

Presentation order but not duration affects binary risky choice

Felix Molter^{1,2,*} and Peter N. C. Mohr^{1,2}

¹School of Business & Economics, Freie Universität Berlin, Berlin, Germany

²Center for Cognitive Neuroscience, Freie Universität Berlin, Berlin, Germany

*E-mail: felixmolter@gmail.com

September 7, 2021

Abstract

Risky choice behaviour often deviates from the predictions of normative models. The information search process has been suggested as a source of some reported "biases". Specifically, gaze-dependent evidence accumulation models, where unfixated alternatives' signals are discounted, propose a mechanistic account of observed associations between eye movements, choices and response times, with longer fixated alternatives being chosen more frequently. It remains debated, however, whether gaze causally influences the choice process, or rather reflects emerging preferences. Furthermore, other aspects the information search process, like the order in which information is inspected, can be confounded with gaze duration, complicating the identification of their causal influences. In our preregistered study 179 participants made repeated incentivized choices between two sequentially presented risky gambles, allowing the experimental control of presentation duration, order, and format (i.e., alternative-wise or attribute-wise). Across presentation formats, we find evidence against an influence of presentation duration on choice. The order in which participants were shown stimulus information, however, causally affected choices, with alternatives shown last being chosen more frequently. Notably, while gaze-dependent accumulation models generally capture effects of gaze duration, causal effects of stimulus order are only predicted by some models, identifying potential for future theory development.

Introduction

A large body of experimental research demonstrates that human risky choices systematically differ from those predicted by optimal, utility-maximising and context-invariant theories of choice like Expected Utility Theory (e.g., Allais, 1953; Hertwig et al., 2004; Kahneman & Tversky, 2012; Mohr et al., 2017; Molter et al., 2021). Many descriptive theories of risky choice ascribe these departures to perceptual or attentional processes (e.g., Busemeyer & Townsend, 1993; Tversky & Kahneman, 1992; Yechiam & Hochman, 2013), with the idea that the decision maker's capacity to process information is limited, and attention serves to

select and amplify information about the choice setting (e.g., different alternatives' attributes, outcomes, or their probabilities). This selection and weighting of information is then assumed to induce biases in the decision process and ultimately affect the choice itself.

While the specific construct of attention differs between theories and its usefulness as a general explanatory device has been questioned (Hommel et al., 2019), eye movements are often taken as an indicator of which information a decision maker processes during the decision. The most ubiquitous finding in eye tracking studies of decision-making (including risky choices) is that alternatives that are looked at

longer are more likely to be chosen (Ashby et al., 2016; Cavanagh et al., 2014; Fiedler & Glöckner, 2012; Glöckner & Herbold, 2011; Gluth et al., 2020; Gluth et al., 2018; Isham & Geng, 2013; Krajbich et al., 2010; Krajbich & Rangel, 2011; Lopez-Persem et al., 2016; Molter et al., 2021; Sepulveda et al., 2020; Smith & Krajbich, 2018; Stewart, Gächter, et al., 2016; Stewart, Hermens, et al., 2016; Thomas et al., 2021; Thomas et al., 2019). Gaze-dependent accumulation models (Glickman et al., 2019; Krajbich et al., 2010; Krajbich et al., 2012; Krajbich & Rangel, 2011) provide a formal account of this *gaze bias* effect and many other details of the empirical association between gaze and choice. They assume that decisions are made by repeated sampling and accumulation of evidence in favour of each alternative until evidence for one alternative reaches a decision threshold. Crucially, accumulation is assumed to depend on gaze allocation, such that information momentarily outside of the decision maker’s gaze is discounted. Notably, gaze-dependent accumulation can explain departures from rational choice (Gluth et al., 2018; Molter et al., 2021), providing theories positing broad attentional effects on choice with concrete process-based evidence.

There remains, however, the question of causality: It is still debated to what extent eye movements causally influence the choice process (and visual attention has a causal role in choice and choice biases), or if they merely reflect an emerging choice (Mormann & Russo, 2021; Westbrook et al., 2020). Note that the accuracy with which gaze-dependent accumulation models predict choices and process measures does not address the issue of causality. Even though these models are frequently interpreted to assume a directed effect of gaze on choice, they only formalize the association between gaze and the choice process, without any directional assumption of causality. For example, it would still be possible for a third variable to influence both choices and gaze. Identifying the direction of causality ultimately requires experimental manipulation.

This issue is of great interest, as a causal effect of gaze on choice would imply that decisions could be influenced by irrelevant, external factors that affect the decision makers’ gaze.

Prior work has investigated different aspects in which changes to the information search process affect choice behaviour, experimentally controlling the perceptual saliency of attributes or alternatives (Milosavl-

jevic et al., 2012; Weber & Kirsner, 1997), the duration for which alternatives or attributes were seen (Armell et al., 2008; Lim et al., 2011; Liu, Lyu, et al., 2020; Pärnamets et al., 2015; Shimojo et al., 2003; Sui et al., 2020; Tavares et al., 2017), the temporal order in which alternatives are inspected (Liu, Zhou, et al., 2020), or the direction in which decision makers can gather information about alternatives and their attributes (Mittone & Papi, 2020; Reeck et al., 2017):

Gaze-dependent accumulation models of risky choice predict that a longer gaze towards an alternative is associated with a higher probability of choosing it (unless its value is aversive; see Smith & Krajbich, 2019). If the effect of gaze on choice was causal, this would imply that experimentally inducing longer gaze towards one choice alternative should increase its likelihood of being chosen. A recent study found support for this effect in risky choices using a gaze-contingent decision prompt paradigm, where participants are free to inspect choice alternatives, but are prompted to make their decision when their viewing patterns favour a target attribute or alternative (Sui et al., 2020). Convergent findings have been reported in other decision-making domains (Armell et al., 2008; Liu, Lyu, et al., 2020; Pärnamets et al., 2015; Shimojo et al., 2003; Tavares et al., 2017).

Similarly, as predicted by some gaze-dependent accumulation models (Glickman et al., 2019) longer relative gaze towards individual attributes (e.g., outcomes or probabilities) is associated with changes in choice behaviour (Glickman et al., 2019; Kim et al., 2012), and there is initial evidence that the effect is causal, too (Liu, Lyu, et al., 2020). These results are in line with other studies that, while not controlling gaze duration, found that manipulation of attribute salience affected choices accordingly (Weber & Kirsner, 1997).

Apart from the duration for which alternatives and attributes were evaluated, the temporal order in which information is processed has also been associated with choice. In particular, the last fixation before a choice is often directed towards the chosen alternative. Again, however, the causal direction of this association is debated. The gaze cascade theory (Shimojo et al., 2003) argues that this finding results from reciprocal positive feedback loops between valuation and information search processes, such that preferred alternatives are attended more, which in turn increases preference for them. Order effects are also captured by multi-

ple computational theories of decision making, albeit in different ways: In the attentional Drift Diffusion Model (aDDM) (Krajovich et al., 2010; Krajovich & Rangel, 2011) this phenomenon is explained by the fact that fixated alternatives are more likely to cross the threshold, as evidence for other alternatives is discounted. Other than this, however, the model does not assume a special role of the last fixation, or the serial position of any piece of inspected information. Mullett and Stewart (2016) showed that an increasing association between gaze and choice before a decision is made, naturally emerges in some gaze-dependent accumulation models and does not imply preference-driven information search. Nevertheless, Liu, Zhou, et al. (2020) found that manipulating which of two snack food items was shown last before a decision is prompted causally biased choices in favor of this item. Other models (Ashby et al., 2016; Glickman et al., 2019) assume a leaky integration process that weighs information acquired later in the decision more heavily, and thereby predicts recency effects such that later presented information should affect choices more strongly. Empirical evidence for recency effects in risky choices comes from decisions from experience (Hertwig et al., 2004), where participants learn about alternatives’ properties by repeated sampling and related work on value integration in a rapidly presented stream of option outcomes (Tsetsos et al., 2012).

Recent work in multiattribute decisions found that the time point at which information about one attribute dimension is considered in a decision affects choice (Amasino et al., 2019; Maier et al., 2020; Sullivan & Huettel, 2021). In contrast to a simple recency effect, however, these models predict that attributes considered earlier than others exert more influence on the choice process, as this also implies a longer total duration of consideration. These results further highlight the important issue that duration and order effects can be closely related (i.e., last seen alternatives are often also seen longer; and both factors are associated with choice) and need to be distinguished carefully.

