Peripheral Visual Information Halves Attentional Choice Biases

Brenden Eum[1,3], Stephanie Dolbier[2,4], and Antonio Rangel[1,5]

[1]Department of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125
[2]Department of Psychology, University of California, Los Angeles, CA, 90095
[3]Corresponding Author; beum@caltech.edu
[4]sydolbier@gmail.com
[5]arangel@caltech.edu

May 4, 2023

Abstract

A growing body of research has shown that simple choices involve the construction and comparison of values at the time of decision. These processes are modulated by attention in a way that leaves decision makers susceptible to attentional biases. Here we study the role of peripheral visual information on the choice process and on attentional choice biases. We use an eye-tracking experiment where subjects ($N = 50$ adults) make binary choices between food items that are displayed in marked screen "shelves" in two conditions: (1) where both items are displayed, and (2) where items are only displayed when subjects fixate within their shelves. We find that removing the nonfixated option approximately doubles the size of the attentional biases. The results show that peripheral visual information is crucial in facilitating good decisions, and suggest that individuals might be influenceable by settings in which only one item is shown at a time, such as e-commerce.

Keywords— decision making, attention, fixations, simple choice, neuroeconomics

Statement of Relevance

Where we look affects what we choose to consume. We find that this attentional effect on choices differs in settings when all available options are presented (e.g. supermarket shelf) versus when only one option is shown at a time (e.g. many shopping websites). Specifically, we find that when only one option is shown at a time, the chances that a consumer will select the last option that they looked at approximately doubles. We argue that this is because consumers discount the value of nonfixated options by more when only one option is shown at a time compared to when all options are presented. In practice, these results suggest that individuals might be substantially more influenceable by point-of-sale marketing in the growing domain of e-commerce than in traditional retail settings. Conceptually, it shows that peripheral visual information plays a critical role in facilitating good decisions.

Introduction

Everyday we face two different types of choice situations. Sometimes we are presented with all of the available options at once, as when we face a supermarket shelf or a buffet table. In other cases, such as many shopping websites, we are presented with one option at a time, which changes sequentially at our own pace. In both cases our overt visual attention is deployed to one option at a time. But the two situations differ on the availability of peripheral visual information about the nonfixated options, which in principle could be used to guide the choice process.

A growing number of experiments have studied the role of visual attention in simple choice and have found that increases in the relative attention received by a desirable option are associated with an increase in the frequency with which it is chosen, all else being equal (Krajbich, Armel, & Rangel, 2010; Krajbich & Rangel, 2011; Krajbich, Lu, Camerer, & Rangel, 2012; Smith & Krajbich, 2018, 2019; J. F. Cavanagh, Wiecki, Kochar, & Frank, 2014; S. E. Cavanagh, Malalasekera, Miranda, Hunt, & Kennerley, 2019; Sepulveda et al., 2020; Thomas, Molter, & Krajbich, 2021; Gluth, Spektor, & Rieskamp, 2018; Gluth, Kern, Kortmann, & Vitali, 2020; Fisher, 2017; Towal, Mormann, & Koch, 2013). Although the exact mechanism behind the attentional bias remains unknown, foveation seems to facilitate the process of value computation and integration in a way that is consistent with overweighting fixated items relative to nonfixated ones. This is formalized in the Attentional Drift-Diffusion-Model (aDDM), which is able to provide a quantitative account of the relationship between fixations, choices, and reaction times (Krajbich et al., 2010; Krajbich & Rangel, 2011; Krajbich et al., 2012; Smith & Krajbich, 2018, 2019). The aDDM predicts that choices can be biased through exogenous manipulations of relative fixation time, consistent with the findings of multiple studies (Armel, Beaumel, & Rangel, 2008; Tavares, Perona, & Rangel, 2017; Hare, Malmaud, & Rangel, 2011; Pärnamets et al., 2015; Ghaffari & Fiedler, 2018; Kunar, Watson, Tsetsos, & Chater, 2017; Peschel, Orquin, & Loose, 2019; Shimojo, Simion, Shimojo, & Scheier, 2003).

Our goal is to study the role of peripheral visual information on the choice process and on attentional choice biases. In particular, do we use the same choice algorithm when all options are presented simultaneously, as when we shop at the market, and when they are presented sequentially, as when we shop online? If not, does the absence of peripheral information change the fixation process and the magnitude of the attentional biases?

We study these questions using an eye-tracking experiment in which subjects make binary choices between foods that are displayed in marked screen "shelves" in two different conditions: (1) a simultaneous condition in which both items are displayed on the screen at the time of choice, and (2) a gaze contingent condition in which items are only displayed when subjects fixate within their shelves. Most previous studies have used choice tasks in which all options are displayed simultaneously, although a handful have used gaze-contingent presentation of stimuli (Simion & Shimojo, 2006; Folke, Jacobsen, Fleming, & De Martino, 2017; Sepulveda et al., 2020; Franco-Watkins & Johnson, 2011). However, none have compared the two situations directly, which is necessary to understand the effect of peripheral visual information on the choice process.

Based on what is known about choice in the simultaneous case, and the seemingly minor change involved in removing the nonfixated options from peripheral vision, it is natural to hypothesize that similar algorithms are at work in both conditions, albeit with some differences. In particular, the aDDM suggests two non-mutually exclusive mechanisms through which removing the nonfixated options from the visual field might affect choices. First, it might increase the overweighting of fixated relative to nonfixated items, which would result in an increased attentional bias. Second, it might change the fixation process in a way that exacerbates the attentional biases, for example by increasing the asymmetry on fixation time across options.

Understanding the role of peripheral visual information in simple choice is important for multiple reasons. First, despite the robustness of the attentional biases identified in previous work, we do not know what are the channels through which covert and overt visual attention influence decisions, nor their relative contribution to choice. Decades of work in visual attention have shown that a substantial amount of information is processed through peripheral visual attention (Carrasco, 2011; Perkovic, Schoemann, Lagerkvist, & Orquin, 2022; Wästlund, Shams, & Otterbring, 2018), which raises the puzzle of why and how fixations matter so much in economic choices, even when making decisions among familiar items. Second, the transition to e-commerce has increased the frequency with which our decisions are made in sequential presentation settings. We need to understand the impact that this has on the choice algorithms and their associated biases in order to design interfaces and nudges that enhance choice quality.

To preview the results, we find that removing the nonfixated options has little impact on the average quality of choices. However, we also find that it approximately doubles the magnitude of attentional choice biases, which make decision makers more susceptible to marketing interventions (e.g., packaging) that affect attention independently of the value of products. We also find that the removal of the nonfixated options slows down the fixation and decision process considerably, but that the impact on attentional biases is driven mostly by an increase in the tendency to overweight the value of fixated options.

All data and code are available for download at the Rangel Neuroeconomics Lab website (`www.rnl.caltech.edu`). The design and analysis plans for this study were not preregistered.

Methods

Task

We investigated the role of peripheral information about nonfixated stimuli using the task depicted in Fig. 1. Subjects made decisions in two conditions: (1) a visible condition in which both items were displayed on the screen at the time of choice, and (2) a hidden condition in which items were displayed only when subjects fixated within the location associated with the stimulus. Subjects were asked to refrain from eating for 2 h before the start of the experiment, and to refrain from eating any foods afterwards during a 1 h waiting period, except for the snack that they chose in a randomly selected trial, which was given to them at the end of the experiment.

Subjects participated in two tasks. First, they were asked to provide liking ratings for 60 snack foods available at local stores ("How much would you LIKE to eat this food?", 1 = "don't like" to 5 = "like a lot", 0.25 intervals). Each item was rated twice, in random order, using a slider bar controlled by the arrow keys, and initialized to a random location to reduce anchoring effects. We use the average of the two ratings as a measure of each item's value.

Second, subjects made choices between two food items, shown on the left and right sides of the screen, in two separate conditions: (1) a visible condition where both items were shown simultaneously, and (2) a hidden condition where items were shown only when subjects fixated within their region of interest (ROI). The ROIs were indicated in both treatments with a white box (Fig. 1). Trials started with an enforced 500 ms central fixation. Subjects indicated their choices with the left and right arrow keys, and responded at their own pace. The selected option was highlighted for 1 s and trials were separated by a 1 s blank screen. Subjects made 360 choices in the exploratory sample and 400 in the confirmatory sample, half on each experimental condition. The task was divided into 4 equal sized blocks, 2 with the hidden condition and 2 with the visible condition, in random order.

The choice pairs in the exploratory data set were randomly selected from all 60 food items. In the confirmatory sample, they were constructed as follows. We used each subject's ratings to prune the stimulus set down to the 40 food items that resulted in the most uniform distribution of ratings, in order to maximize the spread of rating differences across choice pairs. Stimuli for each trial were then randomly selected, subject to the constraint that they be used 4 times per block in the exploratory data set, and 5 times per block in the confirmatory data set. All 60 foods were shown once every 30 trials in the exploratory data set, and all 40 foods were shown once every 20 trials in the confirmatory data set.

