_{95\%}$ $[0.11, 0.620]$), indicating that people showed stronger option-specific attentional biases when they inspected information for a longer time on a given problem.

Appendix D. Is variability in option attention driven by features of the options?

It is conceivable that our analysis reported in the section "Attention Allocation" in the main text did not indicate a substantial contribution of problem-specific effects to attention because we did not code for specific features of options (e.g., the magnitude of outcomes and probabilities) but instead assigned a unique identifier to each problem. To test whether a higher proportion of variability in attention is explained by specific features of choice problems, we conducted analyses using as predictors the difference in expected value (EV) between the ML and the LL options, the most extreme outcomes of the ML and the LL options, and the maximum probabilities of the ML and the LL options (rather than unique identifiers for each choice problem). In the gain (loss) domain, EV differences accounted for 0.004% (0.076%) of variance of attribute attention, the most extreme outcome of the ML option for 0.099% (0.002%), the most extreme outcome of the LL option for 0.033% (0.019%), the highest probability of the ML option for 0.461% (0.008%), and the highest probability of the LL option accounted for 0.004% (0.150%) of the variance in attribute attention. In the gain (loss) domain, EV differences accounted for 0.8% (0.03%) of the variance in option attention, the most extreme outcome of the ML option for 0.02% (0.48%), the most extreme outcome of the LL option for 0.001% (0.157%), the highest probability of the ML option for 0.06% (0.347%), and the highest probability of the LL option accounted for 0.095% (0.018%) of the variance. These very low proportions highlight that the low level of variance explained by differences between choice problems in the analyses reported in the main text are not merely due to our use of unique identifiers for choice problems as predictors.

Appendix E. Analysis of the temporal stability between experimental sessions

In Experiment 1 by Pachur et al. (2018) each participant took part in two sessions, conducted about three weeks apart. As reported by Pachur et al. (2018), individual participants' CPT parameters were clearly correlated across both sessions (with Spearman correlations for the parameters ranging between $r_s = 0.38$ and $r_s = 0.67$), indicating that individual differences in preference were relatively stable across time. The same was true for the attention indices. Here the correlations ranged between $r_s = 0.36$ and $r_s = 0.83$. The authors also showed that differences in attention between sessions predicted individual differences in the CPT parameters α and δ, but not γ.

In our current analyses reported in the main text, data from both sessions were analyzed together. Here we report additional analyses in which we split the data by session in order to test to what extent the links between probability weighting and attention allocation were stable across the two experimental sessions. We estimated the full model for the data from both sessions individually, otherwise applying the same methods as in the analyses reported in the main text. Fig. E1 displays the estimates of the group-level parameters for the effects of attribute attention and option attention on the curvature γ and the elevation δ obtained by modeling data from each session individually, as well as the corresponding estimates obtained in the analyses when both sessions are combined.

First, consider the results regarding the curvature parameter γ (Fig. E1A). In all but one cases, the coefficients that were credible when analyzing all data combined are also credible and point in the same direction when analyzing each session separately. The only exception is the effect of option attention on γ in the loss domain, which was not credible in Session 1, but credible and negative in Session 2. The overall effect across both sessions was credible and slightly less pronounced than the effect in Session 2. Across both sessions, the interaction between attribute attention and option attention was not credible. Overall, the results indicate that the effects of attention on γ replicate across the two sessions.

Next, consider the results regarding the elevation parameter δ (Fig. E1B). Both effects that are credible when analyzing both sessions together—namely $\beta^{\delta}_{a^{attribute}}$ in the domain of gains and $\beta^{\delta}_{a^{option}}$ in the domain of losses—are credible in one of the individual sessions, but not the other one. The remaining effects (which are not credible when analyzing both sessions together) are also not credible in the separate analyses. The effects of attention on the δ parameter are thus not only weaker overall, but also slightly less consistent across the two sessions than the effects of attention on the γ parameter.