Here, we add to this body of research and jointly investigate these multiple potentially causal effects of the information search (presentation duration and order) on choice. We performed a preregistered study using a task that involved incentivized choices between two risky gambles, and was designed to investigate the independent contributions of presentation duration

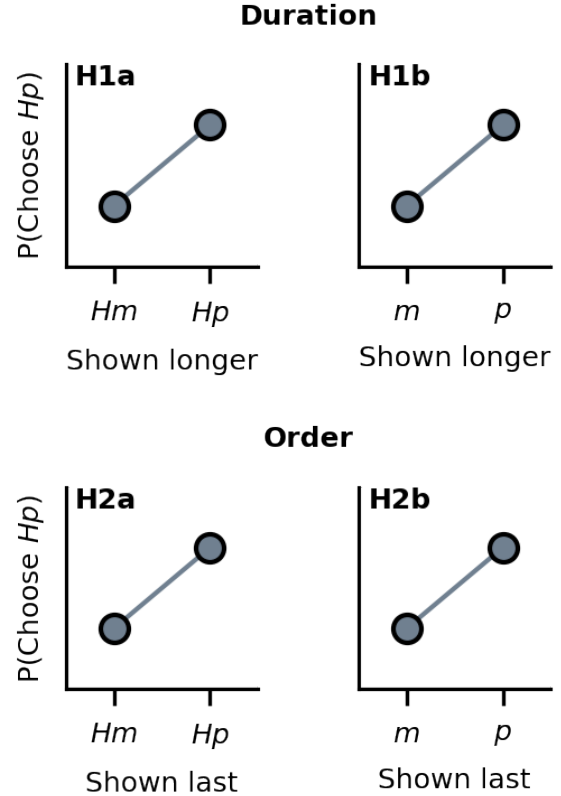


Figure 1. Hypothesized qualitative effects. **H1a:** Higher probability of choosing the alternative with the higher probability (H_p) when it is shown longer. **H1b:** Higher probability of choosing H_p when probabilities are shown longer. **H2a:** Higher probability of choosing H_p when it is shown last. **H2b:** Higher probability of choosing H_p when probabilities are shown last.

and order and their possible interactions in the choice process. Notably, we tested the duration and order effects on two levels, namely, the level of alternatives and the level of attribute dimensions.

Specifically, we hypothesized that presentation *duration* affects choices in two ways: First, in line with gaze-dependent accumulation models, longer presented alternatives should be chosen more frequently (Figure 1 H1a). In addition, we hypothesized that alternatives with higher values on longer presented attributes are chosen more frequently (Figure 1 H1b). Analogously, we hypothesized that presentation *order* affects choices, such that alternatives presented last (Figure 1 H2a), and alternatives with higher values on the attribute presented last (Figure 1 H2b) are chosen more frequently. Our preregistration con-

tained an additional hypothesis and corresponding analysis concerning presentation-dependent changes in value-integration. Details are reported in the [Supplementary Information](#).

In contrast to our predictions and prior literature, our data supports null effects of presentation duration in both attribute- and alternative-wise presentation formats. Instead, we find a causal effect of presentation order on choice, such that alternatives shown last were more likely to be chosen. We discuss the implications of these results.

Results

Behavioural task

In our experiment, 179 participants performed a binary sequential presentation risky choice task (Figure 2), where they made repeated choices between two all-or-nothing risky gambles: One alternative offering a high chance to win a smaller amount (*Hp*) and a second alternative offering a higher amount with a lower probability (*Hm*). Each trial consisted of a presentation phase, where participants learned about the two alternatives' winning probabilities p and outcomes m , and a choice phase. Information was presented across four stages either alternative-wise (Figure 2a) or attribute-wise (Figure 2b). We experimentally controlled presentation duration and order in the task: In each trial one alternative (in alternative-wise presentation) or attribute (in attribute-wise presentation) was shown longer than its competitor (3 vs. 2 seconds across the four stages). In addition, alternatives or attributes could be shown first and third, or second and last in the sequence. In total, participants performed 120 experimental trials based on 15 core choice problems (Figure 2c, d), and 20 additional catch trials with a dominant alternative (Figure 2e). See [Methods](#) for additional details on the behavioural task and stimuli.

Choice behaviour summary

Participants successfully avoided choosing dominated alternatives in catch trials (average \pm s.d. count of dominated choices = 0.35 ± 0.83) with a majority of 140 participants (79%) never choosing a dominated alternative.

In experimental trials, only three participants

Presentation by	Alternatives		Attributes		All
Duration favours	<i>Hm</i>	<i>Hp</i>	<i>Hm</i>	<i>Hp</i>	
Last stage favours					
<i>Hm</i>	0.60	0.60	0.61	0.62	0.61
<i>Hp</i>	0.64	0.63	0.64	0.63	0.63
All	0.62	0.62	0.63	0.62	0.62

Table 1. Mean choice probabilities for the high probability alternative *Hp* across conditions.

(1.7%) exclusively chose the *Hp* (two participants) or *Hm* (one participant) alternatives. All other participants alternated between alternatives to varying degrees, indicating that the selected choice problems covered individual indifference points between high and low probability gambles for a majority of participants (overall mean \pm s.d. $P(\text{choose } Hp) = 0.62 \pm 0.20$; Table 1).

Regression analysis: Effect of presentation order, but not duration

We first analysed choice behaviour using a Bayesian mixed-effects logistic regression model, predicting choice from differences in expected value and effect-coded predictors indicating presentation format, the alternative favoured by presentation duration, the alternative favoured by the last presentation stage, and interaction terms of the duration and order effects with the presentation format (see [Methods](#); Table 2).

Choices were strongly driven by the difference in expected value ($\beta = 1.96$ [1.77, 2.15]) such that alternatives with higher expected values were preferred. In addition, the main effect of presentation order was credibly positive ($\beta = 0.17$ [0.09, 0.25]) indicating a preference towards alternatives favoured by information shown last. No other main or interaction effect was credibly different from zero (presentation duration $\beta = -0.02$ [-0.10, 0.05]; presentation format $\beta = -0.06$ [-0.13, 0.01]; duration-by-format interaction $\beta = 0.01$ [-0.13, 0.16]; order-by-format interaction $\beta = 0.09$ [-0.06, 0.23]).

Next, we addressed our hypotheses regarding effects of presentation duration and order on choice separately, using preregistered directed Bayes factor t -tests and BEST analyses.

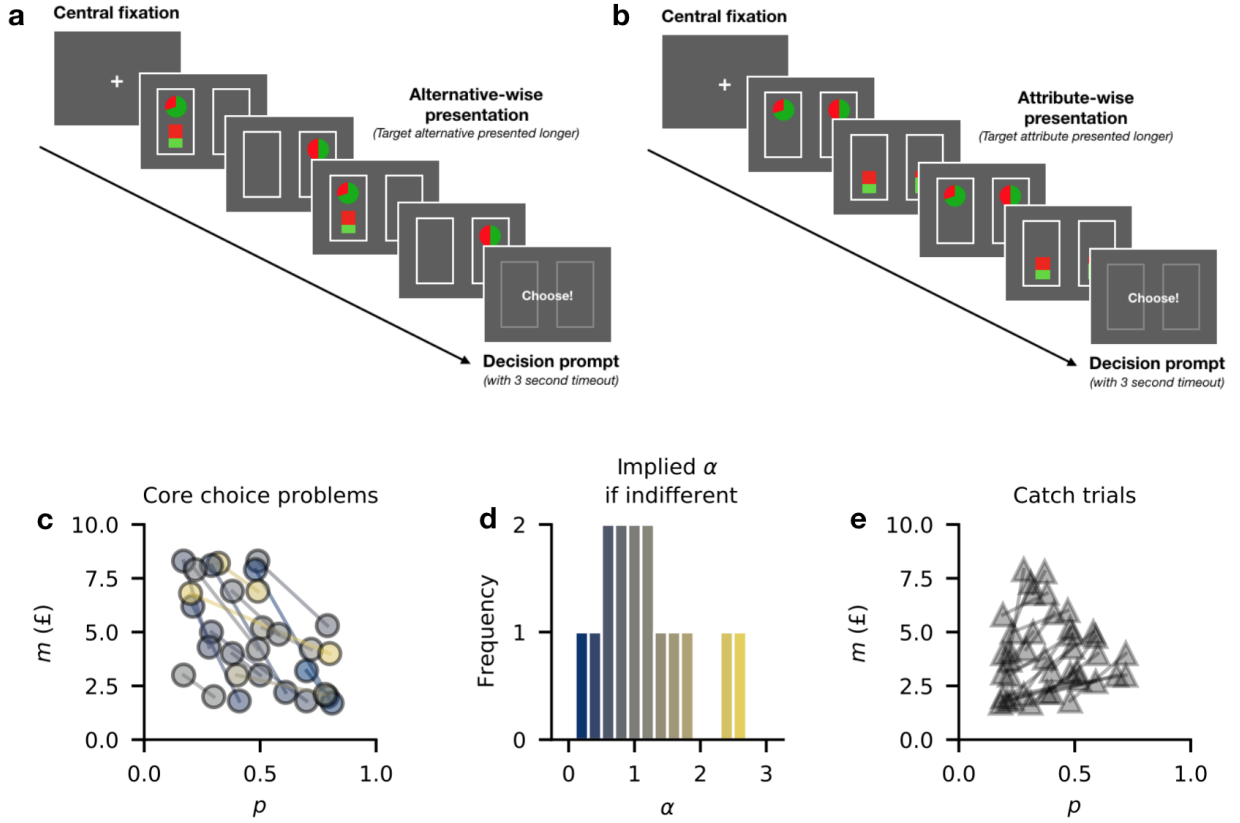


Figure 2. Sequential presentation risky choice task and choice problems. In each trial, participants decided between two all-or-nothing gambles representing the probability p to win an amount m . Information was presented sequentially across four stages so that total presentation duration and presentation order of alternatives (in alternative-wise presentation; **a**) or attributes (in attribute-wise presentation; **b**) was controlled. After the presentation, participants had 3 seconds to indicate their choice. In each trial, the target alternative or attribute was shown longer than the other. **c**) 15 core choice problems used to construct choice trials. Each circle represents one choice alternative, described by its probability p and outcome m . Each connected pair of circles indicates one pair of choices used to construct the 120 experimental trials. Color indicates the α value for which the expected utilities of a pair are equal (see panel **d**). **d**) Distribution of indifference implied α values of the 15 core choice problems from **c**). **e**) Catch trials. Each connected pair of triangles represents one of 20 catch trials, where one alternative is dominated by the other on both attributes.