Participants

50 subjects (mean age = 30.8, 34 female) were recruited from the Caltech and surrounding community using flyers. We pre-screened subjects for a self-reported liking for snack foods (e.g. candy and potato chips) and against requiring glasses for vision correction that might interfere with eye-tracking. Subjects were paid a $35 participation fee. The experiment was approved by Caltech's IRB.

In order to obtain high quality data, we implemented a subject filter at the data collection stage. Immediately after data collection we deleted subjects who failed any of the following criteria: (1) correlation between the two liking ratings of at least 70%, (2) mean RT in choice trials between 0.7 and 6 s, (3) probability of choosing the best item significantly different from chance (based on a binomial test), and (4) at most 10% missing fixation data. Data collection continued until 50 subjects passed the data quality criteria. The first 25 subjects were allocated to the exploratory sample, the other 25 to the confirmatory sample. The number of trials per subject, and the number of subjects, were chosen based on related studies which have shown that this sample size provide reliable estimates of the parameters and effects of interest.

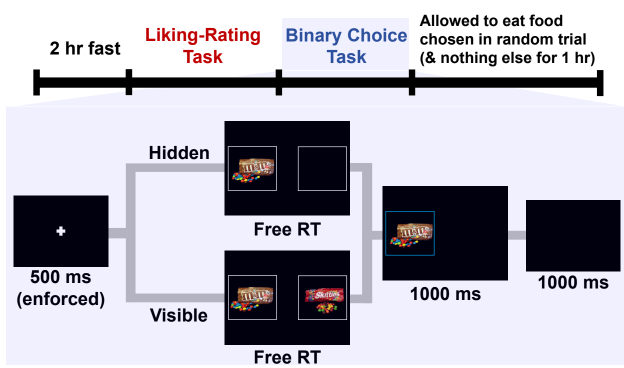


Fig. 1. Task. Subjects had to fixate on a center fixation cross for 500 ms for the trial to start. In the visible condition subjects were presented with two snack food items simultaneously, each located within a white box on the left and right sides of the screen. In the hidden condition subjects had to fixate within the white boxes in order to reveal the snack food item inside. Subjects indicated their response at their own pace with a keyboard press. Once a choice was made, a blue box highlighted the selection for 1 s, followed by a 1 s inter-trial interval.

Eye-tracking

Subjects' fixation patterns were recorded using an EyeLink 100 desk eye-tracker at 500 Hz. Subjects sat approximately 60 cm from a 1920×1080 pixel monitor. Food image sizes were 403×302 pixels. Fixations within the ROI for the left food were classified as "left", those within the right food's ROI were classified as "right", and those outside the two ROIs were classified as "blank". If a sequence of blank fixations was recorded between two fixations of the same type (e.g. left-blank-blank-left), they were re-coded as a fixation of the same type (e.g., left-left-left-left), since blank fixations of this type are typically due to eye-tracking noise and tend to be quite short. Blank fixations recorded between two fixations of different types (e.g. left-blank-right) were coded as a saccade period between fixations. Trials in which any eye-tracking information is missing are dropped from further analysis (mean of 6 and 4 trials per subject in the exploratory and confirmatory datasets, respectively).

Data analysis strategy

In order to be able to explore the data in detail, while avoiding the type of statistical problems that have raised questions about the validity of some published research, we collected two separate datasets with 25 subjects each. We used the first one to carry out exploratory analyses until we understood the data generating process in sufficient detail. Based on this, we pinned down a set of analyses and tests that were carried out in a second confirmatory dataset of equal size. Thus, the confirmatory dataset serves as a replication of our findings, and provides unbiased statistics for hypothesis testing. Given the similarity of the estimates and findings in both samples, and in the spirit of meta-analysis, we also provide results on the pooled sample and describe summary statistics in terms of the pooled estimates.

Computational model

As illustrated in Fig. 2, the Attentional Drift-Diffusion-Model (aDDM) is a version of the Drift-Diffusion-Model of binary choice (Ratcliff & McKoon, 2008; Gold & Shadlen, 2007; Ratcliff, Smith, Brown, & McKoon, 2016) in which value sampling is affected by fixation location. Subjects integrate noisy values signals into an evolving evidence process. Evidence starts every trial at an initial location b, which may include some bias towards one of the options if $b \neq 0$. A choice is made the first time evidence crosses one of two pre-specified barriers, which are fixed at 1 for the left item, and -1 for the right item. The identity of the barrier crossed determines which option is chosen. Critically, evidence evolves as the following diffusion process:

$$Evidence_t = Evidence_{t-1} + \mu_t + e_t \tag{1}$$

where e_t is i.i.d. white Gaussian noise with variance σ^2, and the slope of the process depends on the fixation location. In particular, when the left item is fixated, the slope of integration is $\mu_t = d(V_{left} - \theta V_{right})$, and when the right item is fixated is $\mu_t = d(\theta V_{left} - V_{right})$, where d is a parameter controlling the speed of integration and θ is a parameter controlling the attentional bias. When $\theta = 1$, the fixations do not affect choices, there are no attentional biases, and the process reduces to a standard DDM. In contrast, when $\theta < 1$, the value of the fixated item is overweighted relative to the nonfixated value, which results in an attentional bias that increases as θ gets smaller.

Importantly, the aDDM assumes that the fixation process is orthogonal to the state of evidence in any given trial. Thus, when simulating the model, we sample fixations from the observed fixation distributions, separately for first and middle fixations.

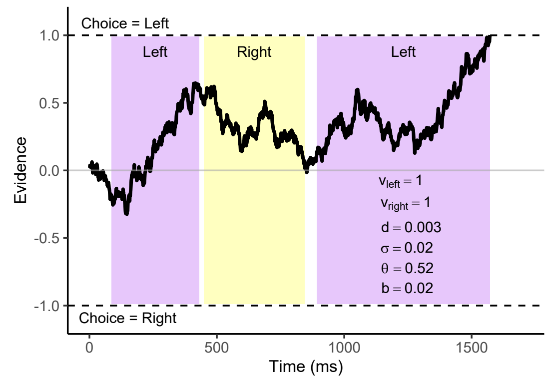


Fig. 2. aDDM Example. An illustration of how the aDDM makes decisions in a sample trial. Colored vertical bands denote fixation locations.

aDDM fitting

We fit the aDDM using a hierarchical Bayesian model, separately for the visible and hidden conditions, using the methods and associated toolbox developed by Lombardi and Hare (Lombardi & Hare, 2021). We estimate the model separately for the exploratory, confirmatory, and pooled datasets. In every case, the model is fit using only the odd trials, as the even trials are reserved for out-of-sample predictions. As described in Fig. S1, the model has the following free parameters, both at the group and individual levels: the evidence accumulation drift rate (d), the standard deviation of the Gaussian noise for the drift process (σ), the attentional bias parameter (θ), and the initial bias of the drift process (b). Posterior distributions were estimated using Markov Chain Monte Carlo methods with 3 chains for a total of 55,000 burn-in samples and 30,000 samples from each of the posteriors. Gelman-Rubin statistics for all estimates are at or below 1.1, indicating convergence.

The model that we estimate and report in the paper specifies priors without any correlation of parameters across the hidden and visible conditions. We do this to maximize the extent to which our posterior estimates are driven by the data. However, in order to investigate the role of the uncorrelated priors on our model fits, we also estimated a version of the model in which the priors for the same parameter in the visible and hidden conditions are correlated. In particular, each parameter (x) consists of two parts: a baseline ($\bar{x}$) and a hidden-condition deviation (Δx). See Fig. S2 for details. As discussed further below, both models generated very similar parameter estimates.

Out-of-sample simulations

Even-numbered trials were set aside as out-of-sample data, in order to compare them to the predictions of aDDM model fitted on the odd trials. We simulate 10 datasets for each subject and condition, using the same rating pairs encountered in the experiment. For each simulated data set we sample a set of parameters from the joint posterior distribution for that subject and condition. Then we simulate each trial as follows. We sample all fixation duration statistics from their observed empirical distributions in the even trials, conditional on the hidden or visible condition. For example, when simulating a hidden condition trial, evidence for the trial is initialized at the bias parameter and a trial-specific drift rate parameter is sampled. Evidence evolves based only on the noise up to the duration of the sampled latency to first fixation. Afterwards, a maximum first fixation duration is sampled from the distribution of first fixations in the hidden condition, and evidence evolves according to the drift rate, noise, and attentional bias parameters depending on the fixation location, as described in the Computational model section. If a barrier is crossed before the maximum fixation duration is reached, the process is terminated and the choice and RT are recorded. Otherwise, a new saccadic duration and maximum fixation duration are sampled from the distributions of saccades and middle fixations in the hidden condition, respectively. The process is repeated until a choice is made. Note that this assumes that the value of the nonfixated item is known during the first fixation, which is unrealistic and interferes with the quality of our fits.