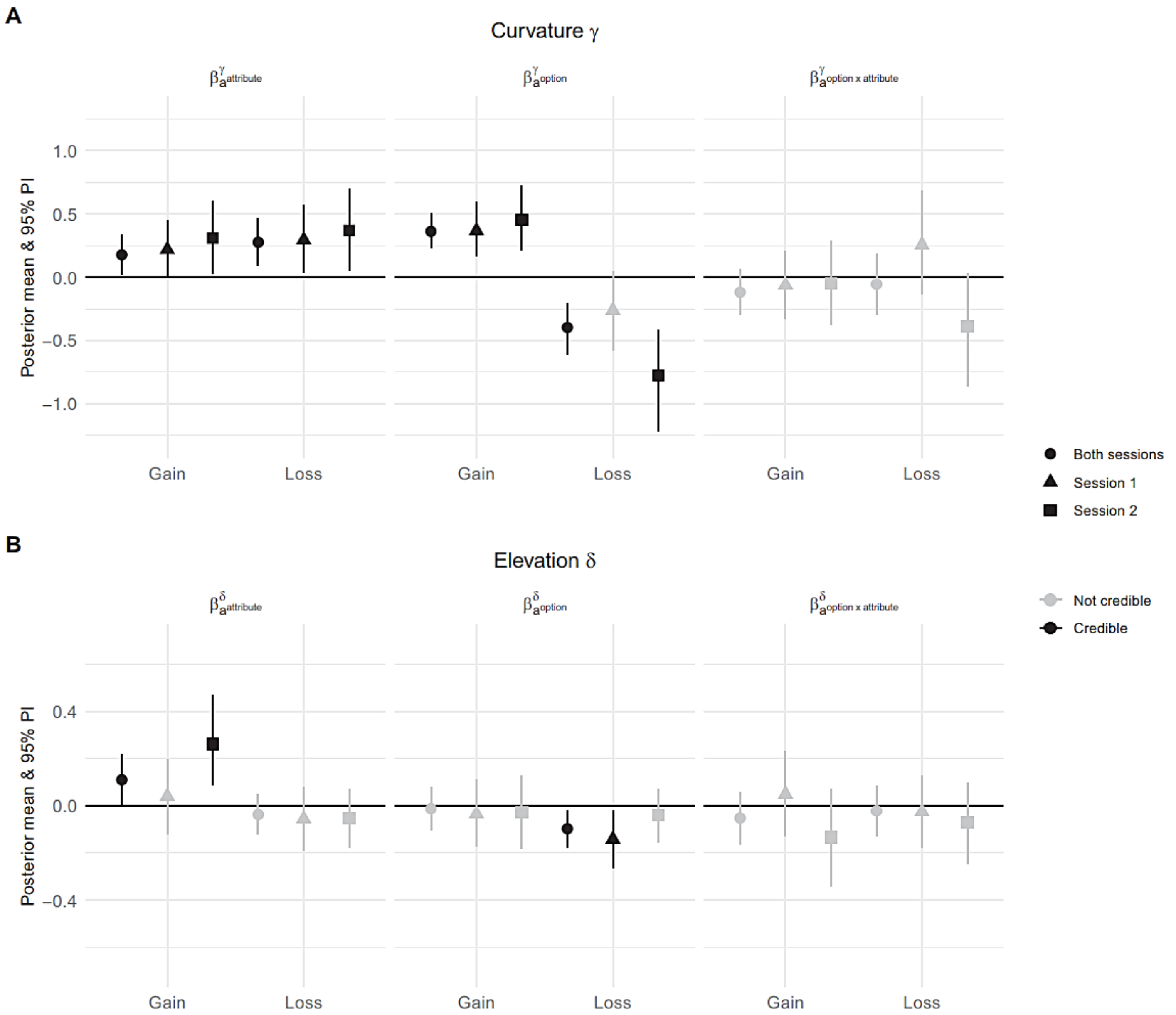

Fig. E1. Results of the full model regarding the association between probability weighting and attention, for each session individually and across both experimental sessions. Shown are the group-level posterior means and 95% posterior intervals (PI) of the regression coefficients for the effects of the attentional variables on the parameters of cumulative prospect theory's probability-weighting function A) γ parameter and B) δ parameter. Attribute attention is defined as the relative amount of attention to probability (vs. outcome) information; option attention is defined as the relative amount of attention to the ML (vs. the LL) option.