No effects of presentation duration across presentation formats

We hypothesized that alternatives shown longer (in alternative-wise presentation) and alternatives with better values on the attribute shown longer (in attribute-wise presentation) would be chosen more frequently. To this end, we computed individuals' probabilities of choosing Hp when it was favoured by presentation duration (because either the alternative Hp or the attribute p was shown longer) and when it was not. We then used directed Bayes Factor t -tests to test the hypotheses of positive versus a null effects in the difference of choice probabilities (see [Methods](#)).

On average and across presentation formats, par-

ticipants chose Hp with a probability of 62.0% when it was favoured by presentation duration, and 62.3% when it was not. There was strong evidence against an effect of presentation duration on choice probability across presentation formats (mean difference = -0.2% [$-1.1\%, 0.6\%$]; mean $d = -0.04$ [$-0.20, 0.11$]; $BF_{+0} = 0.053$, $BF_{0+} = 18.84$). Separate tests for each presentation format confirmed evidence against an effect of presentation duration on choice probability, with strong evidence against an effect in alternative-wise presentation (H1a; mean difference = -0.1% [$-1.5\%, 1.1\%$]; mean $d = -0.02$ [$-0.18, 0.13$]; $BF_{0+} = 15.43$; Figure 3a, e) and strong evidence against an effect in attribute-wise presentation (H1b; mean dif-

Term	Mean	SD	95% HDI		\hat{R}	ESS (bulk)	ESS (tail)
			Lower	Upper			
Intercept	0.35	0.11	0.12	0.56	1.01	351	586
EV diff. ($H_p - H_m$; z-scored)	1.96	0.10	1.77	2.15	1.00	763	1962
Duration (H_p or p longer)	-0.02	0.04	-0.10	0.05	1.00	8644	6667
Order (H_p or p last)	0.17	0.04	0.09	0.25	1.00	7602	6943
Format (by alt.)	-0.06	0.04	-0.13	0.01	1.00	10598	6192
Duration \times Format	0.01	0.08	-0.13	0.16	1.00	9983	6276
Order \times Format	0.09	0.08	-0.06	0.23	1.00	10963	6350

Table 2. Fixed effects estimates from Bayesian mixed-effects logistic regression. Dependent variable: H_p choice. Categorical predictors (Duration, Order, Format) were effect-coded, and positive estimates are associated with the level indicated in parentheses. See [Methods](#) for details on predictor variables. ESS: Effective sample size. Results obtained from two MCMC chains with 5000 posterior samples and 1000 tuning samples. Terms credibly different from zero are shown in boldface.

ference = -0.4% [-1.6%, 0.8%]; mean d = -0.05 [-0.21, 0.10]; BF_{0+} = 19.34; Figure 3b, f).

Strong evidence for effect of presentation order in alternative-wise presentation

We furthermore hypothesized that alternatives shown last (in alternative-wise presentation) and alternatives with better values on the attribute shown last (in attribute-wise presentation) would be chosen more frequently. To test the marginal effect of presentation order on choice (across presentation formats), we computed the individual probabilities of choosing H_p when it was favored by the last presentation stage (because either the alternative H_p or the attribute p was shown last) and when it was not, and performed directed Bayes factor t -tests and Bayesian estimation of their difference. Across presentation formats, participants chose the H_p alternative with a probability of 63.3% when it was favoured in the final presentation stage, and 60.9% when it was not (Table 1). There was extreme evidence in favour of an effect of presentation order on choice across presentation formats (mean difference = 2.4% [1.2%, 3.5%]; mean d = 0.32 [0.16, 0.47]; BF_{+0} = 806.85). Separate tests for each presentation format showed that the effect was more specific to alternative-wise presentation, with extreme evidence for an effect (H2a; mean difference = 3.0% [1.6%, 4.4%]; mean d = 0.33 [0.17, 0.49]; BF_{+0} = 845.71; Figure 3c, g). In attribute-wise presentation, evidence anecdotally favoured a positive effect (H2b; mean difference = 1.5% [0.0%, 3.1%]; mean d = 0.15 [-0.01, 0.31]; 96.67% of posterior mass

Discussion

In this study, we investigated the causal effects of attribute- and alternative-wise presentation duration and order in two-alternative risky choice. In contrast to causal interpretations of simple gaze-dependent accumulation models, our data did not support a causal role of presentation duration in either presentation format. Instead, we found strong evidence for a causal effect of presentation order on choice, especially when information was presented alternative-wise.

Prior work has reported causal effects of viewing- and presentation duration on preferential (and perceptual) choice (Armel et al., 2008; Fisher, 2021; Liu, Lyu, et al., 2020; Pärnamets et al., 2015; Shimojo et al., 2003; Sui et al., 2020; Tavares et al., 2017) using external control of presentation durations or gaze-contingent decision prompts. In free choice paradigms, these effects are well described by gaze-dependent accumulation models like the aDDM, which assume that an alternative’s value representations are discounted while it is not fixated by (or presented to) the decision maker.

In our task, evidence strongly favoured null effects of presentation duration on choice, both in alternative- and attribute-wise presentation formats.

This lack of replication suggests that more work is needed to better understand the conditions under which fixation- or presentation duration can affect choice behaviour. It is possible that duration effects

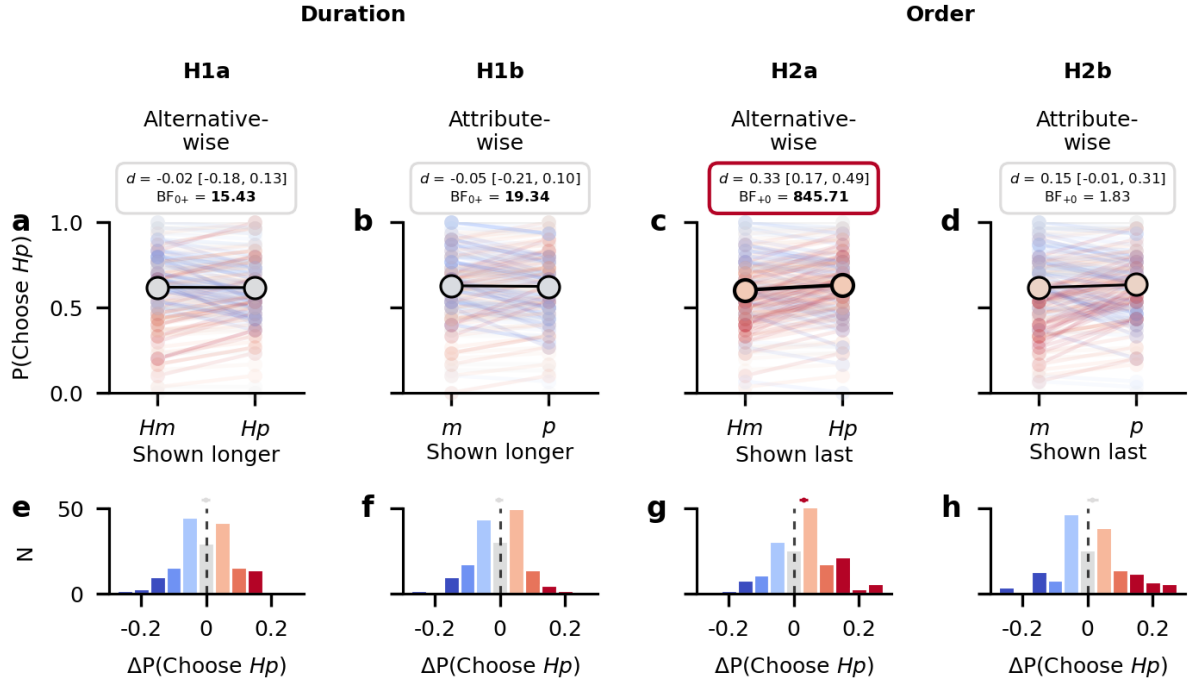


Figure 3. Effects of presentation duration and order on choice. Panels a-d show individual and mean changes in choice probabilities of the higher probability alternative H_p , for duration (a, b) and order manipulations (c, d) in alternative- (a, c) and attribute-wise (b, d) presentation. Each semi-transparent connected pair of dots indicates choice probabilities for H_p of a single participant in trials where H_m vs. H_p was favoured by the presentation manipulation. Group means are indicated by opaque large dots and lines. Cohen's d with HDI_{95%} obtained from paired BEST analysis, and the Bayes factors in favour of a positive directed over a null effect (BF_{+0} ; see Methods) or its reciprocal in favour of a null effect (BF_{0+}) are given for each panel. Panels e-f show corresponding distributions of individual changes (individual slopes in panels a-d), with colors coding the size of individual changes. Small points and horizontal lines above histograms indicate mean and HDI_{95%} change (gray if HDI_{95%} includes 0, red if 0 is excluded).

are moderated by specific aspects of the decision task itself, like presentation durations, and the type and delivery of choice stimuli.