Hierarchical regressions

All the logistic and linear regressions reported in the paper are based on standard hierarchical models with random coefficients for all parameters. The regressions are implemented using the brms R-package (Bürkner, 2017, 2018) and used the default weakly informative priors, occasionally scaled depending on the units of the independent variable. Posterior distributions were estimated using 3 chains for a total of 9,000 burn-in samples and 9,000 samples from each of the posteriors. See the companion data and code package for details (`https://www.rnl.caltech.edu/publications/`).

Results

Basic psychometrics

The top row of Fig. 3 depicts the psychometric choice curve, separately for each experimental condition and dataset. See Table S1 for the associated regression estimates and test statistics. We find a small but significant increase in the responsivity of choices to value differences in the hidden condition. The middle row of Fig. 3 depicts reaction times (RTs) as a function of choice difficulty. We find that RT increases with choice difficulty, that average RTs are about 32% (520 ms) slower in the hidden condition, and that this slowdown does not vary significantly with choice difficulty. The bottom row of Fig. 3 depicts the number of fixations as a function of choice difficulty. We find that the number of fixations increases with choice difficulty, and are approximately similar in both conditions, except for a small flattening in the slope of the fixation curve in the hidden condition.

Together, these results show that removing the nonfixated items slows down the choice process, but has a negligible effect on the quality of average choices (probability best chosen visible = 0.865 $\pm$ 0.007 ($\pm$SEs), probability best chosen hidden = 0.876 $\pm$ 0.007, d = 0.19, t(49) = 1.81, p = 0.08).

Fixation process

Fig. 4 and Table S2 explore the fixation process in more detail. The goal here is to understand the impact that removing nonfixated items has on the fixation process, which is essential to understand how it affects attentional biases.

The first row of Fig. 4 depicts the probability that the first fixation is to the best item, as a function of choice difficulty. The first fixation location is at chance in both conditions. Fig. S3 (top row) shows that there is no difference between conditions on the probability of first fixating left (probability first fix. left visible = 0.759 $\pm$ 0.042, probability first fix. left hidden = 0.803 $\pm$

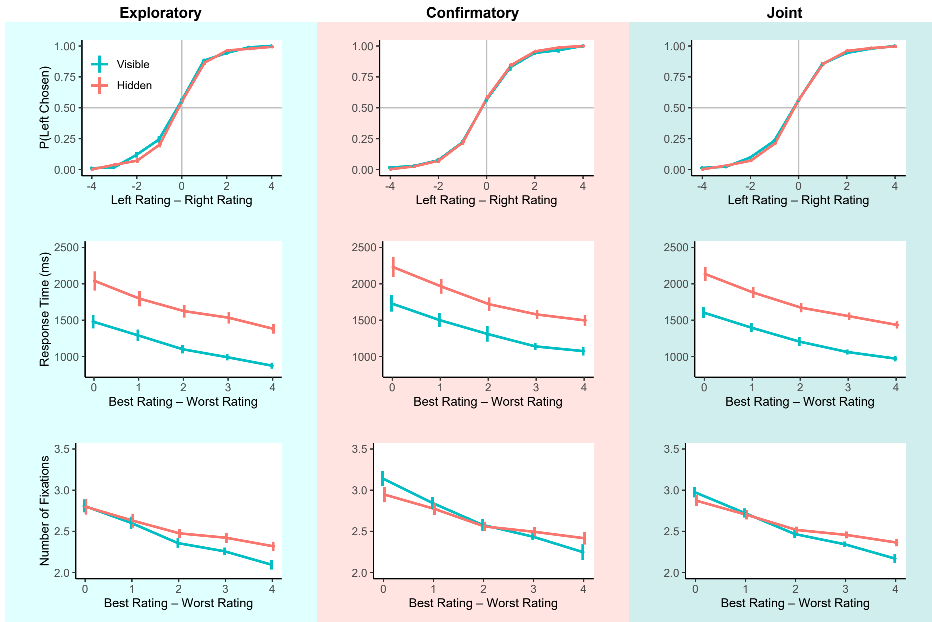


Fig. 3. Basic Psychometrics. (Top) The probability of choosing the left item as a function of its relative value. (Middle) Response time as a function of trial difficulty, as measured by the rating difference between the best and worse items. (Bottom) The number of fixations as a function of trial difficulty. Columns indicate which dataset generated the figures. Error bars show standard errors of the mean across participants.

0.048, d = 0.14, t(49) = 1.64, p = 0.11). Fig. S3 (bottom row) and Table S3 show that there is also no difference between conditions on the latency to the start of the first fixation.

The second row of Fig. 4 depicts the mean duration of first, middle, and last fixations, separately for the two conditions. We find that the three types of fixations are longer in the hidden condition by about 40% on average (Δfirst=160ms, d=1.43, t(49)=11.3, p=3e−15; Δmiddle=145ms, d=0.82, t(49)=8.51, p=3e−11; Δlast=191ms, d=1.86, t(49)=17.02, p=0). Note that this is consistent with the RT results above: an average trial has 3 fixations, and each fixation is on average 165 ms longer in the hidden condition, which implies that decisions should take 495 ms longer, just shy of the observed RT difference.

The third row of Fig. 4 depicts middle fixation durations as a function of choice difficulty. We find that middle fixation durations increase with choice difficulty. Fig. S4 and Table S4 show that this difference is driven by the value of the fixated item: in the hidden condition middle fixation durations increase with the value of the fixated item, whereas the opposite occurs in the visible condition. Interestingly, Fig. S4 also shows that middle fixation durations decrease with the value of the nonfixated item even in the hidden condition.

The fourth row of Fig. 4 depicts the first fixation duration as a function of choice difficulty. We find that duration is independent of value in both conditions, and about 46% (160 ms) longer in the hidden condition. See Fig. S4 and Table S4 for additional results.

The bottom row of Fig. 4 shows the relationship between relative value and relative fixation time. In both conditions, the relationship exhibits an S-shape. Both items are fixated the same amount when they have equal value, but otherwise the better item is fixated longer, with the asymmetry on fixation time increasing in the value advantage. In addition, the effect is stronger in the hidden condition, and as a result the distribution of net fixation times is more asymmetric in favor of the better item in this condition. Note that, since the fixated item is overweighted in the aDDM, this asymmetry in relative fixation time facilitates choosing the better option.

