Journal of Cognitive Psychology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/pecp21

Information sources and congruency modulate preference-based decision-making processes

Aysegul Ozkan & Jiaxiang Zhang

To cite this article: Aysegul Ozkan & Jiaxiang Zhang (29 Jul 2024): Information sources and congruency modulate preference-based decision-making processes, Journal of Cognitive Psychology, DOI: 10.1080/20445911.2024.2384666

To link to this article: https://doi.org/10.1080/20445911.2024.2384666

JOURNAL OF COGNITIVE PSYCHOLOGY
https://doi.org/10.1080/20445911.2024.2384666

Information sources and congruency modulate preference-based decision-making processes

Aysegul Ozkan[a,b,c] and Jiaxiang Zhang [a,d]

[a]Cardiff University Brain Research Imaging Centre, School of Psychology, Cardiff University, Cardiff, UK; [b]Max Planck University College London Centre for Computational Psychiatry and Ageing Research and the Welcome Centre for Human Neuroimaging, University College London, London, UK; [c]Department of Neuroscience, Faculty of Life and Natural Sciences, Abdullah Gul University, Kayseri, Turkey; [d]Department of Computer Science, Swansea University, Swansea, UK

ABSTRACT
Preference-based decisions often need to combine multiple pieces of information. This study investigated how the number of information sources and information congruency affect decision performance. Participants made preference-based choices between two groups of food items. Increasing the number of items in each option led to slower and less accurate decisions. Drift-diffusion modelling showed that more information sources relate to a slower rate of evidence accumulation. Therefore, the additional information impeded rather than improved the decision accuracy. In Experiment 2, each choice option contained either fully congruent information or one piece of incongruent information. Decisions with incongruent information is associated with a lower drift rate than that with congruent information, leading to inferior behavioural performance. Further model simulations support that the change in attention weighting over information sources leads to the observed effects of item numbers and item congruency. Our results suggest a bounded combination of information sources during preference-based decisions.

ARTICLE HISTORY
Received 11 October 2023
Accepted 15 July 2024

KEYWORDS
Decision-making; preference; multiple sources; cognitive modelling; drift diffusion model

1. Introduction

Rational decision-making depends on evaluating multiple options to determine the optimal outcome with the highest gain relative to potential costs. From ordinary decisions in everyday life to complex policymaking, choices are made by integrating pieces of information available to the decision-maker. Previous research has examined how people integrate different sources of information in various types of decisions (Noguchi & Stewart, 2018; Trueblood et al., 2013; Tsetsos et al., 2010; Usher & McClelland, 2004).

In simple and rapid perceptual decisions, sequential sampling models postulate that the evidence supporting each alternative is integrated over time until sufficient evidence in support of one alternative reaches a response threshold (Ratcliff et al., 2016; Ratcliff & McKoon, 2008; Smith & Ratcliff, 2004). This integration process provides an optimal strategy for fast and accurate decisions by reducing the noise in the accumulated evidence (Bogacz, 2007; Zhang & Bogacz, 2010). A large family of sequential sampling models have been proposed (Bogacz et al., 2006; Ratcliff et al., 2016). These models differed in their levels of complexity, the number of evidence accumulators, decision rules, stochastic versus deterministic evidence accumulation, and continuous versus discrete time or evidence representations. One common feature of most sequential sampling models is that they can, or at least attempt to, account for choices as well as the response time of decisions, because response time has been a key dependent variable of interest in perceptual decision research in psychology (Ratcliff, 2006) and neuroscience (Gold & Shadlen, 2007). Most literature on perceptual decisions considers the decision process involving a single source of information, although some have examined how multiple sources of information can influence behaviour (Krzemiński & Zhang, 2022; Palmer, 1995; Shaw, 1982).

Another equally fruitful line of research in psychology, marketing, political science, and behavioural

CONTACT Aysegul Ozkan a.ozkan@ucl.ac.uk; Jiaxiang Zhang jiaxiang.zhang@swansea.ac.uk Department of Computer Science, Swansea University, Swansea, UK

Supplemental data for this article can be accessed online at https://doi.org/10.1080/20445911.2024.2384666.

economics is focused on value-based decision-making. Here, we consider a specific type of value-based decision-making: preference-based decisions (also termed as preferential decisions, for reviews see Spektor et al., 2021; Yoon & Hwang, 1995). Unlike perceptual decisions, in which an objectively correct choice often exists, preference-based decisions are commonly based on subjective preference towards multiple options, where the relevant preference information comprises multiple attributes (Busemeyer et al., 2019; Slovic, 1995). For example, when renting a house, one may consider several attributes such as preferences for room size, price, and location. This raises the issue of how multiple sources of information can be integrated during preference-based decisions. Traditional decision theories suggest that multi-attribute choices are weighted and combined in a way that reaches the maximum utility (Dawes & Corrigan, 1974; Doyle, 1997; Lee & Cummins, 2004). Others have proposed alternative heuristic models, such as the take-the-best model, assuming that decision-makers focus on a few key attributes while disregarding others (Gigerenzer & Gaissmaier, 2010; Gigerenzer & Goldstein, 2011; Gigezenger & Goldstain, 1999).

Empirical studies on multi-attribute preference-based decisions have paid special attention to the contextual influence on choice. That is, when choosing between competitive options, introducing new options to the decision problem can bias the choice if the new options have similar (the similarity effect, Tversky, 1972), inferior (the attraction effect, Huber et al., 1982) or more extreme (the compromise effect, Simonson, 1989) attribute values compared with the original options. These choice context effects are crucial to our understanding of preference-based decisions because they challenge the principles underpin many normative economic choice theories such as the independence axiom (Ray, 1973).

Converging research in perceptual and value-based decision-making promotes the use of sequential sampling models to explain the choice context effects (Busemeyer et al., 2019), which has led to several model extensions specifically for preference-based decisions, including the multi-alternative decision field theory (Roe et al., 2001), the multi-alternative leaky competing accumulator model (Bogacz et al., 2007; Tsetsos et al., 2010; Usher & McClelland, 2004), the multi-attribute linear ballistic accumulator model (MLBA, Trueblood et al., 2014), the associative accumulation model (Bhatia, 2013), and the model of multialternative decision (Noguchi & Stewart, 2018). These extended models predict both choice and response time of preference-based decisions (Evans et al., 2019), providing additional insights into the experimental data (Clithero, 2018; Konovalov & Krajbich, 2019; Webb, 2019). Furthermore, these models open possibilities to using eye movement data (Krajbich et al., 2012; Krajbich & Rangel, 2011) or brain imaging (Mohr et al., 2017) to examine the evidence accumulation process during preference-based decisions.

In multi-attribute decisions, the attributes of each option commonly represent different types of information (e.g. room sizes, prices, and locations in the house renting scenario above). The current study considers a different paradigm, in which all attributes of a choice option contain the *same type* of information (Krzemiński & Zhang, 2022). For example, considering a chocolate assortment box, individual chocolate items in the box convey the same type of information: the subjective preference of individual items. To choose the box with the highest overall preference, the decision-maker needs to combine their preference towards the collection of items in the box. Interestingly, a recent study suggested that, in such a scenario, the decision-maker establishes the group of items as a set, and their preference-based decision can be influenced by the similarity of items within the group (i.e. the set-fit effect, Evers et al., 2014).

The current study builds on the existing literature. In two internet-based experiments, we examined (1) how preference-based decision is affected by the number of items of each option; and (2) whether a decision-maker is sensitive to incongruent information between items. In both experiments, human participants were instructed to make binary choices based on their preferences, whereby each choice option consisted of multiple food items (Figure 1).

Experiment 1 investigated the effect of the number of items per option on behavioural performance. Participants chose between two options at different levels of difficulty, with each containing two or four food items. Importantly, all food items assembled in each choice option were at the same level of preference rating. We hypothesised that such within-option consistency may promote two possible types of behaviour. First, as the number of items per option increases, participants may simply accumulate their preferences for all additional items to build a collective preference for each option. In this case, more items per option will lead to better behavioural performance, i.e. higher accuracy and shorter response time (RT). Alternatively, although previous research suggests that humans do combine multiple sources of information (Krajbich et al., 2012, 2010), such processes are inevitably constrained by limited attentional capacity (Reynolds & Chelazzi, 2004). As a result, more items per option introduce an attentional cost and lead to inferior behavioural

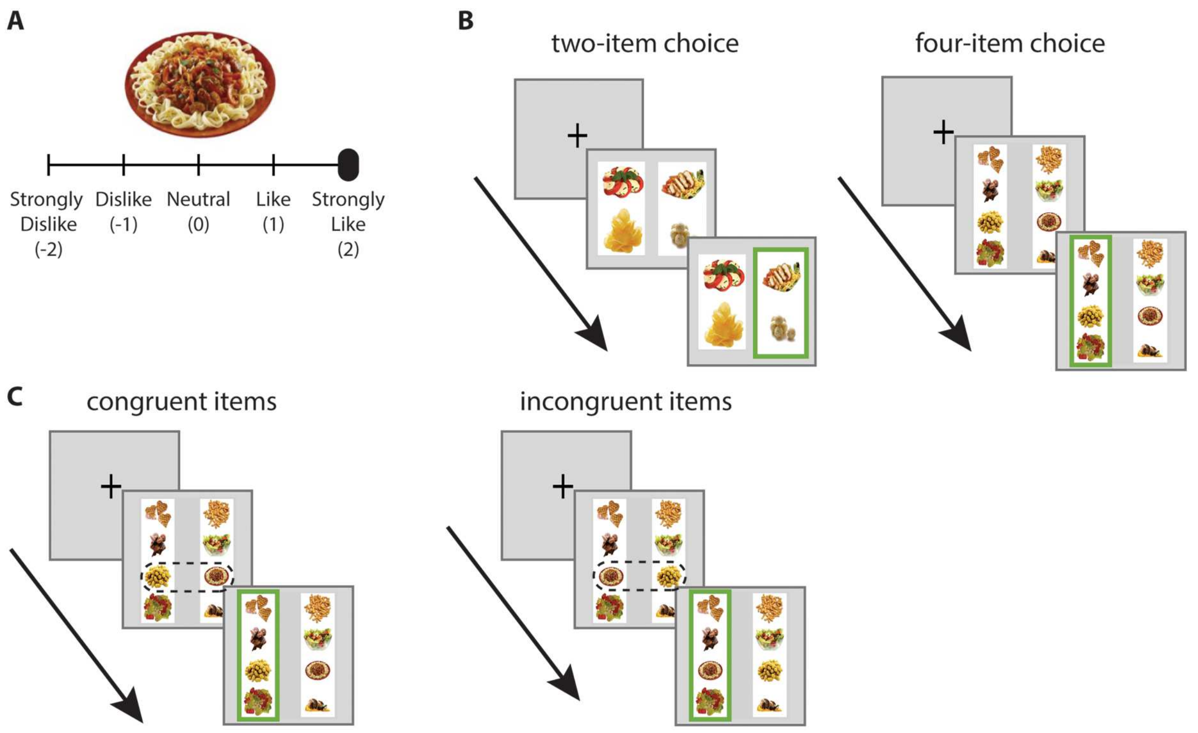

Figure 1. Experimental paradigms. A. The rating task. Participants were instructed to provide a preference rating for each food item, indicating their level of desire to consume each item. In this example, it shows that participants strongly like pasta. B. The main decision-making task in Experiment 1. In two-item trials (left), participants were asked to make a binary choice between two alternatives, each containing two food items. In four-item trials (right), each choice alternative contained four food items. All food items in each option had the same preference rating. C. The main decision-making task in Experiment 2. Participants were instructed to make a binary choice between two options, each consisting of four food items. In congruent trials, all food items in each option had the same preference rating. In incongruent trials, one pair of food items was swapped between the two choice options, introducing incongruent preference information among the items of an option.

performance, i.e. lower accuracy and prolonged RT. In both predictions, the statistical null hypothesis is the lack of behavioural change between the choices with two or four items, implying that participants ignore or are insensitive to the additional (but redundant) information.