It remains, however, an experimental challenge to investigate the causal effects of different aspects of information search on choice: With the external control of presentation parameters, participants are prompted for a decision either before or after they would have made a choice in a free response paradigm. Studies experimentally controlling stimulus presentation (including studies with gaze-dependent decision prompts) thereby differ from simple decision making tasks where participants determine the time of choice themselves; their experimental manipulations interfere with and alter the natural course of decision making. Here, recent work has demonstrated an elegant possibility to induce biases in information search, with downstream effects on choice: Gwinn et al. (2019) had

participants acquire attentional biases in a separate task, which carried over to a choice task otherwise free from interference. Notably, the authors found that not gaze duration, but the location of the first fixation, mediated the effect of the attentional manipulation on choice, providing additional evidence of effects of information order on choice over viewing duration alone.

Regarding the association between the order of information search during decision making and choice, most eye tracking studies of decision making find the last fixation to be predominantly directed towards the chosen alternative (Fiedler & Glöckner, 2012; Glickman et al., 2019; Krajbich et al., 2010; Krajbich & Rangel, 2011; Stewart, Hermens, et al., 2016). This association is predicted by gaze-dependent accumulation models like the aDDM, where relative evidence for an alternative is more likely to cross the decision

boundary while it is fixated. Notably, however, the aDDM does not imply a causal role of information order or the final fixation on choice in particular. In contrast, leaky accumulator models (e.g., Ashby et al., 2016; Glickman et al., 2019; Molter et al., 2021; Usher & McClelland, 2001) predict that information acquired early in the trial decays, and the relative weight of more recent information is increased.

We found strong evidence for a causal effect of presentation order, so that information presented last influences decisions more than information presented earlier. This recency effect was particularly strong in the alternative-wise presentation format.

Our results thereby support theories of gaze-dependent accumulation which include a form of accumulation leak and can account for causal recency effects. Interestingly, prior work has identified recency effects in decisions from experience and related paradigms, where information about risky alternatives' outcomes is actively sampled by the decision maker, and therefore also acquired sequentially (Herwig et al., 2004; Tsetsos et al., 2012). Our results suggest that a similar, and causally directed effect is present during decisions from description, where stimulus information is also experienced sequentially.

One possible explanation for the observed recency effect is that decision makers have imperfect memory about the stimulus information (potentially reinforced by the fast-paced presentation) and then base their choice on more recent, better remembered information. This would constitute a memory-bias within single trials, similar to previously described effects on longer timescales (Gluth et al., 2015; Weilbacher et al., 2021).

Our results further provide evidence that the observed association between decision makers' last fixation and choice goes beyond simple "confirmation" or response locking explanations assuming that choices are already determined before the final fixation and response is made.

In addition to the last fixation preferably being directed to the chosen alternative, eye tracking studies of decision making typically find longer gaze towards the chosen alternative (Cavanagh et al., 2014; Fiedler & Glöckner, 2012; Glickman et al., 2019; Krajbich et al., 2010; Krajbich & Rangel, 2011; Molter et al., 2021; Shimojo et al., 2003; Stewart, Hermens, et al., 2016). Both being associated with choice, these two variables might be confounded frequently, that is, alternatives that are looked at last are also looked at longer during

the decision. Consequently, the effects of duration and order can be mistaken for each other, without explicit control. This issue can also occur in experimental designs using gaze-dependent decision prompts, as – depending on the specific conditions triggering the prompt – the last fixation is potentially more likely to be directed at the alternative determined to receive longer gaze.

While prior work demonstrated that random or externally induced fluctuations in visual attention can have downstream effects on choice mainly through viewing duration, our work highlights the importance of the temporal order in which information is encountered. This suggests that, in situations where presentation order can be controlled, such as television-, cinema- or online video advertisements, choices could be systematically shifted towards certain alternatives. We note, however, that these settings typically do not involve the need to choose (as in our experiment). Yet, it is conceivable that for those cinema-goers who already decided to buy ice-cream, still deliberating which flavour to get, choices are shifted towards the flavour presented last (and closest to their buying decision) in the advertisement clip. Desired (e.g., healthful) choices could then be promoted by reminding decision makers of them just before choices are made.

Similarly, Sullivan and Huettel (2021) argued in favour of time-dependent interventions to promote healthful food choices, as they found health information to enter the decision process with a longer latency than taste information. Reducing overall time pressure could increase healthful choices by reducing the gap in attribute latencies. Alternatively, they suggest that showing health-related information *first* should decrease the latency gap and thereby promote healthy choices. Our results highlight the opposite direction: Attribute-information shown first to participants reduced participants' likelihood of choosing alternatives with better values on that attribute. Note that our study addressed primacy and recency effects on the same one-dimensional construct. We therefore cannot distinguish positive recency- from negative primacy effects, as information shown last was by design *not* shown first. Future work could address possible independent contributions of both effects.

A possible explanation for these apparent differences, apart from the different goods being chosen, is that the latency construct in the model by Sullivan

and Huettel (2021) might be more related to consideration *duration* than *order*: Attributes with lower latencies can, by definition, influence the decision process for a longer duration than others. This way, the order and duration effects are inherently linked in their model.

In conclusion, we showed that presentation order but not presentation duration had a causal effect on risky choice behaviour, in an external presentation paradigm. This has important implications for theories of decision making, which should incorporate mechanisms like accumulation leak, to account for observed recency effects, and suggests an important role of the presentation order – particularly the information presented just before a choice is made – in real world decisions.

Methods

All data collection and analysis procedures were pre-registered on the Open Science Framework ahead of data collection (Molter & Mohr, 2021).

Participants

The experiment was conducted online and participants were recruited via the platform Prolific.co. Our target sample size was 200 participants. We accepted entries on the Prolific platform until 200 complete submissions were reached. Sample size was determined to exceed that of previous laboratory-based experiments addressing related questions (e.g., Armel et al., 2008; Fisher, 2021; Liu, Lyu, et al., 2020; Pärnamets et al., 2015; Sui et al., 2020), accounting for the possibility of lower quality data due to the online setting. Participants were paid £3.75 for completing the study, which took around 30 minutes. They had the chance to win a bonus amount ranging from £0 to £9.85, as one of their chosen gambles was randomly determined and played out after the experiment. The experiment included webcam-based eye tracking (see [Supplementary Information](#)), so only participants with a working webcam using Google Chrome or Mozilla Firefox could participate. Participants gave informed consent prior to participation in the study. Ethics approval for this study was obtained from the German Association for Experimental Economic Research e.V. All methods were performed in accordance with the relevant guidelines and regulations.

Participants were excluded from the analysis without replacement if one of the following criteria was met: i) The participant chose a dominated alternative more than 4 times (20% of catch trials, see below). ii) The participant reported to be red-green colorblind or having difficulty distinguishing the colors in the task. iii) The participant reported having technical difficulties that prevented them from diligent execution of the task. iv) The participant’s self-reported decision strategy suggested that task instructions were misunderstood. v) The participant reported nonserious participation in the task (Aust et al., 2013).

Twenty-one participants were excluded from the analyses (16 for criterion 1, four for criterion 2, one for criterion 4), resulting in a final sample size of 179 (mean \pm s.d. age = 32.46 ± 13.85 ; 84 females, 94 males, one other).

Behavioral task

Participants performed a binary sequential presentation risky choice task, where they made repeated incentivized choices between two graphically displayed all-or-nothing gambles (Figure 2). Each gamble represented the chance to win a monetary amount m with a probability p and nothing otherwise. Winning amounts m were represented by partially filled bars (a fully filled bar represented an amount of 10£), and winning probabilities p by pie charts. Each trial was divided into a presentation phase and a choice phase. During the presentation phase, participants learned about the two available gambles’ attributes. Information was presented sequentially. There were two types of presentation: Alternative-wise and attribute-wise presentation.