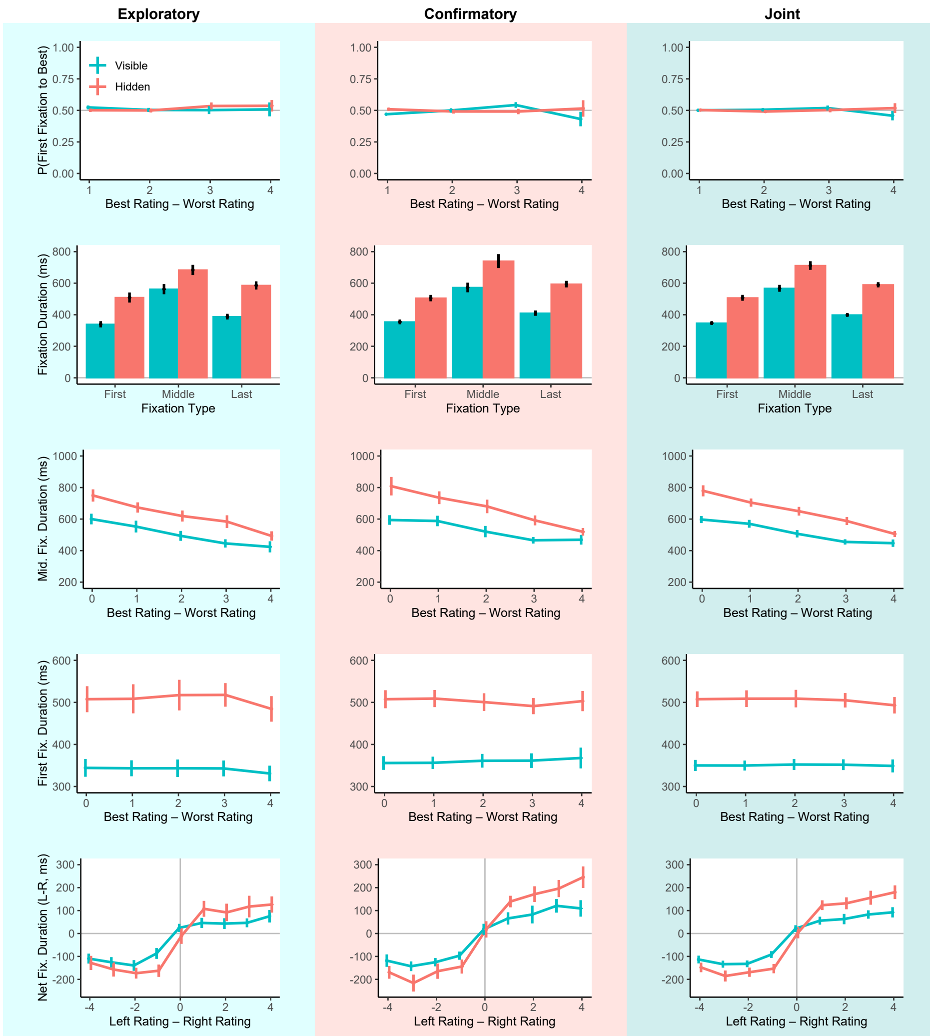


Fig. 4. Fixation Properties. (Row 1) The probability that the first fixation is to the best item as a function of choice difficulty. (Row 2) Fixation durations by fixation type. (Row 3) Middle fixation duration as a function of choice difficulty. (Row 4) First fixation duration as a function of choice difficulty. (Row 5) Net fixation duration to the left item as a function of its relative value. Columns indicate which dataset generated the figures. Error bars show standard errors of the mean across participants.

Choice biases

Fig. 5 and Table S5 depict the attentional bias in both conditions. The goal here is to provide a model-free test of the extent to which removing nonfixated items affects attentional biases.

The top row depicts the probability of choosing the left item as a function of its relative rating and the location of the last fixation. In the absence of an attentional bias, the location of the last fixation should not matter and the choice curves should lie on top of each other. In contrast, and consistent with previous studies (Krajbich et al., 2010; Krajbich & Rangel, 2011; Krajbich et al., 2012; Smith & Krajbich, 2018, 2019; Fisher, 2017; Tavares et al., 2017), we find a substantial attentional bias in the visible condition: on average, when the left and right items are equally valued, the left item is 2.5 times more likely to be chosen when the last fixation is to left than when it is to right. The bias is substantially larger in the hidden condition, where the left item is 5 times more likely to be chosen when the last fixation is to the left than when it is to the right.

The middle row depicts the relationship between net fixation time and the corrected probability of choice. The choice measure is corrected by subtracting from each choice observation (coded as 1 if left chosen, and 0 otherwise) the proportion with which left is chosen at each relative value. As a result, in the absence of an attentional bias, the corrected probability of choice should be 0, independent of net fixation time. In contrast, we find that shifting net fixation time towards the left item by 1 second increases its choice probability by 24% in both conditions.

The bottom row depicts the relationship between excess first fixation durations and the corrected choice probability of the first seen item, using the same correction described above. Excess first fixation duration is defined as first fixation duration minus mean first fixation duration (computed for each subject). In the absence of an attentional bias, the corrected probability should be 0 regardless of excess first fixation duration. In contrast, we find that an increase in the excess first fixation duration by 1 second increases the choice probability by about 22% in the visible condition, but that there is no such effect in the hidden condition.

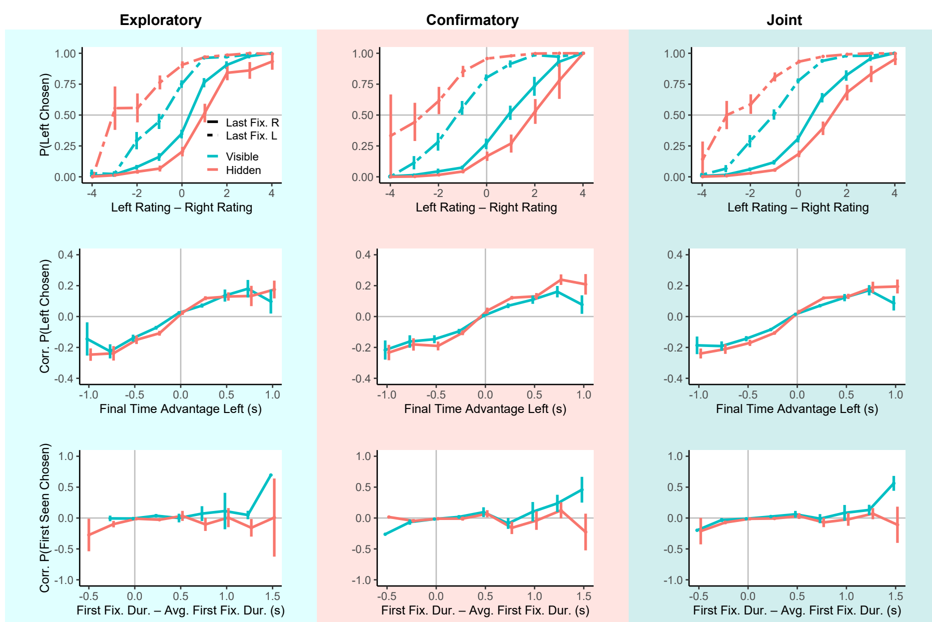


Fig. 5. Choice Biases. (Top) Probability of choosing the left item as a function of its relative value, conditional on last fixation location. (Middle) Corrected probability of choosing the left item as a function of the net fixation time to the left item. The corrected probability is computed by subtracting from each choice observation (coded as 1 if left chosen, and 0 otherwise) the proportion with which left is chosen at each relative value. (Bottom) Corrected probability that the first seen item is chosen as a function of the excess first fixation duration, defined as first fixation duration minus mean first fixation duration (computed for each subject). Columns indicate which dataset generated the figures. Error bars show standard errors of the mean across participants.

aDDM

Given that the aDDM has been shown to provide good quantitative accounts of the relationship between fixations, choices, and RTs, we fit a hierarchical approximation of this model to our data, separately for the visible and hidden conditions. The goal is to investigate the impact of removing the nonfixated items on the parameters of the aDDM and the attentional biases that they predict.

Table 1 summarizes the maximum a posteriori (MAP) estimates for group-level mean parameters. We find that $\theta^V_{group} = 0.52$ and $\theta^H_{group} = 0.29$ ($\theta^V_{group} - \theta^H_{group}$ 95% $CI = [0.12, 0.35]$) which means that the attentional bias parameter in the hidden condition worsens by a factor of two, consistent with the results described above. We also find differences in the estimated parameters for the slope ($d^V_{group} = 0.003$ vs. $d^H_{group} = 0.002$, $d^V_{group} - d^H_{group}$ 95% $CI = [0.0004, 0.0009]$), and noise ($\sigma^V_{group} = 0.022$ vs. $\sigma^H_{group} = 0.017$, $\sigma^V_{group} - \sigma^H_{group}$ 95% $CI = [0.003, 0.006]$). As shown in Table S6, the estimates using the model with correlated priors lead to very similar conclusions.

Table 1. Group-level MAP Parameter Estimates for Model with Uncorrelated Priors across Datasets and Conditions

	Exploratory		**Confirmatory**		**Joint**	
	H	V	H	V	H	V
d_{group}	0.002	0.003	0.002	0.003	0.002	0.003
	[0.002, 0.002]	[0.002, 0.003]	[0.002, 0.002]	[0.002, 0.003]	[0.002, 0.002]	[0.002, 0.003]
σ_{group}	0.018	0.023	0.016	0.021	0.017	0.022
	[0.016, 0.020]	[0.021, 0.025]	[0.015, 0.018]	[0.019, 0.023]	[0.016, 0.018]	[0.021, 0.023]
θ_{group}	0.38	0.54	0.20	0.51	0.29	0.52
	[0.19, 0.54]	[0.41, 0.67]	[0.02, 0.35]	[0.37, 0.63]	[0.17, 0.39]	[0.44, 0.61]
b_{group}	0.02	0.03	0.02	0.00	0.02	0.02
	[-0.07, 0.10]	[-0.05, 0.12]	[-0.07, 0.10]	[-0.08, 0.09]	[-0.03, 0.07]	[-0.03, 0.07]

MAP estimate and 95% HDI of group-level mean.

The hierarchical model also provides individual parameter estimates for each subject, which are shown in Fig. 6. Except for bias, the parameters in the visible condition are larger for most subjects. We estimate θ without the typical bounds at 0 and 1. In the visible condition, 0 out of 50 subject-level MAP estimates for θ fall below zero and 1 falls above one. In the hidden condition, 7 out of 50 fall below zero and 0 fall above one. However, in each of these cases, the 95% highest density intervals (HDIs) include the traditional boundaries. See Figs. S5 and S6 for a comparison of the out-of-sample predictions of the fitted model and the data in the even trials. See Figs. S7 and S8 for a comparison of the fixations and choice biases for subjects with estimated θ^H below and above zero. As shown in Fig. S9, the model with correlated priors leads to very similar individual parameter estimates.