Experiment 2 examined the effect of information congruency among items of each choice option. Participants performed preference-based decisions between two options, each with a collection of four food items. In congruent trials, all food items in each option had the same level of preference rating (same as in Experiment 1). In incongruent trials, an incongruent pair was created by swapping the locations of two food items between the two options. This design creates a scenario in which the number of items per option is the same between congruent and incongruent trials, while the summed or averaged value difference was lower in incongruent than that in congruent trials. Since the presence of the incongruent information should make the decision task more difficult, we expect incongruent trials to have inferior behavioural performance (i.e. lower accuracy and longer RT) than congruent trials.

In both experiments, we fitted a sequential sampling model, the drift-diffusion model (DDM) (Ratcliff & McKoon, 2008), to the behavioural data and inferred the effects of information sources, information congruency, and task difficulty on model parameters. Furthermore, we examined how our results can be interpreted in the context of a multi-attribute choice model, the MLBA (Trueblood et al., 2014).

2. Experiment 1: preference-based decisions with variable item counts

2.1. Participants

A total of 52 participants were recruited from an online recruitment portal Prolific (prolific.co) and took part in the experiment online. Participants' ages ranged from 19–56 years, with a median age of 24, and 16 were females. Supplementary Table 1 shows demographic information about the participants. Prolific users are aware that they participate in research studies and are compensated for their participation based on minimum payment rates (Palan & Schitter, 2018). All participants received monetary payments for their participation. Participants were not paid additional incentives based on their performance. Consent was obtained from all participants. The study was approved by the Cardiff University School of Psychology Research Ethics Committee.

2.2. Apparatus

The experiment was carried out online. Experimental scripts for stimulus presentation and response collection are written in HTML with a JavaScript library jsPsych 6.1.0 (de Leeuw, 2015). The online experiment was run on the Pavlovia web server (pavlovia.org), and participants used web browsers on their computers to complete the experiment. It has been shown that online studies using modern web browsers can be employed as an efficient tool to accurately measure behavioural responses and reaction times (Anwyl-Irvine et al., 2021; de Leeuw & Motz, 2016; Semmelmann & Weigelt, 2017).

2.3. Experimental design

All participants completed two separate experimental sessions spread over two weeks. In each session, choice options were comprised of either two or four food items. Half of the participants completed the session with two-item options first, whereas the other half completed the four-item session first. Each session included a rating part and a decision-making part.

In the rating part, A total of 100 food pictures were chosen from an online food database (Blechert et al., 2019) (See Supplementary Figure 1). Participants were asked to give a preference rating for each food item (i.e. how much they would like to consume the food item). The preference rating was on a Likert-type scale, with five discrete values from −2 to 2, representing five preference levels: “strongly dislike” (−2), “dislike” (−1), “neutral” (0), “like” (1) and “strongly like” (2). Participants were informed that they needed to rate the food items as evenly spread as possible. After the rating part, if there were fewer than 8 items at any preference level, the experiment was terminated without progressing to the main decision-making part, and the participant’s data was discarded without further analysis. A total of 59 participants were rejected on the recruitment platform after the rating part because of their biased ratings. Hereafter, we reported results from the 52 remaining participants.

In the decision-making part, participants were asked to make preference-based decisions between two options, each containing a combination of food items. They were instructed with the following phrases: “Please determine which one of the two/four item combinations you prefer more”. In each trial, two groups of food stimuli were presented vertically on the left and right sides of the screen (Figure 1(B): two-item trial; Figure 1(C): four-item trial). In both two-item and four-item trials, all food items of each option have the same level of preference rating (i.e. from −2 to 2).

For two- and four-item trials, there were four preference difference levels from 1 to 4, determined by the absolute difference in the preference ratings between the items of the two options. That is, for the preference difference of 1, the two choice options contain items rated at 0 vs. 1, 0 vs. −1, 1 vs. 2, or −1 vs. −2. For the preference difference of 4, the two choice options contain items rated at −2 vs. 2. Note that the task difficulty decreases as the preference difference between the two options increases.

2.4. Procedure

Each experimental session comprised 450 trials, which were divided into 15 blocks of 30 trials. Participants took short breaks between blocks. In each block, for each of the preference differences from 1 to 4, there were 12, 9, 6, and 3 trials, respectively. At each preference difference level, each possible pair of preference ratings was presented in an equal number of trials. The order of the task difficulty (i.e. preference difference) was randomised across blocks.

Each trial began with the presentation of a fixation point at the centre of the screen, with a uniformly distributed latency between 250 and 1500 ms. After the fixation, two choice options (each with two or four food items) appeared on the left and right sides of the screen. For each choice option, its associated food items were randomly drawn from the list that satisfies the preference rating required in that trial. Each trial was presented for a maximum of 3000 ms, during which time participants were instructed to click on one option using a mouse to indicate their decision. Immediately after each choice action, the colour of the rectangular border of the chosen option changed colour to indicate the registration of a response and the choice stimulus disappeared after the response. If participants did not respond within 3000 ms, a warning message was given, and the next trial began. The mouse position was reset to the centre of the screen after each trial.

2.5. Data analysis

As highlighted in the introduction, preference-based decisions do not always have an objectively correct choice. This applies to the current study, because the same food item may be evaluated differently between participants. Here, we use the conventional term accuracy to quantify, for each participant, the proportion of trials in which they chose the option with the higher subjective preference rating. Hence, this accuracy measure reflects to what extent participants' choices are consistent with their initial ratings.

We quantified the response time (RT) of each trial as the time between the onset of the food stimulus and the time of the behavioural response. Trials with RTs faster than 300 ms were removed to exclude fast guesses. Furthermore, we removed trials in which participants did not respond before the deadline. Together, 0.91% of all trials were discarded after pre-processing. To make group inferences on mean decision accuracy and RT, we used JASP (jasp-stats.org) to perform both frequentist and Bayesian ANOVA (Wagenmakers et al., 2018), with the difficulty level and the number of information sources as within-subject factors. In addition to the conventional frequentist ANOVA statistics, we reported the inclusion Bayes factor (BFincl) for each effect, which quantifies the evidence in the data for including the effect (Van Den Bergh et al., 2020).

2.6. Cognitive modelling of behavioural data

We used the hierarchical drift-diffusion model (HDDM) toolbox (version 0.9.8) in dockerHDDM (Pan et al., 2022) with Python version 3.8.13 to fit the DDM to each participant's response time distribution and accuracy. HDDM is a hierarchical extension of the DDM (Wiecki et al., 2013). It assumes that model parameters for individual participants are random examples drawn from group-level distributions and uses the Bayesian approach to estimate the posterior distributions of all model parameters at both individual and group levels (Wiecki et al., 2013). DDM assumes that a binary choice is made by a noisy process that accumulates information over time from a starting point until the accumulated information reaches one of two decision boundaries, corresponding to the two choice options (Ratcliff et al., 2016; Ratcliff & McKoon, 2008). When one of the boundaries is reached, a motor response is executed. The model decomposes behavioural data into four components:

- The drift rate (v) refers to the average rate of information accumulation.
- The decision threshold (a) refers to the distance between two response boundaries.
- The non-decision time (T_{er}) refers to the latencies of stimulus encoding and response execution.
- The starting point (z) refers to a priory bias toward one of the two options.

The model predicts choice probabilities and RTs (as the sum of the non-decision time and the duration of the accumulation process). The two decision boundaries correspond to the correct (i.e. consistent with preference rating) and incorrect decisions. Because we presented the position of the correct option (either left or right) randomly across trials, the starting point was fixed at 0.5 during model fitting. In addition, we included trial-by-trial variability in non-decision time s_t as a group-level parameter, which has been shown to improve the model fit to the data (Ratcliff & McKoon, 2008).

To accommodate changes in behavioural performance, one or more model parameters need to vary between conditions. We evaluated 16 variants of the DDM model with different parameter constraints (Figure 3(A)). In all model variants, we allow the drift rate v to vary between preference difference levels, as based on existing literature, the drift rate is sensitive to the task difficulty (Ratcliff & McKoon, 2008). The additional constraints of the 16 DDM variants include: (1) the drift rate v is variable or fixed between 4-item and 2-item trials; (2) the decision boundary a is variable or fixed between 4-item and 2-item trials; and (3) the non-decision time T_{er} is variable between set size (4-item vs. 2-item) and/or preference difference levels, or the non-decision time is fixed in all conditions.

To account for contaminant response, we used the mixture model, in which 5% of observations were assumed to be outliers and were not generated from the drift-diffusion process. The initial values of the sampling process were set to the maximum a-posterior value using a gradient ascent optimisation (Wiecki et al., 2013). We used the default option of informative priors in the HDDM toolbox, constraining parameter estimates within a range of plausible values.

For each model variant, we generated five independent chains of 20,000 samples from the joint posterior distribution of model parameters using Markov chain Monte Carlo (MCMC) sampling. The initial 5000 samples of each chain were discarded as burn-in to provide the stability of posterior estimates (Wiecki et al., 2013). To assess the convergence of the MCMC sampling, for each model variant, we calculated the Gelman-Rubin convergence diagnostic $\hat{R}$ (Gelman & Rubin, 1992) from the five MCMC chains, and used $\hat{R} < 1.1$ as a criterion of convergence.

We used two complementary metrics for model comparison. First, from all MCMC samples, we calculate the deviance information criterion (DIC) value (Spiegelhalter et al., 2002) of each model variant. The DIC value combines a measure of goodness-of-fit and a measure of model complexity (effective number of parameters), where lower values indicate a better model fit.