Appendix F. Analyses using the probability-weighting function by Goldstein and Einhorn (1987)

In the analyses reported in the main text, we used the two-parameter probability-weighting function by Prelec (1998). Several other functional forms of probability-weighting functions have been proposed in the literature. Here we test to what extent the link between attribute attention and probability weighting and the link between option attention and probability weighting also hold when using the two-parameter weighting function proposed by Goldstein and Einhorn (1987):

$$w(p) = \frac{\delta p^{\gamma}}{\delta p^{\gamma} + (1-p)^{\gamma}} \tag{8}$$

Similar to the two-parameter variant of Prelec (1998), this probability-weighting function is shaped by the parameters $\gamma \in [0,2]$ and $\delta \geq 0$. The parameter γ determines the curvature of the probability-weighting function. When $\gamma < 1$, the probability-weighting function is inverse S-shaped (or convex), implying an overweighting of rare events and an underweighting of events with medium to large probabilities. When $\gamma > 1$, the probability-weighting function is S-shaped (or concave), implying an underweighting of small-probability events and an overweighting of events with medium to large probabilities. The parameter δ governs the elevation of the probability-weighting function. However, it is important to note that in Prelec's (1998) probability-weighting function, higher values of δ imply a lower elevation, whereas in Goldstein and Einhorn's (1987) weighting function, higher values of δ imply a lower elevation. Given parameter values of $\gamma = 1$ and $\delta = 1$ the weighting function is linear, implying weighting by objective probabilities.

We applied the same methods as for the computational modeling reported in the main text, with the only difference being that we implemented CPT using the probability-weighting function by Goldstein and Einhorn (1987). Fig. F1 displays the posterior group-level means and 95% posterior intervals (PIs) for the regression coefficients obtained with the attribute-attention model (Fig. F1A) and the option-attention model (Fig. F1B), analogous to Fig. 3 in the main text.

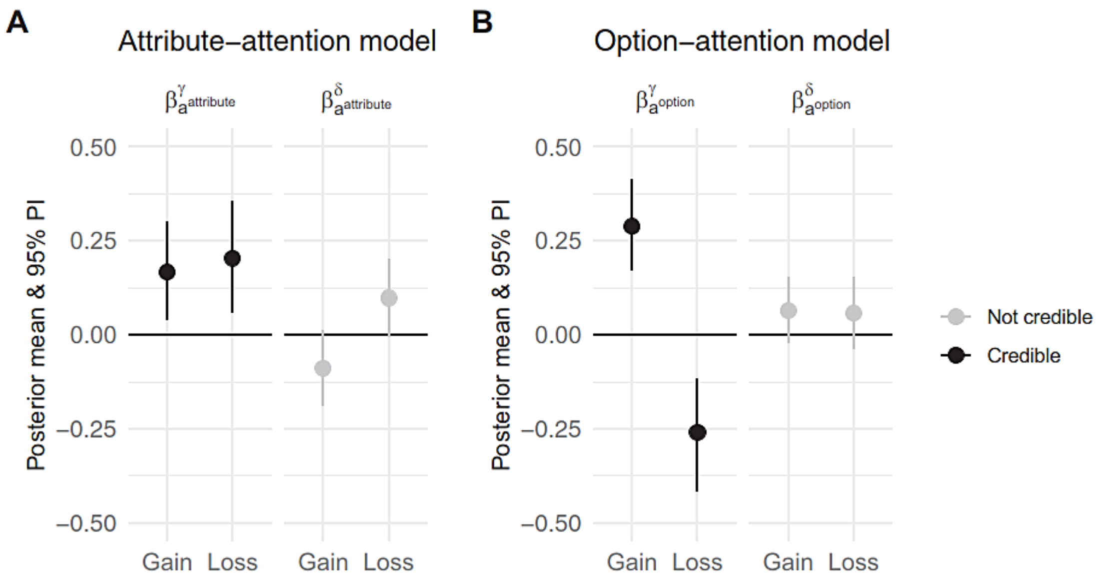

Fig. F1. Results of the attribute-attention model (Panel A) and the option-attention model (Panel B) regarding the association between probability weighting and attention. Shown are the group-level posterior means and 95% posterior intervals (PI) of the regression coefficients for the effects of the attentional variables on the γ and δ parameters of cumulative prospect theory's probability-weighting function, separately for each model. Results are displayed for the analyses using the probability-weighting function by Goldstein and Einhorn (1987). Attribute attention is defined as the relative amount of attention to probability (vs. outcome) information; option attention is defined as the relative amount of attention to the ML (vs. the LL) option.