Mechanisms of choice bias

Our results show that attentional biases are approximately twice as large in the hidden condition, and that this is accompanied by a change in fixation durations, a change in the key attentional bias parameter θ, and changes in other aDDM parameters. In this section, we use out-of-sample simulations to investigate the extent to which the attentional biases are driven by changes in the fixation process, changes in θ, or changes in non-attentional model parameters.

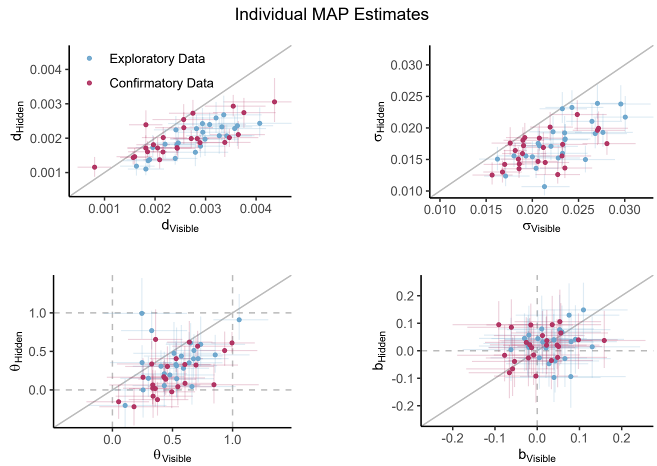


Fig. 6. Comparison of aDDM Parameter Estimates from Model with Uncorrelated Priors. Subject-level aDDM MAP parameter estimates in the visible and hidden conditions, for both the exploratory and confirmatory datasets. Colored lines denote 95% HDIs.

The simulations are shown in Fig. 7. We start the analysis by comparing the observed and simulated attentional bias in the visible condition. To do this, we simulate 10 datasets for every subject in the out-of-sample even trials, using the empirical fixation patterns from the even trials and the aDDM parameters fitted in the odd trials of the visible condition (see Methods for details). As shown in the top panel, we find a good quantitative match between the observed and simulated data.

In row 2, we repeat the exercise by changing one component of the simulations at a time. Panel b depicts data simulated using the fixations from the visible condition, but using the values of θ fitted in the hidden trials. The panel shows that this change by itself generates a good qualitative account of the increased attentional bias in hidden trials.

Panel c depicts data simulated using the parameters fitted in the visible condition but using the fixation process from the hidden condition (ΔFix.). To clarify, when we use the fixation process from the hidden condition, we mean that we are sampling properties of the fixation process (probability of first fixation to the left, latency to first fixation, first fixation duration, middle fixation duration, saccadic duration) from their empirical distributions across the hidden trials, separately for each subject. All properties of the fixation process are independently sampled once per trial, except for middle fixation durations and saccade durations, which are independently sampled until the drift diffusion process terminates. We find that this change, by itself, has a negligible impact on the attentional bias, and thus cannot account for observed data in the hidden condition.

Panel d depicts data simulated using the fixations from the visible condition, but using the values of (d, σ, b) fitted in the hidden trials. Again, this change, by itself, is unable to provide a good qualitative account of the increased attentional bias in hidden trials.

Row 3 depicts simulations in which two of the components are changed at a time. In panel e we use the θ parameters and the fixation process from the hidden condition, but the value of the other parameters are taken from the visible condition. In panel f we use the value of the other parameters (d, σ, b) and the fixation process from the hidden condition, but the value of θ is taken from the visible condition.

Finally, the bottom row depicts a simulation in which all parameters as well as fixations are taken from the hidden condition.

A comparison of these plots shows that the model can account for the large differences in attentional choice biases as long as the change in the θ parameter is taken into account, but not otherwise. This shows that the impact on attentional biases is driven mostly by an increase in the tendency to overweight the value of fixated options.

One natural concern with this simulation analysis is that changes in the θ parameter might be correlated with changes in fixation durations, across subjects. Fig. S10 shows that this is not the case.

For completeness, Fig. S11 and Table S8 show that the estimated model parameters are able to qualitatively account for the observed choice biases associated with net fixation time and excess first fixation duration (in the visible condition) in the bottom rows of Fig. 5. However, they are unable to account for the observed disappearance of excess first fixation bias in the hidden condition.

Discussion

Our experiment was designed to study the impact of peripheral visual information on the decision algorithm and its performance. Removing the nonfixated option has little impact on the quality of average choices, although it slows down the choice process by about 32% (or 520 ms). More importantly, we find that attentional choice biases are approximately twice as large when the nonfixated option is not shown.

The conclusion about the relative magnitude of the attentional biases in the two conditions is based on two different sets of analyses. A model free way of measuring the size of the attentional bias, that does not depend on the assumption that the aDDM is a good description of the data generating process (Mormann & Russo, 2021), is to ask what is the probability of choosing the last fixated item when decisions have equal value (Fig. 5, top row). In the absence of an attentional bias, both items should be chosen with equal probability. In contrast, the last seen item is 2.5 times more likely than the other item to be chosen when all items are shown simultaneously, and 5 times more likely when nonfixated items are hidden. Another way of measuring the attentional bias is based on the aDDM. In this model, the value of nonfixated options at any given time is downweighted by a parameter θ. When $\theta = 1$, there is no attentional bias. When $\theta < 1$, there is an attentional bias in favor of the fixated item, which is stronger for lower values of θ. Our mean estimates are $\theta = 0.52$ when all items are shown and $\theta = 0.29$ when nonfixated options are hidden. In both cases, the results show that removing peripheral visual information doubles the size of the attentional biases.

We find that middle fixations slow down by about 23% and first fixations slow down about 46% in the hidden condition, independently of the stimuli's value. There are two natural hypotheses for this change. One hypothesis, based on bottom-up control of the fixation process, is that the removal of peripheral stimuli changes the priority map that controls fixation durations and locations (Itti & Koch, 2000; Towal et al., 2013). This is consistent with findings from the visual search literature, which have found that a decrease in the saliency of peripheral stimuli, of which removal is an extreme case, increases fixation durations (Machner et al., 2020), as well as with the finding that fixation durations increase in patients with hemispatial neglect (Machner et al., 2012). An alternative hypothesis, based on top-down control of the fixation process, is that fixations slow down to accommodate the increased difficulty of generating value samples for the nonfixated stimuli in the absence of peripheral visual information.

Beyond showing that attentional choice biases increase substantially when only one item is shown at a time, our findings also provide some novel clues about the mechanisms at work in