In trials with alternative-wise presentation (Figure 2a) both attributes of one gamble were shown simultaneously, followed by both attributes of the other gamble. In trials with attribute-wise presentation (Figure 2b), one attribute (e.g., winning probability p) of both gambles was presented simultaneously, followed by the other attribute (e.g., amount m) of both gambles. Presentation always alternated two times between alternatives or attributes (i.e., A-B-A-B). Crucially, one alternative (in trials with alternative-wise presentation) or attribute (in trials with attribute-wise presentation) in each trial was selected to be the target and shown longer than the other one. Target attributes or alternatives were always presented for

1500 ms, whereas other alternatives or attributes were presented for 1000 ms, resulting in a final presentation time advantage for the target of 1000 ms (2x 1500 ms vs. 2x 1000 ms).

After the presentation phase, stimulus information was hidden and participants were prompted to make a choice between the two alternatives within 3 seconds. After completing all choice trials, one gamble chosen by the participant in a randomly determined trial was played out for a real bonus payment. The task can be run at https://moltaire.github.io/causality_task.

All participants made choices for the same 140 choice problems (see below) divided into two blocks. Trial order in each block and the horizontal position of the alternatives in each trial was randomized. Block order was counterbalanced between participants.

Stimuli

We created a set of 15 core choice problems including one alternative with a higher probability of winning a lower amount (Hp) and one alternative with a lower probability of winning a higher amount (Hm). Choice problems were created algorithmically to cover most of the attribute space, be maximally different from each other, and be diagnostic of different risk attitudes. For this last criterion, we controlled the α values for which two alternatives in a pair would have equal expected utility (using a standard power utility function). The distribution of these indifference-implied α values is shown in Figure 2d. Core choice problems are illustrated in Figure 2c.

Then eight trials were created for each core problem by fully crossing the factors (i) presentation format (alternative-wise vs. attribute-wise), (ii) target alternative / attribute (Hp vs. Hm ; p vs. m), and (iii) presentation order (target first and third vs. second and last). This resulted in a total of 120 experimental choice trials.

We added 20 catch trials with one dominant alternative for a total of 140 trials (Figure 2e). In catch trials, individual presentation durations were set to 1250 ms, resulting in the same overall presentation duration (5000 ms), but no presentation time advantage for any alternative or attribute.

Statistical modelling

We performed a Bayesian logistic regression analysis of choice behavior, with choice (Hp vs. Hm) as the dependent variable and the following predictors: Expected value difference ($EV_{Hp} - EV_{Hm}$; z-scored), presentation format (by-attribute vs. by-alternative; effect-coded), duration-favored (Hp favored vs. Hm favored by duration manipulation; effect-coded), last-stage-favored (Hp favored vs. Hm favored in last presentation stage; effect-coded), and interaction terms between presentation format and duration-favored and last-stage-favored. The model included random intercepts and slopes over participants and used the default priors set by the *bambi* library (Westfall, 2017).

We performed directed Bayes factor (BF) t -tests (Morey & Rouder, 2011) of our main hypotheses, testing the directed hypotheses that differences in choice probabilities are larger than zero, over a point null hypothesis. All Bayes factor t -tests used default JZS priors (Cauchy distributed with scale $r = \sqrt{2}/2$) implemented in the BayesFactor package (Morey & Rouder, 2018).

Additionally, we ran paired Bayesian estimation (BEST; Kruschke, 2013, 2014) analyses to compute mean differences, effect sizes d and associated 95% highest posterior density intervals ($HDI_{95\%}$).

We ran two chains with 5000 samples each after a tuning phase of 1000 samples for all BEST analyses and the Bayesian mixed-effects model. Convergence was diagnosed visually and by means of the Gelman-Rubin statistic ($|1 - \hat{R}| \leq 0.05$ for all chains).

We determine parameters in the regression and BEST analyses to be credibly different from zero if $HDI_{95\%}$ exclude zero or at least 95% of the posterior mass is above (below) zero. For interpretation of Bayes factors' evidence strength, we follow the conventional categorization based on Jeffreys (1998): Anecdotal ($1 < BF < 3$); Moderate ($3 \leq BF < 10$); Strong ($10 \leq BF < 30$); Very strong ($30 \leq BF < 100$); Extreme ($100 \leq BF$).

Software

The task was programmed in jsPsych (de Leeuw, 2015). Webcam-based eye tracking was implemented using the webgazer.js library (Papoutsaki et al., 2016). Data processing and analyses were done in Python

with numpy (Harris et al., 2020) and pandas (McKinney, 2012) libraries. Bayesian analyses were implemented in PyMC3 (Salvatier et al., 2016), mixed models used bambi (Capretto et al., 2021). Bayes Factor *t*-tests were performed using the R package BayesFactor (Morey & Rouder, 2018). Figures were created using matplotlib (Hunter, 2007).

Data and code availability

All raw and preprocessed data and scripts to reproduce all processing and analyses steps and figures are available at <https://github.com/moltaire/gaze-choice-causality>.

References

- Allais, M. (1953). L'Extension des Theories de l'Equilibre Economique General et du Rendement Social au Cas du Risque. *Econometrica*, 21(2), 269–290. <https://doi.org/10.2307/1905539>
- Amasino, D. R., Sullivan, N. J., Kranton, R. E., & Huettel, S. A. (2019). Amount and time exert independent influences on intertemporal choice. *Nature Human Behaviour*, 3(4), 383–392. <https://doi.org/10.1038/s41562-019-0537-2>
- Armel, K. C., Beaumel, A., & Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making*, 3(5), 8.
- Ashby, N. J. S., Jekel, M., Dickert, S., & Glöckner, A. (2016). Finding the right fit: A comparison of process assumptions underlying popular drift-diffusion models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(12), 1982–1993. <https://doi.org/10.1037/xlm0000279>
- Aust, F., Diedenhofen, B., Ullrich, S., & Musch, J. (2013). Seriousness checks are useful to improve data validity in online research. *Behavior Research Methods*, 45(2), 527–535. <https://doi.org/10.3758/s13428-012-0265-2>
- 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>
- Capretto, T., Piho, C., Kumar, R., Westfall, J., Yarkoni, T., & Martin, O. A. (2021). Bambi: A simple interface for fitting Bayesian linear models in Python. *arXiv:2012.10754 [stat]*.
- Cavanagh, J. F., Wiecki, T. V., Kochar, A., & Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology. General*, 143(4), 1476–1488. <https://doi.org/10.1037/a0035813>
- de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, 47(1), 1–12. <https://doi.org/10.3758/s13428-014-0458-y>
- Fiedler, S., & Glöckner, A. (2012). The Dynamics of Decision Making in Risky Choice: An Eye-Tracking Analysis. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00335>
- Fisher, G. (2021). Intertemporal Choices Are Causally Influenced by Fluctuations in Visual Attention. *Management Science*. <https://doi.org/10.1287/mnsc.2020.3732>
- 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), e1007201. <https://doi.org/10.1371/journal.pcbi.1007201>
- Glöckner, A., & Herbold, A.-K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making*, 24(1), 71–98. <https://doi.org/10.1002/bdm.684>
- 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*, 1–12. <https://doi.org/10.1038/s41562-020-0822-0>
- Gluth, S., Sommer, T., Rieskamp, J., & Büchel, C. (2015). Effective connectivity between hippocampus and ventromedial prefrontal cortex controls preferential choices from memory [Publisher: Elsevier]. *Neuron*, 86(4), 1078–1090. <https://doi.org/10.1016/j.neuron.2015.04.023>
- Gluth, S., Spektor, M. S., & Rieskamp, J. (2018). Value-based attentional capture affects multi-alternative decision making (M. J. Frank, Ed.). *eLife*, 7, e39659. <https://doi.org/10.7554/eLife.39659>
- Gwinn, R., Leber, A. B., & Krajovich, I. (2019). The spillover effects of attentional learning on value-based choice. *Cognition*, 182, 294–306. <https://doi.org/10.1016/j.cognition.2018.10.012>
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>

- Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from Experience and the Effect of Rare Events in Risky Choice. *Psychological Science*, 15(8), 534–539. <https://doi.org/10.1111/j.0956-7976.2004.00715.x>
- Hommel, B., Chapman, C. S., Cisek, P., Neyedli, H. F., Song, J.-H., & Welsh, T. N. (2019). No one knows what attention is. *Attention, Perception, & Psychophysics*, 81(7), 2288–2303. <https://doi.org/10.3758/s13414-019-01846-w>
- Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Isham, E. A., & Geng, J. J. (2013). Looking Time Predicts Choice but Not Aesthetic Value. *PLOS ONE*, 8(8), e71698. <https://doi.org/10.1371/journal.pone.0071698>
- Jeffreys, S. H. (1998). *The Theory of Probability* (Third Edition). Oxford University Press.
- Kahneman, D., & Tversky, A. (2012). Choices, Values, and Frames. *Handbook of the Fundamentals of Financial Decision Making* (pp. 269–278). World Scientific. https://doi.org/10.1142/9789814417358_0016
- Kim, B., Seligman, D., & Kable, J. W. (2012). Preference Reversals in Decision Making Under Risk are Accompanied by Changes in Attention to Different Attributes. *Frontiers in Neuroscience*, 6. <https://doi.org/10.3389/fnins.2012.00109>
- Krajibich, 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>
- Krajibich, I., Lu, D., Camerer, C., & Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology*, 3, 193. <https://doi.org/10.3389/fpsyg.2012.00193>
- Krajibich, 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 of the United States of America*, 108(33), 13852–13857. <https://doi.org/10.1073/pnas.1101328108>
- Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, 142(2), 573–603. <https://doi.org/10.1037/a0029146>
- Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press.
- Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press
OCLC: ocn861318341.
- 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>
- Liu, H.-Z., Lyu, X.-K., Wei, Z.-H., Mo, W.-L., Luo, J.-R., & Su, X.-Y. (2020). Exploiting the dynamics of eye gaze to bias intertemporal choice. *Journal of Behavioral Decision Making*, n/a(n/a). <https://doi.org/10.1002/bdm.2219>
- Liu, H.-Z., Zhou, Y.-B., Wei, Z.-H., & Jiang, C.-M. (2020). The power of last fixation: Biasing simple choices by gaze-contingent manipulation. *Acta Psychologica*, 208, 103106. <https://doi.org/10.1016/j.actpsy.2020.103106>
- Lopez-Persem, A., Domenech, P., & Pessiglione, M. (2016). How prior preferences determine decision-making frames and biases in the human brain. *eLife*, 5. <https://doi.org/10.7554/eLife.20317>
- Maier, S. U., Raja Beharelle, A., Polanía, R., Ruff, C. C., & Hare, T. A. (2020). Dissociable mechanisms govern when and how strongly reward attributes affect decisions. *Nature Human Behaviour*, 1–15. <https://doi.org/10.1038/s41562-020-0893-y>
- McKinney, W. (2012). *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. "O'Reilly Media, Inc."
- Milosavljevic, M., Navalpakkam, V., Koch, C., & Rangel, A. (2012). Relative visual saliency differences induce sizable bias in consumer choice. *Journal of Consumer Psychology*, 22(1), 67–74. <https://doi.org/10.1016/j.jcps.2011.10.002>
- Mittone, L., & Papi, M. (2020). Inducing alternative-based and characteristic-based search procedures in risky choice. *Judgment and Decision Making*, 10.
- Mohr, P. N. C., Heekeren, H. R., & Rieskamp, J. (2017). Attraction Effect in Risky Choice Can Be Explained by Subjective Distance Between Choice Alternatives. *Scientific Reports*, 7(1), 1–10. <https://doi.org/10.1038/s41598-017-06968-5>
- Molter, F., & Mohr, P. N. C. (2021). Investigating the causal effects of presentation duration, sequence, and format on binary risky choice. *Open Science Framework*. <https://osf.io/bth3m>
- Molter, F., Thomas, A., Huettel, S. A., Heekeren, H., & Mohr, P. N. C. (2021). Gaze-dependent evidence accumulation predicts multi-alternative risky choice behaviour. <https://doi.org/10.31234/osf.io/x6nbf>
- Morey, R. D., & Rouder, J. N. (2011). Bayes factor approaches for testing interval null hypotheses. *Psychological Methods*, 16(4), 406–419. <https://doi.org/10.1037/a0024377>
- Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of bayes factors for common designs*. Manual.

- Mormann, M., & Russo, J. E. (2021). Does Attention Increase the Value of Choice Alternatives? *Trends in Cognitive Sciences*. <https://doi.org/10.1016/j.tics.2021.01.004>
- Mullett, T. L., & Stewart, N. (2016). Implications of visual attention phenomena for models of preferential choice. *Decision*, 3(4), 231. <https://doi.org/10.1037/dec0000049>
- Papoutsaki, A., Sangkloy, P., Laskey, J., Daskalova, N., Huang, J., & Hays, J. (2016). WebGazer: Scalable webcam eye tracking using user interactions. *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*, 3839–3845.
- Pärnamets, P., Johansson, P., Hall, L., Balkenius, C., Spivey, M. J., & Richardson, D. C. (2015). Biasing moral decisions by exploiting the dynamics of eye gaze. *Proceedings of the National Academy of Sciences of the United States of America*, 112(13), 4170–4175. <https://doi.org/10.1073/pnas.1415250112>
- Reeck, C., Wall, D., & Johnson, E. J. (2017). Search predicts and changes patience in intertemporal choice. *Proceedings of the National Academy of Sciences*, 114(45), 11890–11895. <https://doi.org/10.1073/pnas.1707040114>
- Rigoux, L., Stephan, K. E., Friston, K. J., & Daunizeau, J. (2014). Bayesian model selection for group studies — Revisited. *NeuroImage*, 84, 971–985. <https://doi.org/10.1016/j.neuroimage.2013.08.065>
- Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, 2, e55. <https://doi.org/10.7717/peerj-cs.55>
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2), 461–464. <https://doi.org/10.1214/aos/1176344136>
- Sepulveda, P., Usher, M., Davies, N., Benson, A. A., Ortolova, P., & De Martino, B. (2020). Visual attention modulates the integration of goal-relevant evidence and not value (C. Büchel, V. Wyart, A. L. Filipowicz, & S. Gluth, Eds.). *eLife*, 9, e60705. <https://doi.org/10.7554/eLife.60705>
- 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>
- Shinn, M., Lam, N. H., & Murray, J. D. (2020). A flexible framework for simulating and fitting generalized drift-diffusion models (T. Kahnt, J. I. Gold, L. Ding, & E.-J. Wagenmakers, Eds.). *eLife*, 9, e56938. <https://doi.org/10.7554/eLife.56938>
- 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>
- Smith, S. M., & Krajbich, I. (2019). Gaze Amplifies Value in Decision Making. *Psychological Science*, 30(1), 116–128. <https://doi.org/10.1177/0956797618810521>
- Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. *NeuroImage*, 46(4), 1004–1017. <https://doi.org/10.1016/j.neuroimage.2009.03.025>
- Stewart, N., Gächter, S., Noguchi, T., & Mullett, T. L. (2016). Eye Movements in Strategic Choice. *Journal of Behavioral Decision Making*, 29(2-3), 137–156. <https://doi.org/10.1002/bdm.1901>
- Stewart, N., Hermens, F., & Matthews, W. J. (2016). Eye Movements in Risky Choice: Eye Movements in Risky Choice. *Journal of Behavioral Decision Making*, 29(2-3), 116–136. <https://doi.org/10.1002/bdm.1854>
- Sui, X.-Y., Liu, H.-Z., & Rao, L.-L. (2020). The timing of gaze-contingent decision prompts influences risky choice. *Cognition*, 195, 104077. <https://doi.org/10.1016/j.cognition.2019.104077>
- Sullivan, N. J., & Huettel, S. A. (2021). Healthful choices depend on the latency and rate of information accumulation. *Nature Human Behaviour*, 1–9. <https://doi.org/10.1038/s41562-021-01154-0>
Bandiera_abtest: a Cg.type: Nature Research Journals Primary_atype: Research Subject_term: Computational neuroscience;Human behaviour Subject_term.id: computational-neuroscience;human-behaviour
- Tavares, G., Perona, P., & Rangel, A. (2017). The Attentional Drift Diffusion Model of Simple Perceptual Decision-Making. *Frontiers in Neuroscience*, 11, 468. <https://doi.org/10.3389/fnins.2017.00468>
- Thomas, A. W., Molter, F., & Krajbich, I. (2021). Uncovering the computational mechanisms underlying many-alternative choice (T. Verstynen, M. J. Frank, & E. Peterson, Eds.). *eLife*, 10, e57012. <https://doi.org/10.7554/eLife.57012>
- Thomas, A. W., Molter, F., Krajbich, I., Heekeren, H. R., & Mohr, P. N. C. (2019). Gaze bias differences capture individual choice behaviour. *Nature Human Behaviour*, 3(6), 625–635. <https://doi.org/10.1038/s41562-019-0584-8>
- Tsetsos, K., Chater, N., & Usher, M. (2012). Salience driven value integration explains decision biases and preference reversal. *Proceedings of the National Academy of Sciences*, 109(24), 9659–9664. <https://doi.org/10.1073/pnas.1119569109>
- 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>

- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108(3), 550–592. <https://doi.org/10.1037/0033-295X.108.3.550>
- Weber, E., & Kirsner, B. (1997). Reasons for Rank-Dependent Utility Evaluation. *Journal of Risk and Uncertainty*, 14(1), 41–61. <https://doi.org/10.1023/A:1007769703493>
- Weilbacher, R. A., Krajbich, I., Rieskamp, J., & Gluth, S. (2021). The influence of visual attention on memory-based preferential choice. *Cognition*, 215, 104804. <https://doi.org/10.1016/j.cognition.2021.104804>
- Westbrook, A., van den Bosch, R., Määttä, J. I., Hofmans, L., Papadopetraki, D., Cools, R., & Frank, M. J. (2020). Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work. *Science*, 367(6484), 1362–1366. <https://doi.org/10.1126/science.aaz5891>
- Westfall, J. (2017). Statistical details of the default priors in the bambi library. *arXiv:1702.01201 [stat]*. Retrieved February 3, 2020, from <http://arxiv.org/abs/1702.01201>
- Yang, X., & Krajbich, I. (2020). Webcam-based online eye-tracking for behavioral research. <https://doi.org/10.31234/osf.io/qhme6>
- Yechiam, E., & Hochman, G. (2013). Loss-aversion or loss-attention: The impact of losses on cognitive performance. *Cognitive Psychology*, 66(2), 212–231. <https://doi.org/10.1016/j.cogpsych.2012.12.001>

Author contributions

Conceptualization: F.M. and P.N.C.M. Data curation: F.M. Formal analysis: F.M. Funding acquisition: P.N.C.M. Investigation: F.M. Methodology: F.M. Project administration: P.N.C.M. Resources: F.M. Software: F.M. Supervision: P.N.C.M. Validation: F.M. Visualization: F.M. Writing - original draft: F.M. Writing - review & editing: F.M. and P.N.C.M.