F.1. Model performance

The attribute-attention model correctly captured 72% of participants' choices, the option-attention model 72.4%, and the full model 74.6%. That is, also when using Goldstein and Einhorn's (1987) probability-weighting function, accounting for both attribute and option attention allows CPT to better accommodate the data than accounting for only one of the attentional variables. However, the proportion of choices captured by the full model was slightly lower than in the analyses using Prelec's (1998) weighting function. As in the analyses using Prelec's (1998) probability-weighting function, also when using Goldstein and Einhorn's (1987) variant the attribute-attention model had the lowest deviance information criterion (DIC = 8,892), closely followed by the option-attention model (DIC = 8,918) and the more complex full model (DIC = 9,722).

F.2. Probability weighting and attribute attention

Focusing on the attribute-attention model, Fig. F1A shows that the posterior mean of the coefficient $\beta^{\gamma}_{a_{attribute}}$ is positive for the domains of both gains and losses. These effects are credible, indicated by the 95% posterior intervals not including zero. Paying more attention to probabilities (relative to outcomes) is thus associated with higher values on the γ parameter. Regarding the elevation of the probability-weighting function, the regression coefficient $\beta^{\delta}_{a_{attribute}}$ indicates that there were no credible effects of attribute attention in either the gain or the loss domain. These results show that the findings of the analyses reported in the main text generalize to an implementation of CPT with an alternative probability-weighting function.

F.3. Probability weighting and option attention

Focusing on the option-attention model, Fig. F1B shows that the posterior mean of the coefficient $\beta^{\gamma}_{a_{option}}$ is positive and credible in the domain of gains and negative and credible in the domain of losses. That is, paying more attention to the ML option is linked to higher values of γ in the gain domain and lower values of γ in the loss domain.

What about a possible link between option attention and the elevation of the probability-weighting function, that is, the δ parameter? Based on the posterior estimates of $\beta^{\delta}_{a_{option}}$ in the option-attention model, option attention was not credibly associated with δ in the gain domain. This echoes the corresponding finding of the analyses reported in the main text. The only slight difference occurred for the link between option attention and δ in the loss domain. Whereas in the analyses using Prelec's (1998) probability-weighting function there was a very weak negative association between option attention and δ in the loss domain, the two are unrelated when using Goldstein and Einhorn's (1987) probability-weighting function. However, note that when using Prelec's (1998) probability-weighting function, the upper bound of the 95% PI of the effect was extremely close to 0.

F.4. Do attribute attention and option attention have separate effects on probability weighting?

We next consider the full model, in which both attribute attention and option attention as well as their interaction are included as predictors of the parameters of CPT's probability-weighting function. Fig. F2 shows that the effects of attribute attention and option attention on γ were credible in both the gain and the loss domain, and pointed in the same direction as the corresponding effects in the reduced models. These results again echo the findings using Prelec's (1998) probability-weighting function. Moreover, there were again no credible interactive effects of attribute attention and option attention on the curvature parameter γ.

Fig. F2 also shows that in the full model, option attention had a credible but very small effect on δ in the gain domain, and no credible effect on δ in the loss domain. When using Prelec's (1998) probability-weighting function in the full model, option attention had no credible effect on δ in the gain domain, and a small but credible negative effect in the loss domain. That is, although the effects of option attention on δ differ slightly from those obtained using Prelec's (1998) probability-weighting function, these deviations concern very small effects. As in the analyses using Prelec's probability-weighting function, there were again no credible interaction between attribute attention and option attention on the elevation parameter δ.

Overall, the findings show that key conclusions from the analyses presented in the main text hold across alternative formalizations of the probability-weighting function.