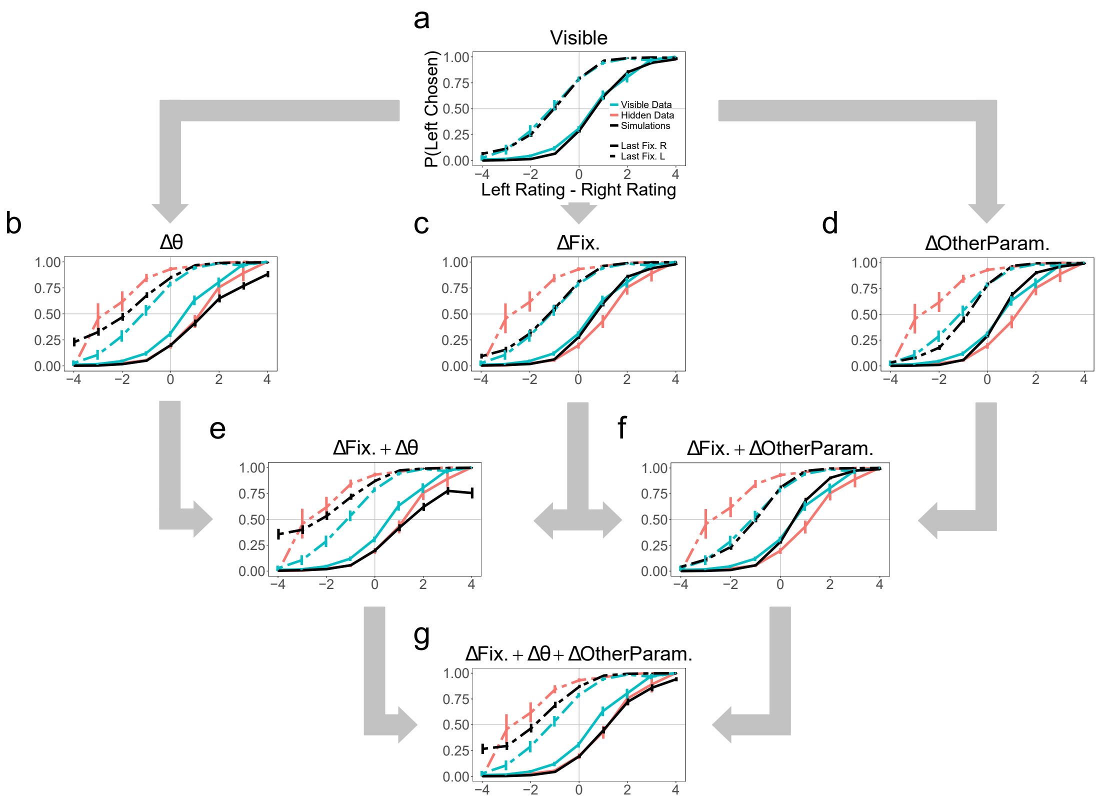


Fig. 7. Mechanisms of Choice Bias. Probability of choosing the left item, as a function of its relative value and last fixation location, in observed (blue, red) and simulated (black) data. Each panel differs on the assumptions that were used to simulate the data. (a) Simulated data for out-of-sample even trials use the empirical fixation patterns and MAP parameters fitted in the visible condition. (b) Simulated data now uses the θ MAP parameter fitted out-of-sample in the hidden condition. (c) Simulated data uses the empirical fixation patterns from the hidden condition (ΔFix.). (d) Simulated data uses the (d, σ, b) MAP parameters fitted from the hidden condition. (e) Simulated data now uses the empirical fixation patterns and the θ MAP parameter from the hidden condition. (f) Simulated data uses the empirical fixation patterns and the (d, σ, b) MAP parameters fitted from the hidden condition. (g) Simulated data now uses fixations and all parameters from the hidden condition. The figures show that the simulations provide a good qualitative match for the difference between the visible and hidden conditions when the attentional bias parameter is modified, but not otherwise. The simulations include 10 observations per trial, per subject.

simple choice.

First, we find that in the absence of peripheral stimuli the attentional bias parameter (θ) is greater than zero on average, which means that the values of nonfixated items are still being processed by some, even if they are underweighted. This suggests that foveation facilitates the extraction of value samples, but that it is not necessary, at least after the second fixation when the identity of both stimuli becomes known. This also implies that covert attention is paid to the nonfixated item, at least after the second fixation. In fact, one interpretation of our results is that removing the nonfixated item reduces the amount of covert attention that it receives (see Carrasco (2011) for an outstanding review of the role of covert visual attention).

Second, the estimated parameters in the visible condition, and specifically the attentional bias parameter, are consistent with related literature. When fitting the aDDM to individuals in their dataset, Krajbich, Armel, and Rangel (Krajbich et al., 2010) found that the average value of attentional bias among subjects was $\theta = 0.52 \pm 0.3$, although their best-fitting model had $\theta = 0.3$. Our estimates of attentional bias in both conditions are similar to the estimated influences of gaze on choice in Weilbächer and co-authors' study (Weilbächer, Krajbich, Rieskamp, & Gluth, 2021) where all options were hidden and had to be recalled from memory at the time of decision-making (attentional discounting parameter: hidden mean = 0.12, visible mean = 0.42). Interestingly, our estimate of θ in the hidden condition is larger than the one in the Weilbächer study, which suggests that the attentional bias is stronger when all information about the choice stimuli has to be recalled from memory (in their study) than when it is available conditional on foveation (as in the hidden condition in this study).

Third, Bayesian models of information sampling in simple choice have proposed that fixations matter because they control which value samples are obtained, and that samples matter because they shift the value estimates from a common initial prior to posteriors that are closer to the true value of each stimulus. As a result, the value estimates of better-than-average items tend to increase with additional fixation time, and the opposite is true for worse-than-average items (Armel & Rangel, 2008; Callaway, Rangel, & Griffiths, 2021; Jang, Sharma, & Drugowitsch, 2021; Li & Ma, 2021). This Bayesian perspective could account for the increased attentional bias when nonfixated items are hidden. Value samples must taken in parallel from both choice options, and either the rate of sampling must be slower, or the sampled information must be noisier for the nonfixated item. These variations should be even more exaggerated when nonfixated items are not present in peripheral vision. Existing Bayesian models do not account for the former, though they do account for the latter (Jang et al., 2021).

Finally, our results also have implications for the growing field of choice architecture, which seeks to understand how seemingly minor changes in the choice environment affect decisions, and how to apply this information to help individuals make better decisions (Johnson et al., 2012). We find substantially larger attentional biases in settings where only one option is shown at a time – as is done on many shopping websites – than in settings where all options are presented simultaneously – such as supermarket shelves. This suggests that individuals might be more susceptible to marketing influences that attract attention (e.g., salient packaging or point-of-sale ads) in the growing domain of e-commerce than in traditional retail settings. Although our experiments only measure the effect of removing peripheral stimuli, similar issues could arise in contexts where choice options are described sequentially using other sensory modalities (e.g., when a waiter describes the menu specials). Extrapolating from our results, we also hypothesize that similar increases in attentional biases could be induced simply by increasing the spatial separation between stimuli, so that it becomes difficult to process nonfixated options using peripheral vision. Consistent with this hypothesis, others have found that subjects with a narrower spatial attention tend to exhibit larger attentional choice biases than those with broader spatial attention (Smith & Krajbich, 2018).

Several aspects of our study might limit the generalizability of the findings. First, our results are limited to the context of binary choice, whereas in many decision contexts more than two options are available for selection. The impact of peripheral visual information on the choice process might depend on the complexity of the environment. Second, based on previous work, we use food stimuli as a basis for understanding attentional effects on value-based choices (Krajbich et al., 2010; Krajbich & Rangel, 2011). However, it is possible that the quantitative influence of peripheral information might depend on the nature of the stimuli (e.g. lotteries, toys, concert tickets), especially if it differs on how easily it can be processed in peripheral vision. Third, in the real-world, it may be more costly and slower for consumers to switch between different options than it is in our simple gaze-contingent paradigm. For instance, consumers may have to walk between two different shelves at a super market or click through a list online; whereas, in our paradigm, they simply need to fixate between two regions of interest. The impact of such variables in choices needs further study.

Author Contributions A.R. designed research; S.D. collected data; B.E. analyzed data; and B.E. and A.R. wrote the paper.

Funding Financial support from the NOMIS Foundation is gratefully acknowledged.

Competing Interests The authors declare no competing interests.

References

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

Armel, K. C., & Rangel, A. (2008). The impact of computation time and experience on decision values. *American Economic Review*, *98*, 163-168. doi: 10.1257/aer.98.2.163

Bürkner, P.-C. (2017). brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software, Articles*, *80*(1), 1-28. doi: 10.18637/jss.v080.i01

Bürkner, P. C. (2018). Advanced bayesian multilevel modeling with the r package brms. *R Journal*, *10*, 395-411. doi: 10.32614/rj-2018-017

Callaway, F., Rangel, A., & Griffiths, T. L. (2021). Fixation patterns in simple choice reflect optimal information sampling. *PLOS Computational Biology*, *17*, e1008863. doi: 10.1371/JOURNAL.PCBI.1008863

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*, 1484-1525. doi: 10.1016/j.visres.2011.04.012

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*, 1476–1488. doi: 10.1037/a0035813

Cavanagh, S. E., Malalasekera, W. M., Miranda, B., Hunt, L. T., & Kennerley, S. W. (2019). Visual fixation patterns during economic choice reflect covert valuation processes that emerge with learning. *Proceedings of the National Academy of Sciences of the U.S.A.*, *116*, 22795-22801. doi: 10.1073/pnas.1906662116

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

Folke, T., Jacobsen, C., Fleming, S. M., & De Martino, B. (2017). Explicit representation of confidence informs future value-based decisions. *Nature Human Behaviour*, *1*, 1-8. doi: 10.1038/s41562-016-0002

Franco-Watkins, A. M., & Johnson, J. G. (2011). Applying the decision moving window to risky choice: Comparison of eye-tracking and mouse-tracing methods. *Judgment and Decision Making*, *6*, 740–749.