However, although the DIC is easy to calculate, it may favour more complex models because of its inaccuracy in estimating the number of parameters (Spiegelhalter et al., 2014). Second, we calculate the Pareto-smoothed importance sampling leave-one-out cross-validation (LOO-CV) deviance for each model variant. The LOO-CV is a more robust measure than DIC, with lower LOO-CV values indicating better out-of-sample predictive quality of the model (Vehtari et al., 2022, 2017). The calculation of LOO-CV requires the pointwise log-likelihood of each MCMC sample and each piece of observed data, which has high computational demands. To address this issue, we calculated the LOO-CV scores from the last 1000 samples of all MCMC chains.

We used the Bayesian hypothesis testing (Gelman et al., 2013) to make inferences between posterior parameters. Similar to previous studies (Szul et al., 2020; Zhang et al., 2016; Zhang & Rowe, 2014), we used the notion $P_{p|D}$ (stands for the posterior probability given observed data) to refer to the proportion of posteriors supporting the testing hypothesis at the group level from the Bayesian hypothesis testing. For example, to test if the DDM drift rate v in condition A is larger than that in condition B (i.e. $v_A > v_B$), we calculate the probability that the difference between the parameter's posterior distributions larger than zero ($P_{p|D} = P(v_A - v_B > 0)$). A high posterior probability indicates strong evidence in favour of the testing hypothesis. Note that $P_{p|D}$ is a continuous measure between 0 and 1. To facilitate discussion and follow the convention (Kelter, 2020), if $P_{p|D} > 0.95$, we consider that there is strong evidence to support the hypothesis of the statistical test.

2.7. Behavioural results

Participants performed binary preference-based choices between options incorporating two or four items in different sessions. Behavioural performance was quantified by accuracy (choice consistency based on participants' preference ratings) and RT.

Preference-based decisions between options with two items had significantly higher accuracy (Figure 2(A)) and faster RT (Figure 2(B)) than options with four items (accuracy: $F(1,50) = 8.340$, $p = 0.006$, $\eta_p^2 = 0.143$, BFincl = 7.981; RT: $F(1,50) = 22.468$, $p < 0.001$, $\eta_p^2 = 0.310$, BFincl = 388.032, repeated measures ANOVA).

As expected, there was a significant main effect of task difficulty (i.e. preference difference between options) in accuracy (accuracy: $F(3,150) = 217.232$, $p < 0.001$, $\eta_p^2 = 0.813$, BFincl = 7.612×10^{50}) and RT (RT: $F(3,150) = 162.672$, $p < 0.001$, $\eta_p^2 = 0.765$, BFincl = 3.836×10^{43}), with larger preference difference leading to better performance. Furthermore, the behavioural performance difference between two-item and four-item decisions became smaller as the task difficulty decreased, as indicated by a significant interaction between item numbers per option and task difficulty (accuracy: $F(3,150) = 3.226$, $p = 0.024$, $\eta_p^2 = 0.061$, BFincl = 0.899, RT: $F(3,150) = 2.940$, $p = 0.035$, $\eta_p^2 = 0.056$, BFincl = 0.944). It is worth noting that the interaction in accuracy is mainly driven by the ceiling effects in easier conditions, as there was no significant interaction in accuracy if the easiest condition (i.e. the condition with the largest preference difference) is removed in the ANVOA.

In an additional analysis, we explored whether choosing between positively rated items differed from choosing between negatively rated items. In a repeated-measures ANOVA, we included a within-subject factor, separating trials with positively rated items (also including neutral, i.e. items rated at 1/0, 2/0, and 1/2) from those with negatively rated items (−1/0, −2/0, and −1/−2). Making choices involving positively rated items was significantly faster across all difficulty levels ($F(1, 50) = 79.446$, $p < 0.001$, $\eta_p^2 = 0.614$, BFincl = 2.055×10^9). No difference was observed in decision accuracy ($F(1, 50) = 0.279$, $p = 0.6$, $\eta_p^2 < 0.006$, BFincl = 0.207).

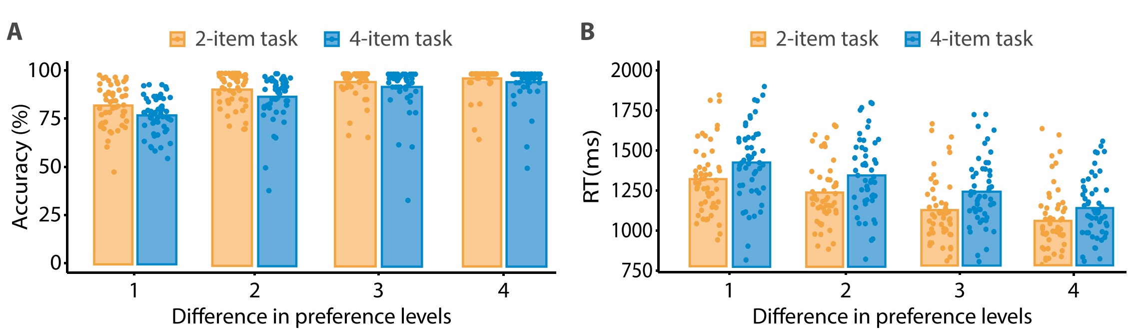

Figure 2. Behavioural results of Experiment 1. A. Mean decision accuracy (choice consistency) between the two-item and four-item conditions at each preference difference level. B. Mean RTs of the two types of choices at each preference difference level. Data points represent individual participants.

2.8. Cognitive modelling results

We used a hierarchical Bayesian version (Cavanagh et al., 2011; Vandekerckhove et al., 2011) of the DDM (Bogacz et al., 2006; Ratcliff & Tuerlinckx, 2002) to decompose individual participants' behavioural data into model parameters to infer their latent cognitive processes. We considered 16 model variants (see Figure 3(A) and *Cognitive Modelling of Behavioural Data* section for all model variants). The model variants systematically allowed model parameters (i.e. the drift-rate v, the decision threshold a, and the non-decision time T_{er}) to vary between preference difference levels (v and T_{er}), set size (v, a, and T_{er}), or both (v and T_{er}). After completing 5 chains of 20,000 samples, the Gelman-Rubin convergence diagnostic (Gelman & Rubin, 1992) is smaller than 1.1 for all parameters in all model variants, supporting that parameter estimates reached convergence.

The model variant that described the data best (i.e. the one with the lowest DIC value and the lowest LOO-CV deviance score) allows all three parameters (v, T_{er}, and a) to vary between two-item and four-item choices, and v and T_{er} to further vary between the preference difference levels (Figure 3(A) and Supplementary Figure 2A). To assess the model's fit, we simulated the model with its posterior parameter estimates. In all conditions, the observed data and model simulations were in good agreement (Figure 4).

Supplementary Table 2 reports the posterior estimates of all parameters of the best-fitted model. We used Bayesian statistics to quantify the proportion of non-overlaps between the posterior distributions of parameters (Gelman et al., 2013; Kruschke, 2011). For the drift rate, there was strong evidence to support that, at each preference difference level, the drift rate in 4-item choices was lower than that in 2-item choices ($P_{p|D} >$ 0.991 in all preference difference levels, Figure 3(B) and Supplementary Table 4). In both 4-item and 2-item choices, the drift rate increases as the preference

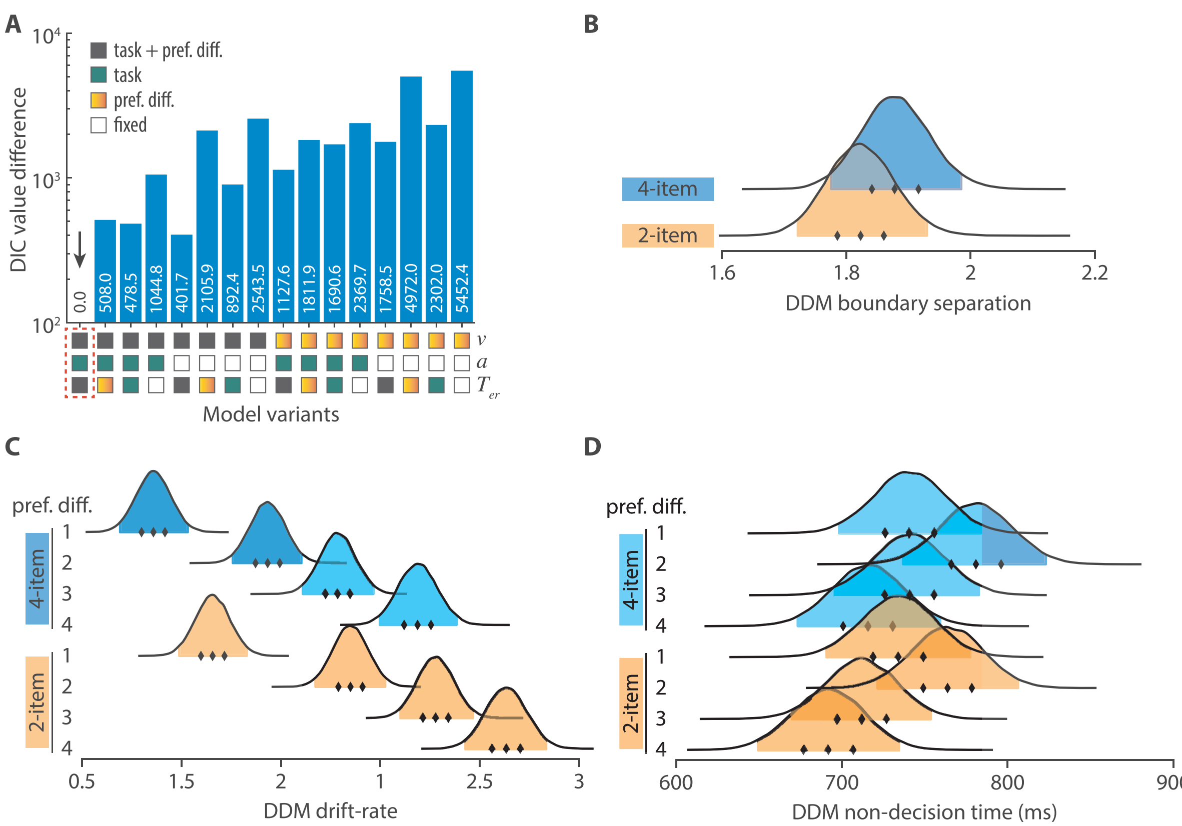

Figure 3. Model comparison and model parameters of Experiment 1. A. The deviance information criterion (DIC) value differences between model variants. The 16 models differ on whether the drift rate v, decision boundary a, and non-decision time T_{er} can vary between 4-item vs 2-item trials and between preference difference levels. The model structures are shown below the bar plot. The value under each bar indicates the DIC value difference between the model variant and the best model. The best model variant was highlighted with a red box and a black arrow. The best model with the minimum DIC value had variable drift rate and non-decision time between set size conditions and preference difference levels, as well as variable decision boundaries between set size conditions. B-D. Posterior group-level parameter estimates from the best-fitted model (B: decision boundary; C: drift rate; D: non-decision time). In each posterior estimate, the solid black line indicates the full posterior distribution. The coloured area represents 94% highest density interval. The markers indicate the 0.25, 0.5 and 0.75 quartiles.