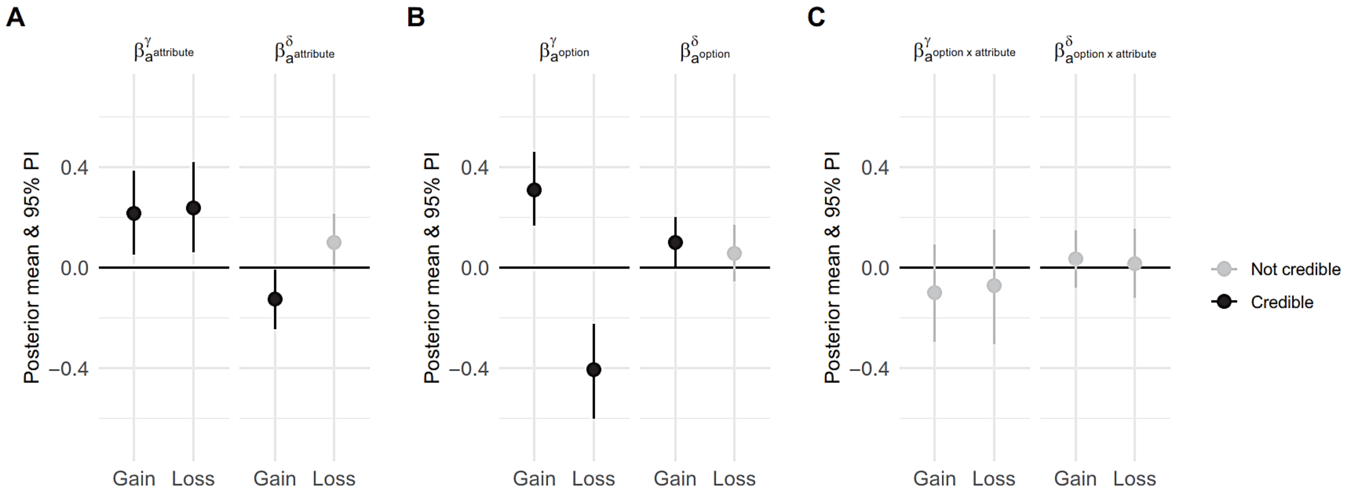

Fig. F2. Results of the full model regarding the association between probability weighting and attention. Shown are the group-level posterior means and 95% posterior intervals (PI) of the regression coefficients for (A) attribute attention, (B) option attention, and (C) their interaction on the γ and δ parameters of cumulative prospect theory's probability-weighting function. Results are displayed for the analyses using the probability-weighting function by Goldstein and Einhorn (1987). Attribute attention is defined as relative amount of attention to probability (vs. outcome) information; option attention is defined as the relative amount of attention to the ML (vs. the LL) option.

References

Allais, M. (1953). L'extension des théories de l'équilibre économique général et du rendement social au cas du risque [Extension of the theories of general economic equilibrium and social output to the case of risk]. *Econometrica, 21*(2), 269–290. https://doi.org/10.2307/1905539

Armel, K. C., Beaumel, A., & Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision making, 3*(5), 396–403.

Bhatia, S. (2014). Sequential sampling and paradoxes of risky choice. *Psychonomic Bulletin & Review, 21*(5), 1095–1111. https://doi.org/10.3758/s13423-014-0650-1

Bhatnagar, R., & Orquin, J. L. (2022). A meta-analysis on the effect of visual attention on choice. *Journal of Experimental Psychology: General, 151*(10), 2265–2283.

Bird, G. D., Lauwereyns, J., & Crawford, M. T. (2012). The role of eye movements in decision making and the prospect of exposure effects. *Vision Research, 60*, 16–21. https://doi.org/10.1016/j.visres.2012.02.014

Birnbaum, M. H. (2008). New paradoxes of risky decision making. *Psychological Review, 115*(2), 463–501. https://doi.org/10.1037/0033-295X.115.2.463

Boehm, U., Marsman, M., Matzke, D., & Wagenmakers, E.-J. (2018). On the importance of avoiding shortcuts in applying cognitive models to hierarchical data. *Behavior Research Methods, 50*(4), 1614–1631. https://doi.org/10.3758/s13428-018-1054-3

Boehm, U., Steingroever, H., & Wagenmakers, E.-J. (2018). Using Bayesian regression to test hypotheses about relationships between parameters and covariates in cognitive models. *Behavior Research Methods, 50*(3), 1248–1269.

Bruhin, A., Fehr-Duda, H., & Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica, 78*(4), 1375–1412. https://doi.org/10.3982/ECTA7139

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100* (3), 432–459. https://doi.org/10.1037/0033-295X.100.3.432

Camerer, C. F. (2000). Prospect theory in the wild: Evidence from the field. In D. Kahneman, & A. Tversky (Eds.), *Choices, values, and frames* (pp. 288–300). Cambridge University Press.

Cavagnaro, D. R., Pitt, M. A., Gonzalez, R., & Myung, J. I. (2013). Discriminating among probability weighting functions using adaptive design optimization. *Journal of Risk and Uncertainty, 47*(3), 255–289. https://doi.org/10.1007/s11166-013-9179-3

Coombs, C. H., Dawes, R. M., & Tversky, A. (1970). *Mathematical psychology: An elementary introduction*. Prentice-Hall.