Competing Interests Statement

The authors declare no competing interests.

Supplementary Information

Webcam-based eye tracking

Participants' gaze during the task was recorded using webcam-based eye tracking implemented in the JavaScript library webgazer.js (Papoutsaki et al., 2016; Yang & Krajbich, 2020). To this end, the task included a webcam setup at the start of the experiment, a 13-point calibration routine, and a 5-point validation routine before the start of each block, where participants were instructed to focus their gaze on indicated screen locations for three seconds. Validation points were chosen to correspond to the locations where stimulus information would be presented during experimental choice trials.

We assessed the quality of the collected eye tracking data using multiple complementing metrics: First, we measured the sampling rates obtained in each validation, which heavily depend on the device that runs the experimental task. Mean \pm s.d. sampling rates were 15.96 ± 7.95 Hz and ranged from 0.49 Hz to 29.55 Hz, and 71% of blocks had sampling rates above 10 Hz during validation.

Second, as a measure of bias, we computed the average absolute distance of the estimated gaze location samples from the corresponding validation targets in x- and y directions (in percent of screen width and height, respectively). Mean \pm s.d. error (clustered by block) was $5.58 \pm 7.11\%$ in x- and $8.79 \pm 9.76\%$ in y-direction.

Third, as a measure of accuracy, we computed the sample variability as the standard deviation of samples corresponding to a single validation target in x- and y-directions. Mean \pm s.d. variability was $5.81 \pm 3.07\%$ in x- and $7.46 \pm 4.70\%$ in y-direction.

Fourth, we computed the proportion of samples within an elliptical Area of Interest (AoI) around the corresponding validation target. For this, we set the AoI-width and -height to 15% of the screen width and height, respectively (note that widths and heights of 25% would leave no white space between AoIs due to the stimulus spacing used in the task). On average, only $51.80 \pm 26.12\%$ of the samples were contained within a validation target's AoI.

Next, we set minimum thresholds on each of these parameters to classify validation results as valid or invalid: We required a minimum sampling rate of 10 Hz, a maximum average error and a maximum variability

of 20% in x- and y-directions, and a minimum of 60% of samples within the validation target AoI. Only 33 of 356 blocks (9.27%) fulfilled these minimal quality requirements. In addition, it can be assumed that eye tracking quality only *decreased* during a block, due to head movement or other sources of variability (e.g., changes in lighting conditions, etc.). We therefore did not perform any further explorative analyses on the eye tracking data. We note that it might be possible to correct for systematic biases in validation (e.g., using clustering approaches). Future studies could benefit from improved webcam-based eye tracking data by performing online checks of validation accuracy and repeating calibration and validation steps until an acceptable level of quality is reached.

Influence of presentation format on value integration

Our preregistration contained the additional hypothesis that presentation format (alternative- vs. attribute-wise presentation) elicits different forms of value integration in the decision process. Specifically, we hypothesized, that alternative-wise presentation would elicit more integrative, within-alternative processing, whereas attribute-wise presentation would be associated with more comparative processing within attribute dimensions and between alternatives. To test this hypothesis, we fit two behavioral models, which used within-alternative and within-attribute integration of attributes, respectively, to the choice and response time data of each participant. Crucially, the models were fit separately for trials with alternative- and attribute-wise presentation. We then computed the relative fit of the models for each participant and presentation format by taking the difference of the models' Bayesian Information Criterion (BIC) (Schwarz, 1978). Finally, we performed directed paired Bayes factor *t*-tests of the differences, testing the directed hypothesis that the relative fit of the within-alternative integration model over the between-alternative model is increased in alternative-wise vs. attribute-wise presentation.

Behavioural modelling

We analyzed participants' choice and response time data using two different behavioral models. Both models shared a similar general structure: During the

presentation phase of the trial, a relative evidence signal R is assumed to be formed (Figure S1a-b), depending on the presentation format, duration, and order. After the choice prompt, a noisy diffusion process between two decision bounds (corresponding to choosing Hp and Hm , respectively) is initiated that elicits a choice at a time point t (Figure S1c). Crucially, the drift rate of the diffusion process is proportional to the relative evidence signal R at the end of the presentation phase. For both models, the diffusion process after the choice prompt is parameterized by a drift constant v (that scales the final relative evidence signal R), a noise parameter s , while the boundary separation is kept constant at a value of 1. The two models only differed in the process of calculating the relative evidence signal R during the presentation phase.

Alternative-wise integration model The alternative-wise integration model (slate-gray in Figure S1) computes alternative-wise expected utilities, using a standard utility function ($U_i = p_i m_i^\alpha$). During the presentation phase, the difference between the expected utilities is assumed to accumulate over time. Critically, the momentarily not presented alternative's utility is discounted by an alternative-wise gaze-discount θ . The final relative evidence signal $R_{altwise}$ is given by

$$R_{altwise} = g_{Hp}(U_{Hp} - \theta U_{Hm}) + g_{Hm}\theta U_{Hp} - U_{Hm}$$

Where g_{Hp} and g_{Hm} are the relative presentation durations of Hp and Hm . Note that in trials with attribute-wise presentation, g_{Hp} and g_{Hm} are set to 0.5, as both alternatives' attributes are presented equally long. Additionally, the model uses a parameter b_{last} to predict order effects in trials with alternative-wise presentation: Positive b_{last} shift R towards the last-presented alternative, negative b_{last} shift it to the alternative presented first.

Attribute-wise integration model The attribute-wise model (orange in Figure S1) assumes that the relative evidence R is computed through attribute comparisons between alternatives, weighted addition of attribute differences, and accumulation of differences over time. Importantly, it assumes that the momentarily not presented attributes are discounted

by an attribute-wise gaze-discount η . The final relative evidence signal $R_{attwise}$ is given by

$$R_{attwise} = g_p(w_p \Delta_p + \eta(1 - w_p) \Delta_m) + g_m(\eta w_p \Delta_p + (1 - w_p) \Delta_m)$$

Where Δ_p and Δ_m are the attribute differences between Hp and Hm alternatives. g_p and g_m are relative presentation durations for p and m attributes, respectively. w_p controls the relative weighting between probability and outcome attributes. Note that attributes are also normalized in each trial (by dividing by the sum of values on the attribute). During trials with alternative-wise presentation, the relative gaze durations towards attributes are set to 0.5, since at every point during the presentation phase, information of both attributes is presented. Additionally, the model uses a parameter b_{last} to predict order effects in trials with attribute-wise presentation: Positive b_{last} shift R towards the alternative with the higher value on the last-presented attribute, negative b_{last} shift it to the alternative with the higher value on the attribute presented first.

Parameter estimation Both models were implemented in pyddm (Shinn et al., 2020) with a temporal resolution $dt = 0.01$, and a resolution of the evidence space $dx = 0.01$. Both models were fit separately to trials with alternative-wise and attribute-wise presentation using pyddm's default differential evolution algorithm, minimizing the Bayesian Information Criterion (BIC; Schwarz, 1978).

Model validation To ensure interpretability and validity of the models' parameter estimates, we performed a parameter recovery study as follows: First, we estimated each model's parameters from the empirical data of each participant. Then we simulated a synthetic data set of the same size as the empirical one, using the individually obtained estimates. Then we re-fit the models to the synthetic data and compared known generating to the obtained recovered parameters by means of Bayesian linear regression (dependent variable: recovered parameter; independent variables: Intercept, generating parameter) and correlation analyses (Kruschke, 2013; Lee & Wagenmakers, 2013). The models' parameters could be recovered to a satisfying degree, with the gaze-discount parameters θ and η showing the largest differences (Figure S2).