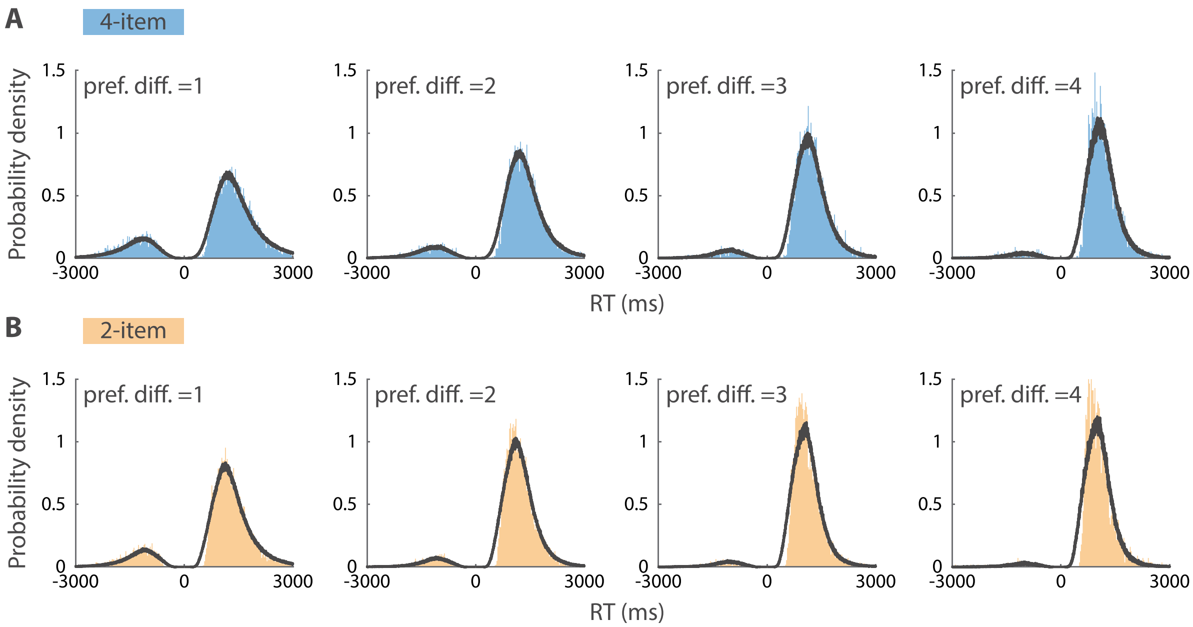

Figure 4. Posterior predictive RT distributions for 4-item (A) and 2-item (B) choices in Experiment 1. Each panel shows the normalized histograms of the observed RT distributions and the model predictions (black lines) across participants. The RT distribution plots are proportional to decision accuracy: the distribution along the positive *x*-axis indicates RTs of correct responses, and the distribution along the negative *x*-axis indicates RTs of incorrect responses. Posterior predictive were generated from the best-fit model (model 1 in Figure 3). For each participant, we drew 500 samples of all model parameters from the participant's joint posterior parameter distribution. Each parameter set of the 500 samples was used to simulate the same amount of model-predicted data as observed in the experiment. The simulated RT distributions were then used to calculate posterior model predictions across all parameter sets.

difference between the two options becomes larger ($P_{p|D} > 0.995$ in all pairwise comparisons, Supplementary Table 4).

For the non-decision time, we did not observe strong evidence supporting a difference between 4-item and 2-item choices at each level of preference difference ($P_{p|D} < 0.937$ in all comparisons, Figure 3(C) and Supplementary Table 4), nor between preference difference levels within each type of choice ($P_{p|D} < 0.908$ in all comparisons, Supplementary Table 4). For the decision boundary, there was no evidence supporting that it differs between 4-item and 2-item choices ($P_{p|D} = 0.771$ for 4-item vs 2-item comparison, Figure 3(D) and Supplementary Table 4).

3. Experiment 2: preference-based decisions with information congruency

3.1. Participants

We recruited another 52 participants from the Prolific online recruitment portal (prolific.co). Participants' ages ranged from 19 to 56 years, with a median age of 23.5 years, and 17 participants were female (Supplementary Table 1). Informed consent was obtained from all participants. The study was approved by the Cardiff University School of Psychology Research Ethics Committee.

3.2. Experimental design

Similar to Experiment 1, Experiment 2 comprised two parts: an initial rating part and a main decision-making part. The rating part was the same as in Experiment 1.

In the decision-making part, two groups of food items were presented on the left and right sides of the screen in each trial. Each group consisted of four food items (Figure 1(C)). Participants were asked to choose their preferred group of food items with the following instruction: "Please determine which one of the four food item combinations you prefer more".

Half of the decision-making trials followed a similar design as in Experiment 1: all food items in a group had the same level of preference rating (hereinafter referred to as "congruent trials"). Different from Experiment 1, Experiment 2 had three preference differences levels from 1 to 3.

In the other half of the trials, we first generated two groups of food items in the same way as the congruent trial. We then swapped the position of food items in a random row, hereafter referred to as "incongruent trials". As a result, the swapped row contains incongruent value information compared with the other rows. Participants were not informed that the food items may contain congruent and incongruent information.

3.3. Procedure

The decision-making task comprised 576 trials, including 288 congruent and 288 incongruent trials in a randomised order. The task was divided into 15 blocks of 48 trials. After each block of trials, the decision accuracy (choice consistency based on the initial preference rating) was provided on the screen. We included this feedback between blocks to help participants maintain focus on the experiment.

Each trial began with the presentation of a fixation point at the centre of the screen with a uniformly distributed latency between 250 and 1500 ms. After the fixation, two groups of food items appeared on the left and right sides of the screen until a response was received. Different from Experiment 1, participants used a keyboard (left and right arrow keys) to register their decisions, and there was no time limit for responses. The choice stimulus disappeared after the response.

3.4. Data analysis and cognitive modelling

As in Experiment 1, statistical analyses were performed on RT and accuracy (choice consistency based on participants' preference ratings). Trials with RTs faster than 300 ms were removed to exclude fast guesses. Because there was no response deadline in Experiment 2, we further removed trials with RTs longer than 10,000 ms. Together, 1.69% of all trials were discarded after pre-processing.

We fitted 8 variants of the DDM to behavioural data in Experiment 2 (Figure 6(A)). In all model variants, the drift rate can vary between preference difference levels. The additional constraints of the 8 DDM variants include: (1) the drift rate v is variable or fixed between congruent and incongruent trials; (2) the non-decision time T_{er} is variable between congruency conditions (congruent vs. incongruent) and/or preference difference levels, or the non-decision time is fixed in all conditions. We used the same mode fitting procedure, convergence check, and model comparison methods as in Experiment 1.

3.5. Behavioural results

Participants made binary preference-based choices between two groups of food items. In half of the trials, incongruent information was introduced by swapping a pair of items between the two groups (i.e. incongruent trials). Compared with congruent trials, incongruent trials had lower accuracy ($F(1,51) = 365.036$, $p < 0.001$, $\eta_p^2 = 0.877$, BFincl $= 8.385 \times 10^{21}$, repeated-measures ANOVA) and slower RT ($F(1,51) = 163.222$, $p < 0.001$, $\eta_p^2 = 0.762$, BFincl $= 4.481 \times 10^{14}$). Hence, the presence of incongruent information hinders behavioural performance (Figure 5).

We further replicated the effect of task difficulty observed in Experiment 1. Across incongruent and congruent conditions, a larger value difference was associated with higher accuracy ($F(2,102) = 281.650$, $p < 0.001$, $\eta_p^2 = 0.847$, BFincl $= 4.332 \times 10^{38}$) and faster RT ($F(2,102) = 156.963$, $p < 0.001$, $\eta_p^2 = 0.755$, BFincl $= 6.205 \times 10^{28}$). There were significant interactions between congruency and task difficulty. These results suggest that, as the task difficulty decreases, the congruency effect becomes smaller for accuracy ($F(2,102) = 4.625$, $p = 0.012$, $\eta_p^2 = 0.083$, BFincl $= 3.824$) but larger for RT ($F(2,102) = 35.399$, $p < 0.001$, $\eta_p^2 = 0.410$, BFincl $= 1.482 \times 10^9$). Similar to the Results of Experiment 1, there was no significant interaction in accuracy if the easiest condition is removed in the ANVOA. Hence, the interaction in accuracy is mainly driven by the ceiling effects in conditions with larger value differences.

Same as in Experiment 1, we explored whether choosing between positively rated items differed from choosing between negatively rated items across congruency and difficulty conditions. Making choices involving positively rated items was significantly faster ($F(1, 51) = 20.983$, $p < 0.001$, $\eta_p^2 = 0.292$, BFincl $= 615.513$). No difference was observed in decision accuracy ($F(1, 51) = 0.147$, $p = 0.703$, $\eta_p^2 < 0.003$, BFincl $= 0.212$). Therefore, we replicated the facilitation effect in RT with positive options observed in Experiment 1.

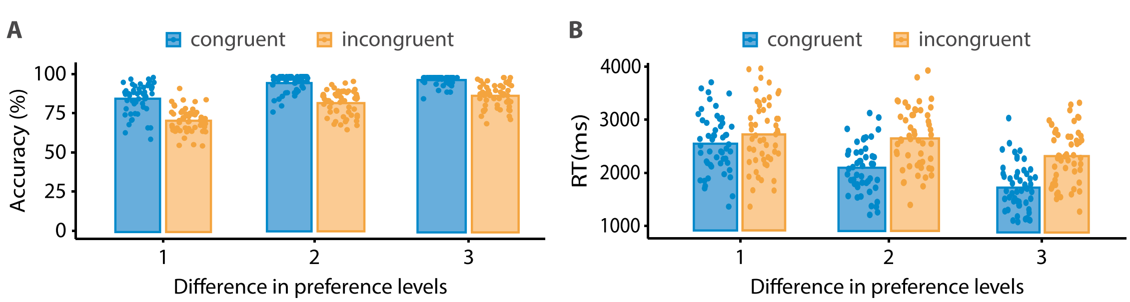

Figure 5. Behavioural results of Experiment 2. A. Mean decision accuracy (choice consistency) between the congruent and incongruent conditions at each preference difference level. B. Mean RTs of congruent and incongruent conditions at each preference difference level. Data points represent individual participants.