Fiedler, S., & Glöckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in Psychology, 3*, 335. https://doi.org/10.3389/fpsyg.2012.00335

Fisher, G. (2017). An attentional drift diffusion model over binary-attribute choice. *Cognition, 168*, 34–45. https://doi.org/10.1016/j.cognition.2017.06.007

Fisher, G. (2021a). Intertemporal choices are causally influenced by fluctuations in visual attention. *Management Science, 67*(8), 4961–4981. https://doi.org/10.1287/mnsc.2020.3732

Fisher, G. (2021b). A multiattribute attentional drift diffusion model. *Organizational Behavior and Human Decision Processes, 165*, 167–182. https://doi.org/10.1016/j.obhdp.2021.04.004

Friedman, M., & Savage, L. J. (1948). The utility analysis of choices involving risk. *Journal of Political Economy, 56*(4), 279–304.

Gelman, A., Goodrich, B., Gabry, J., & Vehtari, A. (2019). R-squared for Bayesian regression models. *The American Statistician, 73*(3), 307–309.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science, 7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

Glaholt, M. G., & Reingold, E. M. (2011). Eye movement monitoring as a process tracing methodology in decision making research. *Journal of Neuroscience, Psychology, and Economics, 4*(2), 125–146. https://doi.org/10.1037/a0020692

Glickman, M., Sharoni, O., Levy, D. J., Niebur, E., Stuphorn, V., & Usher, M. (2019). The formation of preference in risky choice. *PLoS Computational Biology, 15*(8), Article e1007201. https://doi.org/10.1371/journal.pcbi.1007201

Glöckner, A., & Pachur, T. (2012). Cognitive models of risky choice: Parameter stability and predictive accuracy of prospect theory. *Cognition, 123*(1), 21–32. https://doi.org/10.1016/j.cognition.2011.12.002

Gluth, S., Kern, N., Kortmann, M., & Vitali, C. L. (2020). Value-based attention but not divisive normalization influences decisions with multiple alternatives. *Nature Human Behaviour, 4*(6), 634–645. https://doi.org/10.1038/s41562-020-0822-0

Goldstein, W. M., & Einhorn, H. J. (1987). Expression theory and the preference reversal phenomena. *Psychological Review, 94*(2), 236–254. https://doi.org/10.1037/0033-295X.94.2.236

Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology, 38*(1), 129–166. https://doi.org/10.1006/cogp.1998.0710

Goodrich, B., Gabry, J., Ali, I., Brilleman, S., & Buros Novik, J. (2018). rstanarm: Bayesian applied regression modeling via Stan. [R package version 2.18.2]. https://cran.r-project.org/package=rstanarm

Harrison, G. W., & Swarthout, J. T. (2019). Eye-tracking and economic theories of choice under risk. *Journal of the Economic Science Association, 5*(1), 26–37. https://doi.org/10.1007/s40881-019-00063-3

Hirmas, A., & Engelmann, J. (2022). *Impulsiveness moderates the effects of exogenous attention on the sensitivity to gains and losses in risky lotteries* (Tinbergen Institute Discussion Papers No. 22-046/I). Tinbergen Institute https://ideas.repec.org/p/tin/wpaper/20220046.html

Jasper, J. D., Bhattacharya, C., & Corser, R. (2017). Numeracy predicts more effortful and elaborative search strategies in a complex risky choice context: A process-tracing approach. *Journal of Behavioral Decision Making, 30*(2), 224–235. https://doi.org/10.1002/bdm.1934