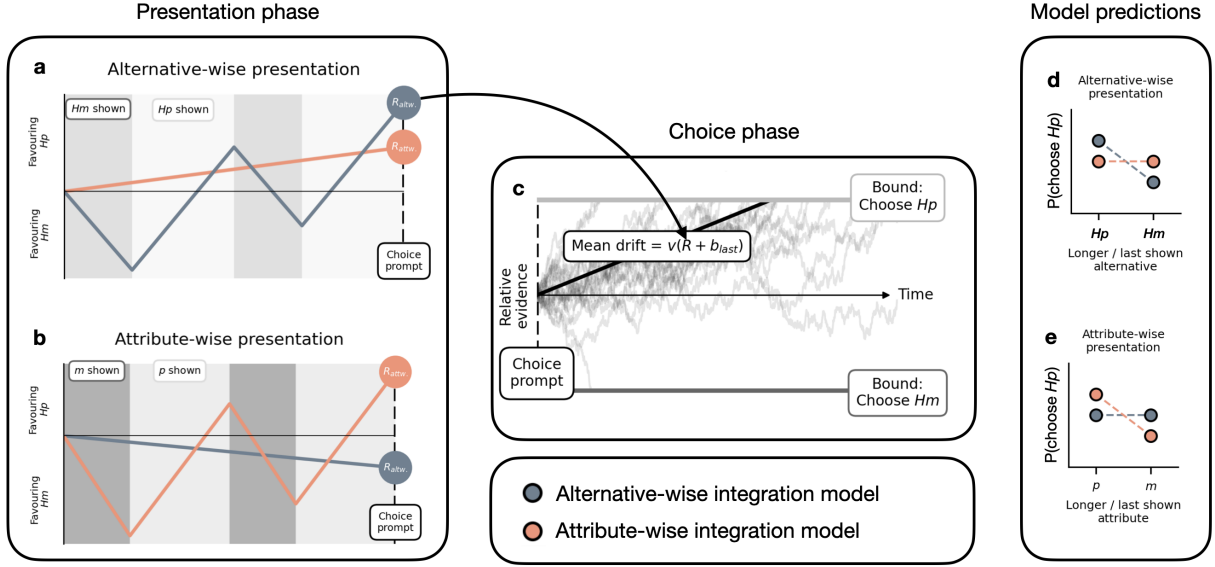


Figure S1. Behavioural models and their predictions. **a)** Construction of drift-rates in trials with alternative-wise presentation. Only the alternative-wise model's drift rate is sensitive to effects of presentation duration in alternative-wise presentation. Analogously, only the alternative-wise model produces order effects in alternative-wise presentation, controlled by the b_{last} parameter. **b)** Construction of drift rates in attribute-wise presentation. Here, only the attribute-wise model's drift rate is sensitive to effects of presentation duration. Similarly, only the attribute-wise model produces order effects in attribute-wise presentation, controlled by the b_{last} parameter. **c)** After construction of the drift rate in the presentation phase, choices and response times result from a diffusion process with the previously constructed drift rate. **d-e)** Both models' predicted effects of alternative-wise (**d**) and attribute-wise (**e**) presentation duration. Slate-gray color refers to alternative-wise model. Orange color refers to attribute-wise model.

Similarly, we performed model recovery analyses by fitting each model to the synthetic data generated from all models. Then, for each generating model, we performed Bayesian model selection (Rigoux et al., 2014; Stephan et al., 2009) to identify the most likely generating model. The models could be recovered almost perfectly (Figure S3).

Results

For trials with alternative-wise presentation, the mean \pm s.d. BIC of the alternative-wise model was 102.16 ± 56.19 (range -24.20 to 344.44). Mean \pm s.d. BIC of the attribute-wise model was 93.27 ± 57.03 (range -45.06 to 338.90), indicating a better fit of the attribute-wise model in alternative-wise presentation.

Conversely, in trials with attribute-wise presentation, the mean \pm s.d. BIC of the alternative-wise model was 90.97 ± 55.24 (range -45.39 to 310.33), while the attribute-wise model achieved mean \pm s.d. BIC of 97.41 ± 55.36 (range -50.91 to 321.62), indicating better fit of the alternative-wise model.

Next, we took the difference between BIC of the models for each participant and presentation format and tested whether they credibly differed from zero, using a directed paired Bayes factor t -test (and BEST analysis to obtain estimates of the effect size d). Results strongly supported a result *opposite* to our hypothesis, namely, that the attribute-wise model was favoured in alternative-wise presentation, while the alternative-wise model was favoured in attribute-wise presentation (mean BIC difference = -15.34 [-14.57, -16.20]; mean $d = 2.75$ [2.29, 3.23]; $BF_{+0} = 0.0011$, $BF_{0+} = 907.95$).

In sum, these results are contrary to our preregistered hypothesis: In trials with alternative-wise presentation, the simpler attribute-wise model was preferred. Vice versa, in trials with attribute-wise presentation, the simpler alternative-wise model was preferred.

We note, however, that the preregistered analysis could not address our hypothesis optimally for multiple reasons. First, model complexity differed systematically between conditions: In trials with alternative-

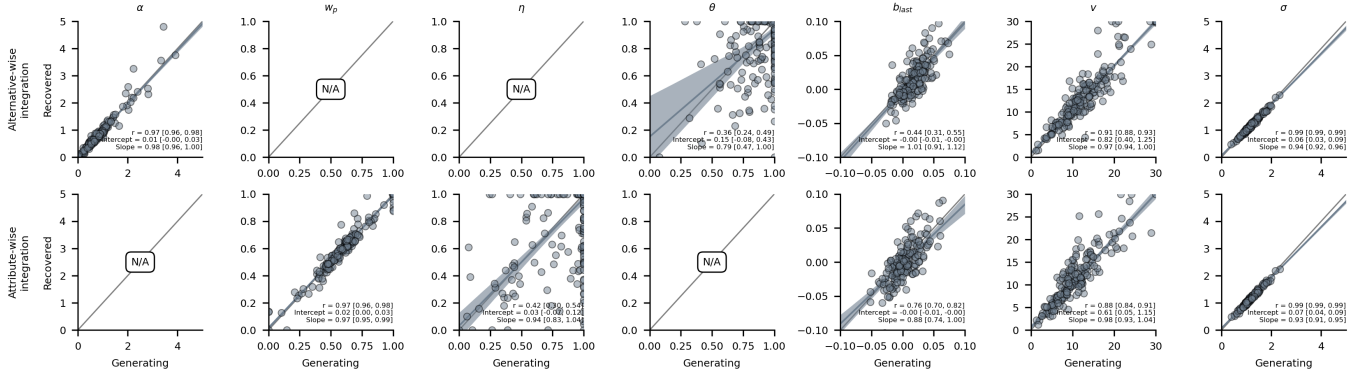


Figure S2. Parameter recovery of the two diffusion models. Each panel shows the relationship between true data generating, and recovered parameter values for a single model parameter. Upper and lower rows show results from the *alternative-wise* and *attribute-wise* integration models, respectively. Generating parameters were obtained from fitting the models to the empirical data. Annotations report Bayesian correlation coefficient r , and the slope and intercept estimates of a Bayesian regression analysis with HDI_{95%} given in brackets. Perfect, unbiased recovery would show an intercept of 0 and a slope of 1, with all points on the diagonal.

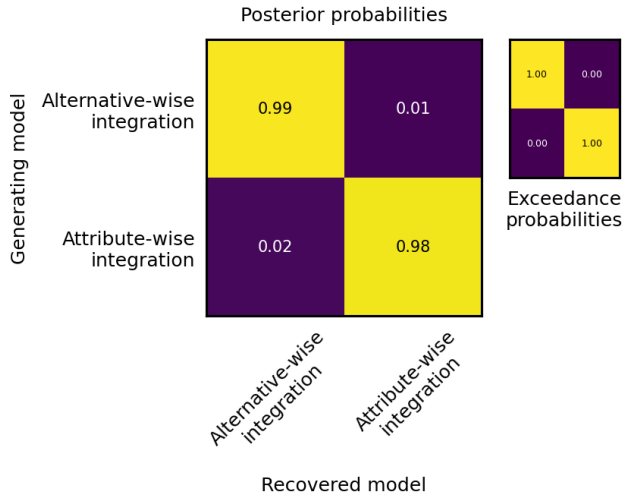


Figure S3. Model recovery results. The large panel shows a confusion matrix from the model recovery analysis. Each cell shows the posterior model probability of a fitted model (in each column) for a given generating model (in each row). Perfect recovery would show only values of 1 on the diagonal. The smaller confusion matrix shows exceedance probabilities, analogously.

wise presentation, the alternative-wise model uses two free parameters more than the attribute-wise model (namely, the alternative-wise gaze discount θ , and the alternative-wise b_{last} parameter), while the reverse is true in trials with attribute-wise presentation. While the BIC takes model complexity into account, this issue highlights that the models differ in more aspects than their value-integration process (namely, the different gaze discounts and last-stage effects), prohibiting conclusions about changes in value-integration. Furthermore, more complex models are penalized for their inclusion of duration-dependent gaze-discount mechanisms to explain associations of presentation duration and choice, which our behavioural analyses showed to be absent in our data.