3.6. Cognitive modelling results

Similar to Experiment 1, the HDDM model was used to decompose each participant's behavioural data into internal components of cognitive processing. We allowed two model parameters (i.e. the drift-rate v and the non-decision time T_{er}) to be fixed or vary between preference difference levels, congruency type (incongruent or congruent options), or both. All parameters in all models converged after 5 MCMC chains of 20,000 samples (Gelman-Rubin convergence diagnostic $\hat{R} < 1.1$).

We examined both DIC values and LOO-CV scores for model comparison (Figure 6(A) and Supplementary Figure 2B). The top two model variants based on DIC values only had a small difference of 9.1 in their DIC values (the 2nd and the 4th model in Figure 6(A)). The two model variants were also the top two based on the LOO-CV score (Supplementary Figure 2B), and the one with a slightly higher DIC value (i.e. the 2nd model in Figure 6(A)) had a substantially lower LOO-CV score. Taken together, we took the model with the lowest LOO-CV score as the best model and reported below the model fit and posterior analyses.

The best model variant allows the drift rate v and the non-decision time T_{er} to vary between incongruent and congruent conditions, and T_{er} to further vary between preference difference levels. Posterior predictive RT distributions from model simulation were in good agreement with the observed data in all conditions (Figure 7).

Supplementary Table 3 and Figure 6 report the posterior estimates of all parameters of the best-fitted model. For the drift rate, there was strong evidence to support that, at each preference difference level, the drift rate in incongruent choices was lower than that in congruent choices ($P_{p|D} = 1$ in all preference difference levels, Figure 6(B) and Supplementary Table 5). In both congruent and incongruent trials, the drift rate increases as the preference difference between two options becomes larger ($P_{p|D} = 1$ in all pairwise comparisons, Supplementary Table 5).

For the non-decision time, we did not observe strong evidence supporting a difference between preference difference levels ($P_{p|D} < 0.730$ in all comparisons, Figure 6(C) and Supplementary Table 5).

4. Interpreting drift-rate changes with a multi-attribute choice model

Our modelling results suggested that increasing the number of items per option (Experiment 1), as well as introducing an incongruent pair of items (Experiment 2), led to a decrease in the drift rate, which in turn resulted in lower decision accuracy and longer RT. The DDM has been used in many decision paradigms (Ratcliff et al., 2016), including preference-based decisions. Nevertheless, the standard DDM is not specifically designed for multi-attribute decisions (cf. Dai & Busemeyer, 2014; Harris et al., 2018), or more precisely for the current study, decisions with options comprised of multiple items. Since many combinations of factors can change the drift rate of the standard DDM, it is not straightforward to interpret the drift

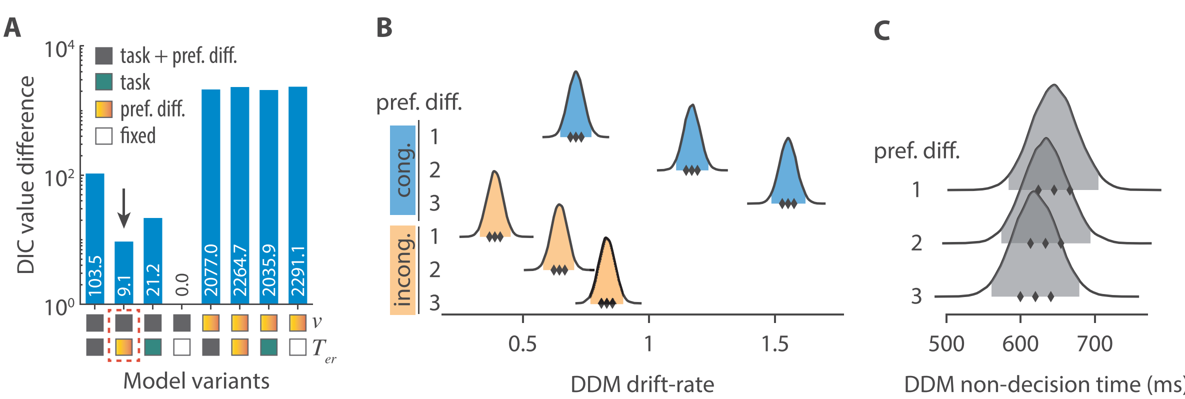

Figure 6. Model comparison and model parameters of Experiment 2. A. The deviance information criterion (DIC) value differences between model variants. The 8 models differ on whether the drift rate v and non-decision time T_{er} can vary between congruent vs. incongruent trials and between preference difference levels. The model structures are shown below the bar plot. The value under each bar indicates the DIC value difference between the model variant and the best model. We reported results from the model with the second lowest DIC value (the 2nd model from the left in panel A, marked with a red box and a black arrow) because that model also has the lowest LOO-CV score (Supplementary Figure 2B). The best model had variable drift rates between congruency conditions and preference difference levels, as well as variable non-decision times between preference levels. B-C. Posterior group-level parameter estimates from the best-fitted model (B: drift rate; D: non-decision time). In each posterior estimate, the solid black line indicates the full posterior distribution. The coloured area represents 94% highest posterior density interval. The markers indicate the 0.25, 0.5 and 0.75 quartiles.

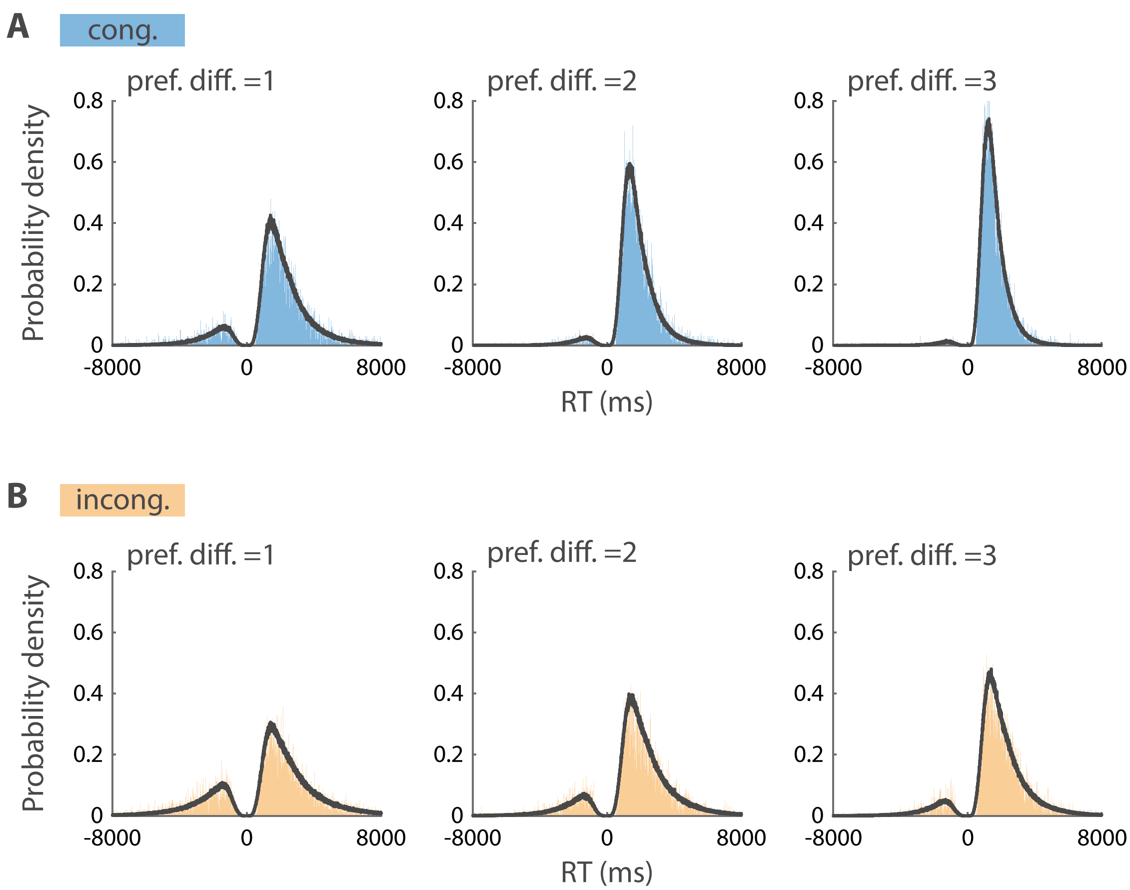

Figure 7. Posterior predictive RT distributions for choices with congruent (A) and incongruent (B) items in Experiment 2. Posterior model predictions were generated in the same way as in Figure 4.

rate change in relation to the current experimental design.

To address this issue, we attempt to interpret our results with a proper multi-attribute choice model, the MLBA (Trueblood et al., 2014). Unlike DDM, MLBA contains multiple independent linear accumulators, each representing one choice option. Each accumulator has a drift rate randomly sampled across trials, and the accumulation process is linear with no momentary noise (i.e. ballistic). The decision process is simulated as a horse race among the accumulators, from a uniformly sampled starting point to a common decision threshold (Figure 8(A)). As an extension of the original linear ballistic accumulator (LBA) model (Brown & Heathcote, 2008), MLBA in addition defines how the mean drift rates depend on the attributes of choice options.

The drift rates of DDM and LBA closely correspond to each other (Donkin et al., 2011). Hence, instead of fitting the MLBA to behavioural data, we aim to provide a synergy between models, by examining how the MLBA can reproduce the effect of item numbers and congruency on drift rates. Following the definition of MLBA (Trueblood et al., 2014), for a binary decision, the mean drift rates for the two accumulators are:.

$$\begin{cases} d_1 = I_0 + V_{12} \ , \\ d_2 = I_0 + V_{21} \ , \end{cases}$$

where I_0 is the baseline drift rate. V_{12} and V_{21} represent the comparison between the two options, which is a weighted sum of all pairwise comparisons between the n attributes:.

$$\begin{cases} V_{12} = \sum_{k=1}^{n} exp(-\lambda \cdot |P_{1,k} - P_{2,k}|) \cdot (P_{1,k} - P_{2,k}) \ , \\ V_{21} = \sum_{k=1}^{n} exp(-\lambda \cdot |P_{2,k} - P_{1,k}|) \cdot (P_{2,k} - P_{1,k}) \ . \end{cases}$$

$P_{1,k}$ and $P_{2,k}$ represent the subjective preference of the k-th attribute in options 1 and 2. The exponential term refers to the weight of attention given to each attribute comparison, and the parameter λ is the decay constant for attention weights.[1] More specifically, the attention

[1]In the full version of the MLBA, the decay constant differs between positive and negative value differences to allow for similarity asymmetry (Trueblood et al., 2014; Tversky, 1977). Here, for simplicity, we assume a single decay constant for positive and negative value differences.