Ghaffari, M., & Fiedler, S. (2018). The power of attention: Using eye gaze to predict other-regarding and moral choices. *Psychological Science*, *29*, 1878-1889. doi: 10.1177/0956797618799301

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

Gluth, S., Spektor, M. S., & Rieskamp, J. (2018). Value-based attentional capture affects multi-alternative decision making. *eLife*, *7*, e39659. doi: 10.7554/eLife.39659

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*, 535-574. doi: 10.1146/annurev.neuro.29.051605.113038

Hare, T., Malmaud, J., & Rangel, A. (2011). Focusing attention on the health aspects of foods changes value signals in vmpfc and improves dietary choice. *Journal of Neuroscience*, *31*, 11077-11087. doi: 10.1523/JNEUROSCI.6383-10.2011

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489-1506. doi: 10.1016/S0042-6989(99)00163

-7

Jang, A. I., Sharma, R., & Drugowitsch, J. (2021). Optimal policy for attention-modulated decisions explains human fixation behavior. *eLife*, *10*, e63436. doi: 10.7554/eLife.63436

Johnson, E. J., Shu, S. B., Dellaert, B. G., Fox, C., Goldstein, D. G., Häubl, G., … Weber, E. U. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters*, *23*, 487–504. doi: 10.1007/s11002-012-9186-1

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*, 1292-1298. doi: 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. doi: 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 of the U.S.A.*, *108*, 13852-13857. doi: 10.1073/pnas.1101328108

Kunar, M. A., Watson, D. G., Tsetsos, K., & Chater, N. (2017). The influence of attention on value integration. *Attention, Perception, and Psychophysics*, *79*, 1615-1627. doi: 10.3758/s13414-017-1340-7

Li, Z.-W., & Ma, W. J. (2021, 08). An uncertainty-based model of the effects of fixation on choice. *PLOS Computational Biology*, *17*(8), 1-17. doi: 10.1371/journal.pcbi.1009190

Lombardi, G., & Hare, T. (2021). *Piecewise constant averaging methods allow for fast and accurate hierarchical bayesian estimation of drift diffusion models with time-varying evidence accumulation rates.* Unpublished manuscript. doi: 10.31234/osf.io/5azyx

Machner, B., Dorr, M., Sprenger, A., von der Gablentz, J., Heide, W., Barth, E., & Helmchen, C. (2012). Impact of dynamic bottom-up features and top-down control on the visual exploration of moving real-world scenes in hemispatial neglect. *Neuropsychologia*, *50*, 2415-2425. doi: 10.1016/j.neuropsychologia.2012.06.012

Machner, B., Lencer, M. C., Möller, L., von der Gablentz, J., Heide, W., Helmchen, C., & Sprenger, A. (2020). Unbalancing the attentional priority map via gaze-contingent displays induces neglect-like visual exploration. *Frontiers in Human Neuroscience*, *14*, 41. doi: 10.3389/fnhum.2020.00041

Mormann, M., & Russo, J. E. (2021). Does attention increase the value of choice alternatives? *Trends in Cognitive Sciences*, *25*(4), 305-315. doi: https://doi.org/10.1016/j.tics.2021.01.004

Perkovic, S., Schoemann, M., Lagerkvist, C.-J., & Orquin, J. L. (2022). Covert attention leads to fast and accurate decision-making. *Journal of Experimental Psychology: Applied*. doi: 10.1037/xap0000425

Peschel, A. O., Orquin, J. L., & Loose, S. M. (2019). Increasing consumers' attention capture and food choice through bottom-up effects. *Appetite*, *132*, 1-7. doi: 10.1016/j.appet.2018.09.015

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 U.S.A*, *112*, 4170-4175. doi: 10.1073/

pnas.1415250112

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922. doi: 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*, 260-281. doi: 10.1016/j.tics.2016.01.007

Sepulveda, P., Usher, M., Davies, N., Benson, A. A., Ortoleva, P., & De Martino, B. (2020). Visual attention modulates the integration of goal-relevant evidence and not value. *eLife*, *9*, e60705. doi: 10.7554/eLife.60705

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, *6*, 1317–1322. doi: 10.1038/nn1150

Simion, C., & Shimojo, S. (2006). Early interactions between orienting, visual sampling and decision making in facial preference. *Vision Research*, *46*, 3331-3335. doi: 10.1016/j.visres.2006.04.019

Smith, S. M., & Krajbich, I. (2018). Attention and choice across domains. *Journal of Experimental Psychology: General*, *147*, 1810–1826. doi: 10.1037/xge0000482

Smith, S. M., & Krajbich, I. (2019). Gaze amplifies value in decision making. *Psychological Science*, *30*, 116-128. doi: 10.1177/0956797618810521

Tavares, G., Perona, P., & Rangel, A. (2017). The attentional drift diffusion model of simple perceptual decision-making. *Frontiers in Neuroscience*, *11*, 1-36. doi: 10.3389/fnins.2017.00468

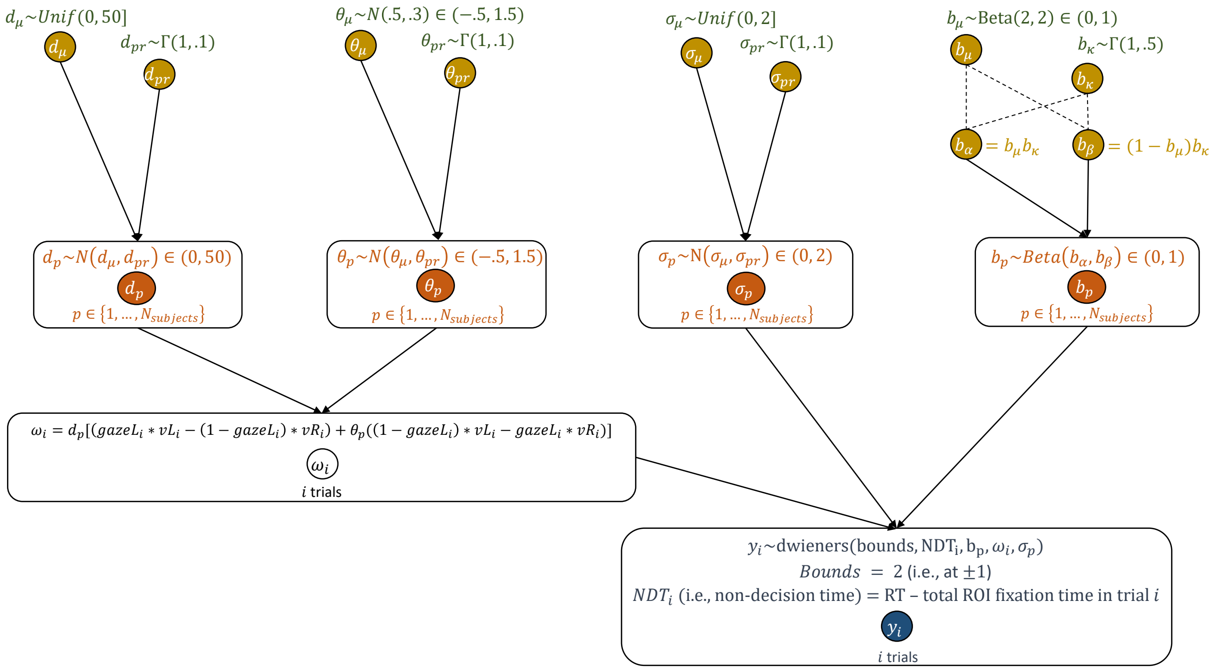
Thomas, A. W., Molter, F., & Krajbich, I. (2021). Uncovering the computational mechanisms underlying many-alternative choice. *eLife*, *10*, 1-27. doi: 10.7554/ELIFE.57012

Towal, R. B., Mormann, M., & Koch, C. (2013). Simultaneous modeling of visual saliency and value computation improves predictions of economic choice. *Proceedings of the National Academy of Sciences of the U.S.A.*, *110*, e1304429110. doi: 10.1073/pnas.1304429110

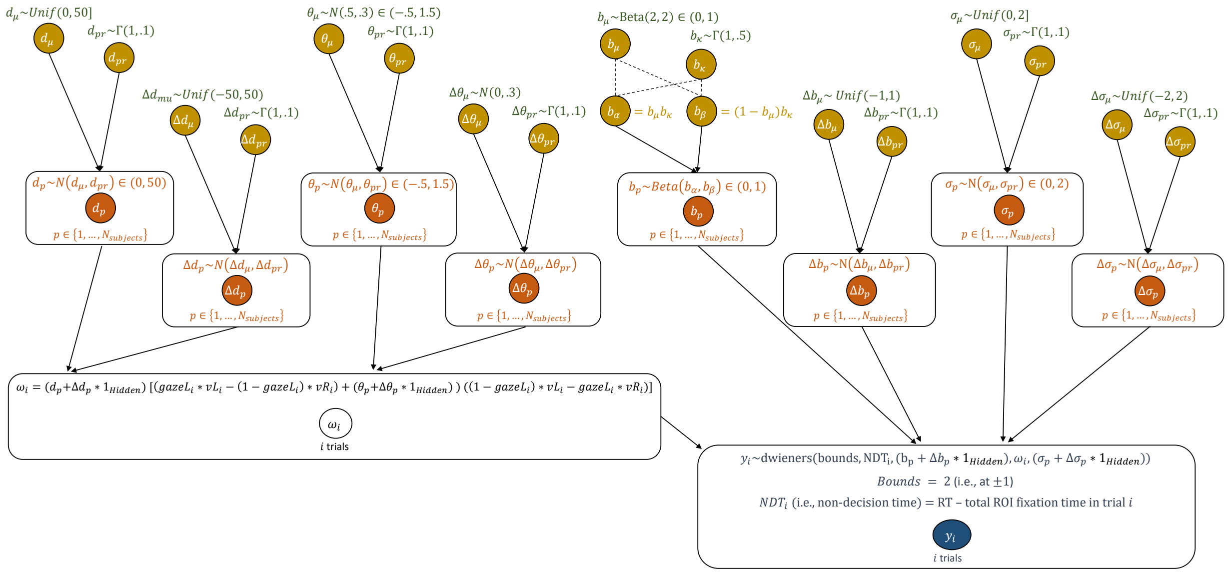
Weilbächer, R. A., Krajbich, I., Rieskamp, J., & Gluth, S. (2021, October). The influence of visual attention on memory-based preferential choice. *Cognition*, *215*, 104804. doi: 10.1016/j.cognition.2021.104804