Johnson, J. G., & Busemeyer, J. R. (2016). A computational model of the attention process in risky choice. *Decision, 3*(4), 254–280. https://doi.org/10.1037/dec0000050
Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–292. https://doi.org/10.2307/1914185
Kellen, D., Pachur, T., & Hertwig, R. (2016). How (in)variant are subjective representations of described and experienced risk and rewards? *Cognition, 157*, 126–138. https://doi.org/10.1016/j.cognition.2016.08.020
Keller, C., Kreuzmair, C., Leins-Hess, R., & Siegrist, M. (2014). Numeric and graphic risk information processing of high and low numerates in the intuitive and deliberative decision modes: An eye-tracker study. *Judgment and Decision making, 9*(5), 420–432.
Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience, 13*(10), 1292–1298. https://doi.org/10.1038/nn.2635
Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences, 108*(33), 13852–13857. https://doi.org/10.1073/pnas.1101328108
Krefeld-Schwalb, A., Pachur, T., & Scheibehenne, B. (2022). Structural parameter interdependencies in computational models of cognition. *Psychological Review, 129*(2), 313–339. https://doi.org/10.1037/rev0000285
Lejarraga, T., Hertwig, R., & Gonzalez, C. (2012). How choice ecology influences search in decisions from experience. *Cognition, 124*(3), 334–342. https://doi.org/10.1016/j.cognition.2012.06.002
Lim, S.-L., O'Doherty, J. P., & Rangel, A. (2011). The decision value computations in the vmpfc and striatum use a relative value code that is guided by visual attention. *Journal of Neuroscience, 31*(37), 13214–13223. https://doi.org/10.1523/JNEUROSCI.1246-11.2011
Lopes, L. L. (1987). Between hope and fear: The psychology of risk. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 255–295). Elsevier. https://doi.org/10.1016/S0065-2601(08)60416-5
Marr, D. (1982). *Vision. A computational investigation into the human representation and processing of visual information*. W. H. Freeman.
Molter, F., & Mohr, P. N. (2021). Presentation order but not duration affects binary risky choice. https://psyarxiv.com/gcthj/
Mullett, T. L., & Stewart, N. (2016). Implications of visual attention phenomena for models of preferential choice. *Decision, 3*(4), 231–253. https://doi.org/10.1037/dec0000049
Nilsson, H., Rieskamp, J., & Wagenmakers, E.-J. (2011). Hierarchical Bayesian parameter estimation for cumulative prospect theory. *Journal of Mathematical Psychology, 55*(1), 84–93. https://doi.org/10.1016/j.jmp.2010.08.006
Nittono, H., & Wada, Y. (2009). Gaze shifts do not affect preference judgments of graphic patterns. *Perceptual and Motor Skills, 109*(1), 79–94. https://doi.org/10.2466/PMS.109.1.79-94
Orquin, J. L., Lahm, E. S., & Stojić, H. (2021). The visual environment and attention in decision making. *Psychological Bulletin, 147*(6), 597–617. https://doi.org/10.1037/bul0000328
Orquin, J. L., Perkovic, S., & Grunert, K. G. (2018). Visual biases in decision making. *Applied Economic Perspectives and Policy, 40*(4), 523–537. https://doi.org/10.1093/aepp/ppy020
Pachur, T., Hertwig, R., Gigerenzer, G., & Brandstätter, E. (2013). Testing process predictions of models of risky choice: A quantitative model comparison approach. *Frontiers in Psychology, 4*(646), 1–22. https://doi.org/10.3389/fpsyg.2013.00646
Pachur, T., Hertwig, R., & Wolkewitz, R. (2014). The affect gap in risky choice: Affect-rich outcomes attenuate attention to probability information. *Decision, 1*(1), 64–78. https://doi.org/10.1037/dec0000006
Pachur, T., Mata, R., & Hertwig, R. (2017). Who dares, who errs? Disentangling cognitive and motivational roots of age differences in decisions under risk. *Psychological Science, 28*(4), 504–518. https://doi.org/10.1177/0956797616687729
Pachur, T., & Scheibehenne, B. (2012). Constructing preference from experience: The endowment effect reflected in external information search. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(4), 1108–1116. https://doi.org/10.1037/a0027637
Pachur, T., Schulte-Mecklenbeck, M., Murphy, R. O., & Hertwig, R. (2018). Prospect theory reflects selective allocation of attention. *Journal of Experimental Psychology: General, 147*(2), 147–169. https://doi.org/10.1037/xge0000406
Pachur, T., Suter, R. S., & Hertwig, R. (2017). How the twain can meet: Prospect theory and models of heuristics in risky choice. *Cognitive Psychology, 93*, 44–73. https://doi.org/10.1016/j.cogpsych.2017.01.001
Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). The adaptive decision maker. Cambridge University Press. https://doi.org/10.1017/CBO9781139173933