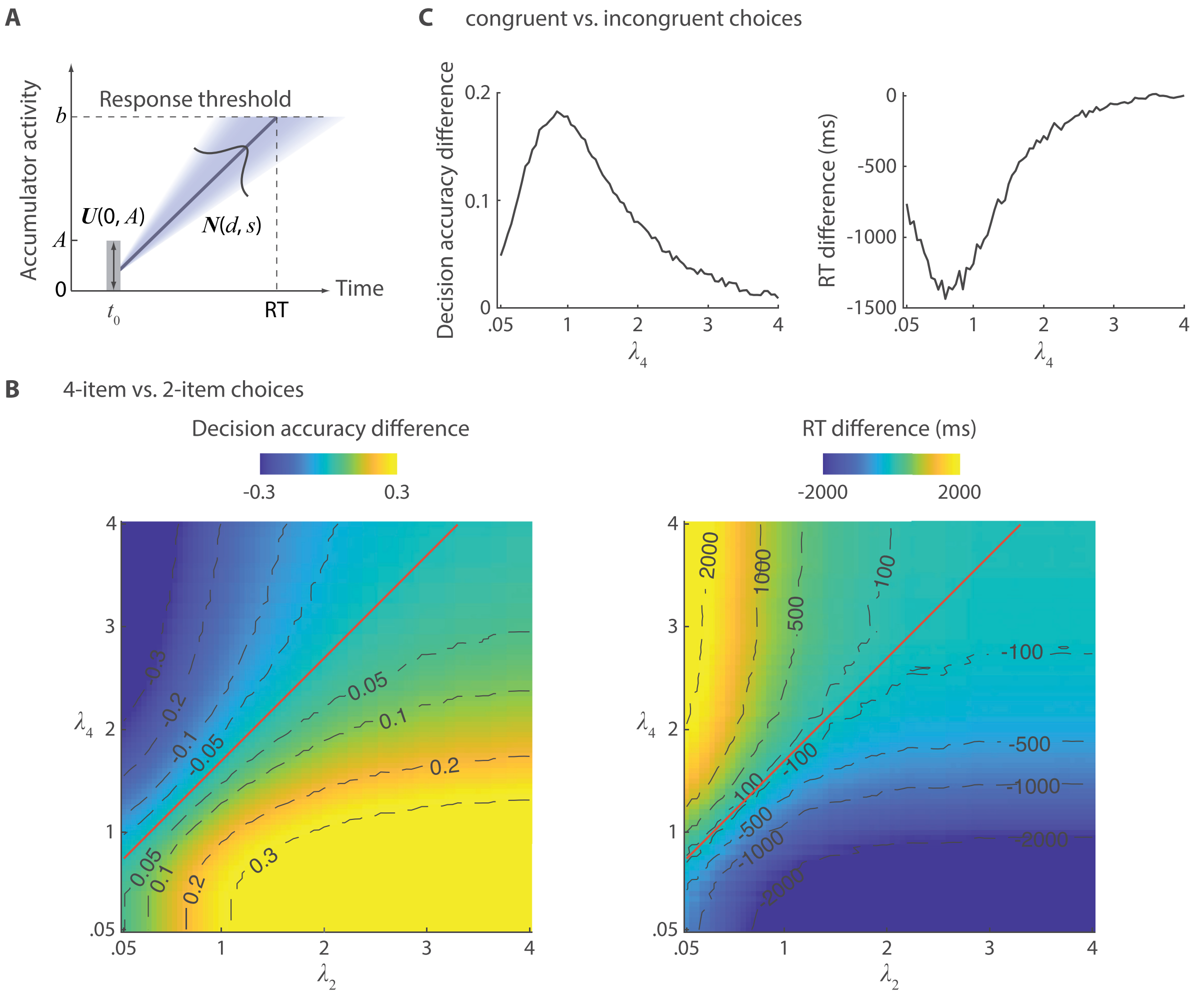

Figure 8. MLBA model simulation. A. The LBA model. On each trial, for each accumulator, the drift rate is sampled from a normal distribution with mean d and standard deviation s. The starting point is sampled from a uniform distribution between 0 and A. The accumulation process is a linear race towards the response threshold b with no noise. The sum of the accumulation time and the non-decision time t_0 is the predicted RT. B. MLBA model simulation comparing decision accuracy (left) and RT (right) difference between hypothetical 4-item and 2-item choices. The mean drift rates in 4-item and 2-item choices were set according to Equations 1 and 2, respectively. The decay constants λ_4 and λ_2 varied from 0.05 to 4 with a step size of 0.05. For each pair of λ_4 and λ_2 values, the MLBA model was simulated for 1,000,000 trials, from which the difference in decision accuracy and mean RT were calculated. Other MLBA parameters were chosen based on a previous study (Trueblood et al., 2014) and were fixed during the simulation ($b = 2$, $A = 1$, $s = 1$, $t_0 = 0.3$, $I_0 = 0.31$). The value difference between items P_{diff} was set to 1. Note that the simulation results largely depend on λ_4 and λ_2 and their difference. The red line on the contour plots indicates the critical threshold $\lambda_4 - \lambda_2 = \log(2)/P_{diff}$. For λ_4 larger than the critical value, MLBA yields consistent results observed in Experiment 1: 4-item choices have lower accuracy and longer RT than 2-item choices. The simulation results are consistent with this theoretical prediction. **C.** MLBA model simulation comparing decision accuracy (left) and RT (right) difference between hypothetical choices with congruent vs. incongruent item pairs. The decay constant λ_4 varied from 0.05–4 with a step size of 0.05. The rest model parameters were the same as in Panel B. For each λ_4 value, the mean drift rates for options including congruent and incongruent items were calculated according to Equation 3. The MLBA model was then simulated for 1,000,000 trials to obtain the difference in decision accuracy and RT. Across all λ_4 values, model simulations are consistent with the results in Experiment 2: incongruent item pairs hinder behavioural performance.

weight is larger when the attribute values of the two options are more similar, and smaller when the attribute values are more distinct. Therefore, λ controls relative attention to small vs. large differences in attribute values.

Without loss of generality, assume that the first alternative is the correct choice (i.e. the one with higher rated preference $d_1 > d_2$), and denote the preference difference between each pair of items as $P_{diff} = P_{1,k} - P_{2,k}$ ($P_{diff} > 0$). In Experiment 1, if participants accumulate the preference difference from all pairs of items, for trials with 4 items per option, the mean drift rates for the correct ($d_{1,4}$) and incorrect ($d_{2,4}$) choices are given by:.

$$\begin{cases} d_{1,4} = I_0 + 4exp(-\lambda_4 \cdot P_{diff}) \cdot P_{diff} \ , \\ d_{2,4} = I_0 - 4exp(-\lambda_4 \cdot P_{diff}) \cdot P_{diff} \ . \end{cases} \quad (1)$$

Similarly, the mean drift rates for options with 2 items ($d_{1,2}$ and $d_{1,2}$) are given by:.

$$\begin{cases} d_{1,2} = I_0 + 2exp(-\lambda_2 \cdot P_{diff}) \cdot P_{diff} \quad , \\ d_{2,2} = I_0 - 2exp(-\lambda_2 \cdot P_{diff}) \cdot P_{diff} \quad . \end{cases} \tag{2}$$

λ_4 and λ_2 represent the decay constants in 4-item and 2-item trials, respectively. Comparing Equations 1 and 2 above, it is clear that, when $\lambda_4 - \lambda_2 = \log 2 / P_{diff}$, options including 4 and 2 items have the same set of mean drift rates (i.e. $d_{1,4} = d_{1,2}$ and $d_{2,4} = d_{2,2}$). If all other model parameters remain unchanged between 4-item and 2-item trials (as observed in our DDM results), for the MLBA to reproduce the effect of item numbers in decision accuracy and RT, the decay constant λ needs to increase as the number of items per option becomes larger. In the case of options with 4 – vs. 2-item, the increase in λ must satisfy the minimum amount of $\lambda_4 - \lambda_2 > \log 2/P_{diff} > 0$.

To validate this theoretical conclusion, we vary λ_4 and λ_2 over a range of values. For each combination of λ_4 and λ_2, we used Equations 1 and 2 to define the mean drift rates and simulated the MLBA for 1,000,000 trials of 4-item and 2-item decisions. We then identify parameter regimes that qualitatively satisfy the behavioural performance observed in Experiment 1 (lower accuracy and longer RT in 4-item than 2-item trials). Model simulations confirmed our prediction (Figure 8(B)): as the number of items per option increases, if the amount of attention allocated to each item is reduced below a critical value (via the increase in the decay constant), the decision process becomes less accurate and lasts longer. It has been proposed that the attention weighting of value comparison is associated with visual fixation (Krajbich & Rangel, 2011). In the context of the current experiments, when the value difference between item pairs remains the same, increasing the set size (i.e. 4-item vs. 2-item per option) may constrain the frequency and duration of fixation on each item, which leads to reduced attentional weighting, and in turn, a more error-prone decision process.

Note that the drift rate definition in Equation 1 naturally predicts the effect of information congruency observed in Experiment 2. For the congruent condition, the mean drift rates are the same as in Equation 1. For the incongruent condition, the presence of the incongruent item pair leads to a change in the accumulated preference information, and the mean drift rates are given by:.

$$\begin{cases} d_{1,4} = I_0 + 2exp(-\lambda_4 \cdot P_{diff}) \cdot P_{diff} \quad , \\ d_{2,4} = I_0 - 2exp(-\lambda_4 \cdot P_{diff}) \cdot P_{diff} \quad . \end{cases} \tag{3}$$

Compared with Equation 1, for the incongruent condition, the accumulator representing the correct option had a lower mean drift rate (d_1), and the difference between the two accumulators was also reduced. This change in the mean drift rate will lead to lower decision accuracy and longer RT, as observed in Experiment 2 and confirmed by model simulation (Figure 8(C)).

We conducted further model simulations, in which we varied the value difference P_{diff}. Increasing P_{diff} led to a higher decision accuracy and faster mean RT (Supplementary Figure 3), in line with the observed behavioural results from different difficulty conditions in Experiments 1 and 2.

5. Discussion

In two independent experiments, we investigated how the existence of multiple information sources impacts preference-based decisions in terms of behavioural performance and its underlying cognitive mechanism. Experiment 1 investigated the impact of the number of information sources on decision-making. When the number of items in each choice option increased, human participants made slower and less accurate choices. Experiment 2 extended the main results of the first experiment. When the number of items remains the same, incongruent information among each option leads to less accurate and slower decisions. In both experiments, decisions were slower and less accurate in more difficult conditions, in which preference ratings between options were closer.

Our experimental design and procedure are similar to those used by (Philiastides & Ratcliff, 2013), who sought to identify how branding bias affects behavioural and decision processes. When making preference-based decisions between options associated with single items, they reported that behavioural performance varied according to the difference in the preference ratings of items. Instead, both experiments in the current study replicated the main finding of (Philiastides & Ratcliff, 2013), with the extension to options associated with two and four items. Taken together, these results suggest that the value difference influences both the speed and accuracy of preference-based decisions, which calls for the need for computational modelling to combine these behavioural measures.

One noteworthy addition is that our research was carried out in an online setting, suggesting the validity and reproducibility of online experiments to investigate the integration of subjective value during decision-making. When compared with trials with negatively rated items, the presence of positively rated items with the same value difference facilitates RT, but not decision accuracy. These results are akin to the effect of reward magnitude, which also demonstrates a facilitating effect on RT in probabilistic reward tasks (Chen & Kwak, 2017; Schurman & Belcher, 2013) and preference-based decisions (Shevlin et al., 2022).

As highlighted above, in Experiment 1, the number of items per option affected behavioural performance, and the negative impact of multiple information sources on accuracy is more prominent in difficult trials. Since choices with four-item options consist of more pieces of information than those with two-item options, the prolonged RT associated with four-item options may reflect the additional time required to evaluate more information sources. However, more items per option also led to less accurate decisions. This may appear to be counterintuitive, as all items within an option had the same level of subjective value (i.e. preference rating). Previous research on consumer behaviour lends conceptual support to our results (Jacoby et al., 1974a, 1974b): when consumers are provided with more information, such as more choice alternatives and attributes per choice, the increase in information load undermines their decision quality (Lurie, 2004).

Experiment 1 manipulated task difficulty by having different levels of preference difference between options. In Experiment 2, we examined the effect of task difficulty further with the additional manipulation of information congruency. In the incongruent condition, one pair of items had their value difference opposite to the rest of the item pairs, but the magnitude of their value difference was the same as the remaining pairs. As expected, in addition to the sensitivity of the average value difference between options, participants showed lower accuracy and longer RT in the incongruent than the congruent trials.

Previous studies support the integration of multiple information sources during food choices (Krajbich et al., 2010; Krajbich & Rangel, 2011). Similarly, information from different domains, such as price and preference, can jointly guide decision-making (Krajbich et al., 2012). Indeed, small attentional variations during the decision process, measured by visual fixation, impact the final choice, suggesting that people tend to consider all items when making a choice. This hypothesis is also closely linked to theoretical models of multi-attribute choice: preference formation is driven by attention switching between different attributes, as suggested by the decision field theory (Roe et al., 2001), and the value-based LCA model (Usher & McClelland, 2004).

Using a Bayesian hierarchical implementation of the DDM, our findings confirm that sequential sampling models provide a good fit for response accuracy and RT data in preference-based decisions, expanding the application of sequential sampling models (Bhatia, 2013; Krajbich & Rangel, 2011; Noguchi & Stewart, 2018; Trueblood et al., 2014; Tsetsos et al., 2010, 2012). Bayesian inferences from the best-fitted model support that the number of information sources and item congruency affect the drift rate of the DDM.

First, increasing the amount of information reduces the drift rate across all difficulty levels. In other words, as the number of information sources increases, participants would, on average, accumulate evidence at a slower pace to reach the decision threshold, and the accumulation process is more susceptible to the influence of momentary noise. The magnitude of the drift rate has been associated with the allocation of attention (Schmiedek et al., 2007). It is possible that an additional cost of attention allocation is present with more information sources (Palmer, 1995; Reynolds & Chelazzi, 2004), which in turn leads to a lowered drift rate. Our simulation of the MLBA model provides additional support to this proposition: when the number of items per option increases, if the decay constant of attention weighting in the MLBA model is increased over a critical value, the model consistently predicts hindered behavioural performance over a large range of parameter values.

Second, in Experiment 2, the incongruent condition had a lower drift rate than the congruent condition at all difficulty levels. In the incongruent condition, the four items contained conflicting information. In addition to decreasing the total value score, this conflict of information may have a distracting effect on attention. Thereby the rate of evidence accumulation was adversely affected.

This would be in accordance with the findings of a previous multi-attribute study, which investigated how differential attention to positive and negative features of a product affects purchasing decisions (Fisher, 2017). It was found that consumers give more weight to negative features than positive features in their choices, and attention is paid to negative features for a longer period during the choice process. In our case, the incongruent condition involved one non-preferred item in each choice option; hence, there may be an additional attentional cost associated with the incongruent pair during the integration of values (Fisher, 2017).

Third, in both experiments, the drift rates vary with the difference in the preference level between options (i.e. task difficulty). The easiest task (with the highest difference in preference ratings) had the highest drift rate, and the drift rate decreased as task difficulty increased. These expected results are in line with the definition of the drift rate, which represents the difference in the average evidence in favour of two choice alternatives.

Fourth, non-decision time is considered as the delay period during the decision process (Ratcliff & McKoon, 2008). Cognitive modelling often considers non-decision time to be fixed between experimental conditions within a session. Brain imaging studies suggested that the non-decision time estimated from accumulation models represents the latencies of early sensory processing (Nunez et al., 2017) and motor preparation (Karahan et al., 2019), both of which are external to the evidence accumulation process but susceptible to value-based information. Hence, we considered an extended model space, including model variants that allow variable non-decision time between conditions. In both experiments, the models with variable non-decision time produced a superior fit (as confirmed by the DIC and LOO-CV scores). However, there was no strong evidence to support that non-decision time changes between task difficulty levels or different information sources. Hence, our experimental manipulation did not influence visual encoding and motor preparation latencies during preference-based decisions.

One issue requires further consideration. Our two experiments used different response modalities and response deadlines (mouse with a 3000 ms deadline in Experiment 1, and keyboard with no response deadline in Experiment 2). Mouse response latencies are known to be longer and more variable than key presses (Gatti et al., 2024; Plant et al., 2003). With the presence of a response deadline, participants may alter their baseline level speed-accuracy trade-off, and the RT measures from mouse clicks may be further contaminated by mouse movements. In Experiment 1, all participants quickly adapted to the response deadline, because they only missed the deadline in a small fraction of trials (<1%). Nevertheless, further studies are needed to confirm whether the effect of item numbers is robust across a wide range of response deadlines, and whether participants' decision strategies are altered by their adaptation to response deadlines.

In summary, when choosing between options comprised of multiple items, both the number of information sources and the averaged value difference influence preference-based decisions. Such behavioural change relates to the quality of evidence required for rational and speeded actions, but not to the latency of sensorimotor encoding.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by HORIZON EUROPE European Research Council [grant number 716321].

ORCID

Jiaxiang Zhang http://orcid.org/0000-0002-4758-0394

References

Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2021). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods*, 1407–1425. https://doi.org/10.3758/s13428-020-01501-5

Bhatia, S. (2013). Associations and the accumulation of preference. *Psychological Review*, *120*(3), 522–543. https://doi.org/10.1037/a0032457

Blechert, J., Lender, A., Polk, S., Busch, N. A., & Ohla, K. (2019). Food-pics_extended—An image database for experimental research on eating and appetite: additional images, normative ratings and an updated review. *Frontiers in Psychology*, *10*, 307. https://doi.org/10.3389/fpsyg.2019.00307

Bogacz, R. (2007). Optimal decision-making theories: Linking neurobiology with behaviour. *Trends in Cognitive Sciences*, *11*(3), 118–125. https://doi.org/10.1016/j.tics.2006.12.006

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700–765. https://doi.org/10.1037/0033-295X.113.4.700

Bogacz, R., Usher, M., Zhang, J., & McClelland, J. L. (2007). Extending a biologically inspired model of choice: Multialternatives, nonlinearity and value-based multidimensional choice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1485), 1655–1670. https://doi.org/10.1098/rstb.2007.2059

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

Busemeyer, J. R., Gluth, S., Rieskamp, J., & Turner, B. M. (2019). Cognitive and neural bases of multi-attribute, multi-alternative, value-based decisions. *Trends in Cognitive Sciences*, *23*(3), 251–263. https://doi.org/10.1016/j.tics.2018.12.003

Cavanagh, J. F., Wiecki, T. V., Cohen, M. X., Figueroa, C. M., Samanta, J., Sherman, S. J., & Frank, M. J. (2011). Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. *Nature Neuroscience*, *14*(11), 1462–1467. https://doi.org/10.1038/nn.2925

Chen, X. J., & Kwak, Y. (2017). What makes you go faster?: The effect of reward on speeded action under risk. *Frontiers Psychology*, *8*. https://doi.org/10.3389/FPSYG.2017.01057/PDF

Clithero, J. A. (2018). Response times in economics: Looking through the lens of sequential sampling models. *Journal of Economic Psychology*, *69*, 61–86. https://doi.org/10.1016/j.joep.2018.09.008

Dai, J., & Busemeyer, J. R. (2014). A probabilistic, dynamic, and attribute-wise model of intertemporal choice. *Journal of Experimental Psychology: General*, *143*(4), 1489–1514. https://doi.org/10.1037/a0035976

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*(2), 95–106. https://doi.org/10.1037/h0037613

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

de Leeuw, J. R., & Motz, B. A. (2016). Psychophysics in a Web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods*, *48*(1), 1–12. https://doi.org/10.3758/s13428-015-0567-2

Donkin, C., Brown, S., Heathcote, A., & Wagenmakers, E.-J. (2011). Diffusion versus linear ballistic accumulation: different models but the same conclusions about psychological processes?. *Psychonomic Bulletin & Review*, *18*, 61–69. https://doi.org/10.3758/s13423-010-0022-4

Doyle, J. (1997). *Rational decision making*.

Evans, N. J., Holmes, W. R., & Trueblood, J. S. (2019). Response-time data provide critical constraints on dynamic models of multi-alternative, multi-attribute choice. *Psychonomic Bulletin & Review*, *26*(3), 901–933. https://doi.org/10.3758/s13423-018-1557-z

Evers, E. R. K., Inbar, Y., & Zeelenberg, M. (2014). Set-fit effects in choice. *Journal of Experimental Psychology: General*, *143*(2), 504–509. https://doi.org/10.1037/a0033343

Fisher, G. (2017). An attentional drift diffusion model over binary-attribute choice. *Cognition*, (168), 34–45.