Payne, J. W., & Venkatraman, V. (2011). Opening the black box: Conclusions to A handbook of process tracing methods for decision research. In M. Schulte-Mecklenbeck, A. Kühberger, & R. Ranyard (Eds.), *A handbook of process tracing methods for decision research* (pp. 223–249). Psychology Press.
Peters, E. (2020). *Innumeracy in the wild: Misunderstanding and misusing numbers*. Oxford University Press.
Prelec, D. (1998). The probability weighting function. *Econometrica, 66*(3), 497–527. https://doi.org/10.2307/2998573
Rosen, L. D., & Rosenkoetter, P. (1976). An eye fixation analysis of choice and judgment with multiattribute stimuli. *Memory & Cognition, 4*(6), 747–752. https://doi.org/10.3758/BF03213243
Rouder, J. N., & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin & Review, 12*(4), 573–604. https://doi.org/10.3758/BF03196750
Savage, L. J. (1951). The theory of statistical decision. *Journal of the American Statistical Association, 46*(253), 55–67. https://doi.org/10.2307/2280094
Scheibehenne, B., & Pachur, T. (2015). Using Bayesian hierarchical parameter estimation to assess the generalizability of cognitive models of choice. *Psychonomic Bulletin & Review, 22*(2), 391–407. https://doi.org/10.3758/s13423-014-0684-4
Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience, 6*(12), 1317–1322. https://doi.org/10.1038/nn1150
Simon, H. A. (1978). Rationality as process and as product of thought. *The American Economic Review, 68*(2), 1–16.
Smith, S. M., & Krajbich, I. (2018). Attention and choice across domains. *Journal of Experimental Psychology: General, 147*(12), 1810–1826. https://doi.org/10.1037/xge0000482
Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B: Statistical Methodology, 64*(4), 583–639. https://doi.org/10.1111/1467-9868.00353
Stewart, N., Hermens, F., & Matthews, W. J. (2016). Eye movements in risky choice. *Journal of Behavioral Decision Making, 29*(2–3), 116–136. https://doi.org/10.1002/bdm.1854
Stewart, N., Scheibehenne, B., & Pachur, T. (2018). *Psychological parameters have units: A bug fix for stochastic prospect theory and other decision models.* https://doi.org/10.31234/osf.io/qvgcd
Su, Y.-S., & Yajima, M. (2015). R2jags: Using R to run 'JAGS' [R package version 0.5-7]. http://CRAN.R-project.org/package=R2jags
Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297–323. https://doi.org/10.1007/BF00122574
Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2011). Hierarchical diffusion models for two-choice response times. *Psychological Methods, 16*(1), 44–62. https://doi.org/10.1037/a0021765
Vanunu, Y., Hotaling, J. M., Le Pelley, M. E., & Newell, B. R. (2021). How top-down and bottom-up attention modulate risky choice. *Proceedings of the National Academy of Sciences, 118*(39). https://doi.org/10.1073/pnas.2025646118. Article e2025646118.
Vincent, B. T., & Stewart, N. (2020). The case of muddled units in temporal discounting. *Cognition, 198*. https://doi.org/10.1016/j.cognition.2020.104203. Article 104203.
Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annual Review of Psychology, 60*, 53–85. https://doi.org/10.1146/annurev.psych.60.110707.163633
Wedell, D. H., & Senter, S. M. (1997). Looking and weighting in judgment and choice. *Organizational Behavior and Human Decision Processes, 70*(1), 41–64. https://doi.org/10.1006/obhd.1997.2692
Willemsen, M. C., & Johnson, E. J. (2011). Visiting the decision factory: Observing cognition with MouselabWEB and other information acquisition methods. In M. Schulte-Mecklenbeck, A. Kühberger, & R. Ranyard (Eds.), *A handbook of process tracing methods for decision research* (pp. 21–42). Psychology Press.
Wulff, D. U., Mergenthaler-Canseco, M., & Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description–experience gap. *Psychological Bulletin, 144*(2), 140–176. https://doi.org/10.1037/bul0000115
Yang, X., & Krajbich, I. (2022). A dynamic computational model of gaze and choice in multi-attribute decisions. *Psychological Review*. https://doi.org/10.1037/rev0000350. Advance online publication.
Zilker, V., Hertwig, R., & Pachur, T. (2020). Age differences in risk attitude are shaped by option complexity. *Journal of Experimental Psychology: General, 149*(9), 1644–1683. https://doi.org/10.1037/xge0000741
Zilker, V., & Pachur, T. (2022a). Attribute attention and option attention in risky choice. https://doi.org/10.17605/OSF.IO/QHMTW
Zilker, V., & Pachur, T. (2022b). Toward an attentional turn in research on risky choice. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.953008
Zilker, V., & Pachur, T. (2022c). Nonlinear probability weighting can reflect attentional biases in sequential sampling. *Psychological Review, 129*(5), 949–975. https://doi.org/10.1037/rev0000304