Gatti, D., Marelli, M., & Rinaldi, L. (2024). Predicting hand movements with distributional semantics: Evidence from mouse-tracking. *Cognitive Science*, *48*(1), e13372. https://doi.org/10.1111/cogs.13399

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. https://doi.org/10.1201/b16018

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

Gigerenzer, G., & Gaissmaier, W. (2010). Heuristic decision making. *Annual Review of Psychology*, *62*, 451–482. https://doi.org/10.1146/ANNUREV-PSYCH-120709-145346

Gigerenzer, G., & Goldstein, D. G. (2011). Reasoning the fast and frugal way: Models of bounded rationality. *Heuristics: The Foundations of Adaptive Behavior*. https://doi.org/10.1093/acprof:oso/9780199744282.001.0001

Gigezenger, G., & Goldstain, D. G. (1999). Betting on one good reason: The take the best heuristic. - PsycNET [WWW Document]. Retrieved from September 12, 2021. https://psycnet.apa.org/record/1999-04366-004

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. https://doi.org/10.1146/annurev.neuro.29.051605.113038

Harris, A., Clithero, J. A., & Hutcherson, C. A. (2018). Accounting for taste: A multi-attribute neurocomputational model explains the neural dynamics of choices for self and others. *The Journal of Neuroscience*, *38*, 7952–7968. https://doi.org/10.1523/JNEUROSCI.3327-17.2018

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, *9*, 90–98. https://doi.org/10.1086/208899

Jacoby, J., Speller, D. E., & Berning, C. K. (1974a). Brand choice behavior as a function of information load: Replication and extension. *Journal of Consumer Research*, *1*(1), 33–42. https://doi.org/10.1086/208579

Jacoby, J., Speller, D. E., & Kohn, C. A. (1974b). Brand choice behavior as a function of information load. *Journal of Marketing Research*, *11*(1), 63–69. https://doi.org/10.1177/002224377401100106

Karahan, E., Costigan, A. G., Graham, K. S., Lawrence, A. D., & Zhang, J. (2019). Cognitive and white-matter compartment models reveal selective relations between corticospinal tract microstructure and simple reaction time. *The Journal of Neuroscience*, *39*(30), 5910–5921. https://doi.org/10.1523/JNEUROSCI.2954-18.2019

Kelter, R. (2020). Analysis of Bayesian posterior significance and effect size indices for the two-sample t-test to support reproducible medical research. *BMC Medical Research Methodology*, *20*(1), 88. https://doi.org/10.1186/s12874-020-00968-2

Konovalov, A., & Krajbich, I. (2019). Revealed strength of preference: Inference from response times. *Judgment and Decision Making*, *14*(4), 381–394. https://doi.org/10.1017/S1930297500006082

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., 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

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

Kruschke, J. (2011). Doing bayesian data analysis: a tutorial with R and BUGS. *Europe's Journal of Psychology*, *7*(4). https://doi.org/10.5964/EJOP.V7I4.163

Krzemiński, D., & Zhang, J. (2022). Imperfect integration: Congruency between multiple sensory sources modulates decision-making processes. *Attention, Perception, & Psychophysics*, *84*(5), 1566–1582. https://doi.org/10.3758/s13414-021-02434-7

Lee, M. D., & Cummins, T. D. R. (2004). Evidence accumulation in decision making: Unifying the "take the best" and the "rational" models. *Psychonomic Bulletin & Review*, *11*(11), 343–352. https://doi.org/10.3758/BF03196581

Lurie, N. H. (2004). Decision making in information-rich environments: The role of information structure. *Journal of Consumer Research*, *30*(4), 473–486. https://doi.org/10.1086/380283

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), 8942. https://doi.org/10.1038/s41598-017-06968-5

Noguchi, T., & Stewart, N. (2018). Multialternative decision by sampling: A model of decision making constrained by process data. *Psychological Review*, 125(4), 512–544. https://doi.org/10.1037/REV0000102

Nunez, M. D., Vandekerckhove, J., & Srinivasan, R. (2017). How attention influences perceptual decision making: Single-trial EEG correlates of drift-diffusion model parameters. *Journal of Mathematical Psychology*, *76*, 117. https://doi.org/10.1016/J.JMP.2016.03.003

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Palmer, J. (1995). Attention in visual search: Distinguishing four causes of a set-size effect. *Current Directions in Psychological Science*, *4*(4), 118–123. https://doi.org/10.1111/1467-8721.ep10772534

Pan, W., Geng, H., Zhang, L., Fengler, A., Frank, M., Zhang, R.-Y., & Chuan-Peng, H. (2022). A hitchhiker's guide to Bayesian hierarchical drift-diffusion modeling with dockerHDDM. https://doi.org/10.31234/osf.io/6uzga

Philiastides, M. G., & Ratcliff, R. (2013). Influence of branding on preference-based decision making. *Psychological Science*, *24*(7), 1208–1215. https://doi.org/10.1177/0956797612470701

Plant, R. R., Hammond, N., & Whitehouse, T. (2003). How choice of mouse may affect response timing in psychological studies. *Behavior Research Methods, Instruments, & Computers*, *35*(2), 276–284. https://doi.org/10.3758/BF03202553

Ratcliff, R. (2006). Modeling response signal and response time data★. *Cognitive Psychology*, *53*(3), 195–237. https://doi.org/10.1016/j.cogpsych.2005.10.002

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481. https://doi.org/10.3758/BF03196302

Ray, P. (1973). Independence of Irrelevant alternatives. *Econometrica*, *41*, 987–991. https://doi.org/10.2307/1913820

Reynolds, J. H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annual Review of Neuroscience*, *27*(1), 611–647. https://doi.org/10.1146/annurev.neuro.26.041002.131039

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionst model of decision making. *Psychological Review*, *108*(2), 370–392. https://doi.org/10.1037/0033-295X.108.2.370

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H. M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, *136*(3), 414–429. https://doi.org/10.1037/0096-3445.136.3.414

Schurman, D. L., & Belcher, J. P. (2013). Noncontingent reward magnitude effects on reaction time: A replication and extension. *Bulletin of the Psychonomic Society*, *3*(3), 104–106. https://doi.org/10.3758/BF03333408

Semmelmann, K., & Weigelt, S. (2017). Online psychophysics: reaction time effects in cognitive experiments. *Behavior Research Methods*, *49*(4), 1241–1260. https://doi.org/10.3758/s13428-016-0783-4

Shaw, M.L. (1982). Attending to multiple sources of information: I. The integration of information in decision making. *Cognitive Psychology*, *14*(3), 353–409. https://doi.org/10.1016/0010-0285(82)90014-7

Shevlin, B. R. K., Smith, S. M., Hausfeld, J., & Krajbich, I. (2022). High-value decisions are fast and accurate, inconsistent with diminishing value sensitivity. *Proceedings of the National Academy of Sciences*, *119*, e2101508119. https://doi.org/10.1073/pnas.2101508119

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, *16*(2), 158–174. https://doi.org/10.1086/209205

Slovic, P. (1995). The construction of preference. *American Psychologist*, *50*(5), 364–371. https://doi.org/10.1037/0003-066X.50.5.364

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, *27*, 161–168. https://doi.org/10.1016/j.tins.2004.01.006

Spektor, M. S., Bhatia, S., & Gluth, S. (2021). The elusiveness of context effects in decision making. *Trends in Cognitive Sciences*, *25*(10), 843–854. https://doi.org/10.1016/j.tics.2021.07.011

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *76*(3), 485–493. https://doi.org/10.1111/rssb.12062

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

Szul, M. J., Bompas, A., Sumner, P., & Zhang, J. (2020). The validity and consistency of continuous joystick response in perceptual decision-making. *Behavior Research Methods*, *52*(2), 681–693. https://doi.org/10.3758/s13428-019-01269-3

Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological Review*, *121*(2), 179–205. https://doi.org/10.1037/a0036137

Trueblood, J. S., Brown, S. D., Heathcote, A., & Busemeyer, J. R. (2013). Not just for consumers. *Psychological Science*, *24*(6), 901–908. https://doi.org/10.1177/0956797612464241

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

Tsetsos, K., Usher, M., & Chater, N. (2010). Preference reversal in multiattribute choice. *Psychological Review*, *117*(4), 1275–1291. https://doi.org/10.1037/a0020580

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, *79*(4), 281–299. https://doi.org/10.1037/h0032955

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*(4), 327–352. https://doi.org/10.1037/0033-295X.84.4.327

Usher, M., & McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychological Review*, *111*(3), 757–769. https://doi.org/10.1037/0033-295X.111.3.757

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

Van Den Bergh, D., Van Doorn, J., Marsman, M., Draws, T., Van Kesteren, E.-J., Derks, K., Dablander, F., Gronau, Q. F., Kucharský, Š., Gupta, A. R. K. N., Sarafoglou, A., Voelkel, J. G., Stefan, A., Ly, A., Hinne, M., Matzke, D., & Wagenmakers, E.-J. (2020). A tutorial on conducting and interpreting a Bayesian ANOVA in JASP. *L'Année psychologique*, *120*(1), 73–96. https://doi.org/10.3917/anpsy1.201.0073

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Vehtari, A., Simpson, D., Gelman, A., Yao, Y., & Gabry, J. (2022). Pareto smoothed importance sampling.

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E. J., van Doorn, J., Šmíra, M., Epskamp, S., Etz, A., Matzke, D., … Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, *25*(1), 58–76. https://doi.org/10.3758/s13423-017-1323-7

Webb, R. (2019). The (Neural) dynamics of stochastic choice. *Management Science*, *65*(1), 230–255. https://doi.org/10.1287/mnsc.2017.2931

Wiecki, T., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the drift-diffusion model in Python. *Frontiers in Neuroinformatics*, *7*, 14. https://doi.org/10.3389/fninf.2013.00014

Yoon, K. P., & Hwang, C.-L. (1995). *Multiple attribute decision making: An introduction*. SAGE Publications.

Zhang, J., & Bogacz, R. (2010). Optimal decision making on the basis of evidence represented in spike trains. *Neural Computation*, *22*(5), 1113–1148. https://doi.org/10.1162/neco.2009.05-09-1025

Zhang, J., Rittman, T., Nombela, C., Fois, A., Coyle-Gilchrist, I., Barker, R. A., Hughes, L. E., & Rowe, J. B. (2016). Different decision deficits impair response inhibition in progressive supranuclear palsy and Parkinson's disease. *Brain*, *139*(1), 161–173. https://doi.org/10.1093/brain/awv331

Zhang, J., & Rowe, J. B. (2014). Dissociable mechanisms of speed-accuracy tradeoff during visual perceptual learning are revealed by a hierarchical drift-diffusion model. *Frontiers Neuroscience*, 69. https://doi.org/10.3389/FNINS.2014.00069/BIBTEX