Wästlund, E., Shams, P., & Otterbring, T. (2018). Unsold is unseen ... or is it? Examining the role of peripheral vision in the consumer choice process using eye-tracking methodology. *Appetite*, *120*, 49–56. doi: 10.1016/j.appet.2017.08.024

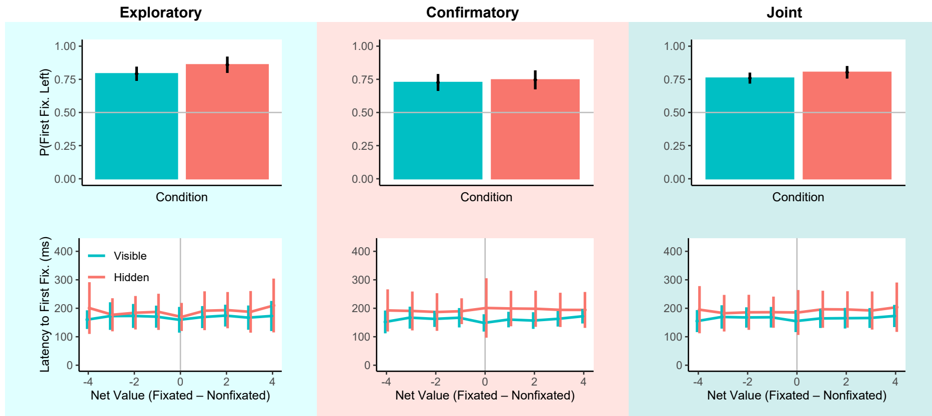
Supplementary Material



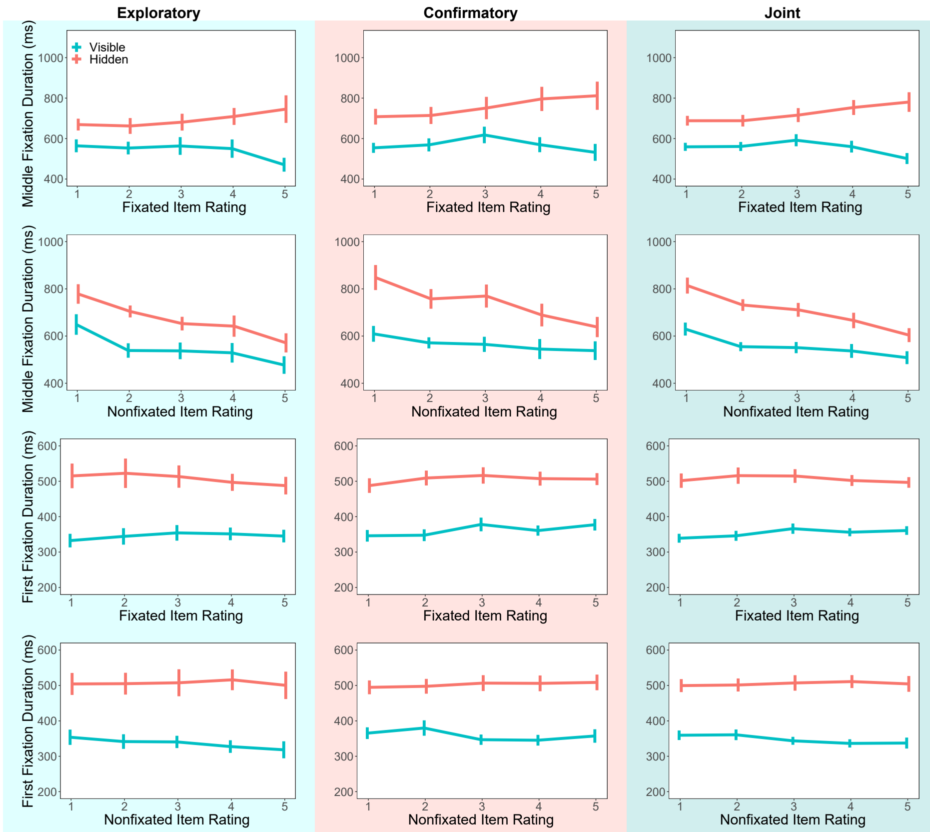
Supplementary Fig. 1. Directed acyclic graph of hierarchical aDDM with uncorrelated priors. The hierarchical model estimates group and individual parameters for the aDDM with uncorrelated priors. The 10 group parameters are depicted in the top row of yellow circles. The 4 individual parameters estimated for each subject are depicted in orange in the middle row. The distribution of individual parameters as a function of the group parameters is specified using a transformation of some of the parameters, denoted by the dashed lines. The choice and RT outcome y_i of trial i for a subject p is modeled as a Drift-Diffusion-Model with bounds at ± 1, non-decision time NDT_i, bias b_p, trial specific slope ω_i, and noise σ_p. The trial specific slope ω_i depends on the subject's drift rate parameter d_p, attentional bias parameter θ_p, gaze data for the trial $gazeL_i$, and item liking ratings for the foods used in the trial (vL_i,vR_i). $gazeL_i$ denotes the proportion of time spent fixating on the left item during the trial. The hyperpriors for the group parameters are described at the top of the graph. "$\in (X, Y)$" indicates truncation to bounds. $N_{subjects} = 25$ in the exploratory and confirmatory datasets, and 50 in the joint dataset.



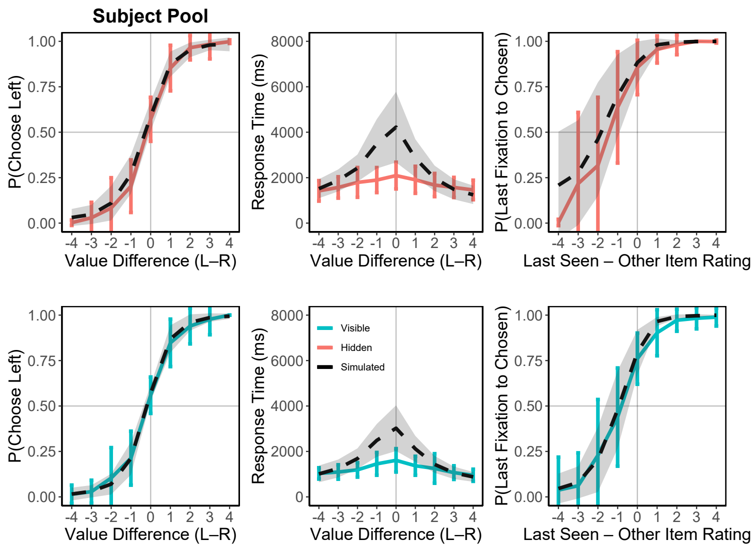
Supplementary Fig. 2. Directed acyclic graph of hierarchical aDDM with correlated priors. The hierarchical model estimates group and individual parameters for the aDDM with correlated priors. The 18 group parameters are depicted in the top row of yellow circles. The 8 individual parameters estimated for each subject are depicted in orange in the middle row. The distribution of individual parameters as a function of the group parameters is specified using a transformation of some of the parameters, denoted by the dashed lines. The choice and RT outcome y_i of trial i for a subject p is modeled as a Drift-Diffusion-Model with bounds at ± 1, non-decision time NDT_i, bias b_p, conditional difference in bias Δb_p, trial specific slope ω_i, noise σ_p, and conditional difference in noise $\Delta\sigma_p$. The trial specific slope ω_i depends on the subject's drift rate parameter d_p, conditional difference in drift rate parameter Δd_p, attentional bias parameter θ_p, conditional difference in attentional bias parameter $\Delta\theta_p$, gaze data for the trial $gazeL_i$, and item liking ratings for the foods used in the trial (vL_i,vR_i). $gazeL_i$ denotes the proportion of time spent fixating on the left item during the trial. The hyperpriors for the group parameters are described at the top of the graph. "$\in (X,Y)$" indicates truncation to bounds. $N_{subjects} = 25$ in the exploratory and confirmatory datasets, and 50 in the joint dataset.



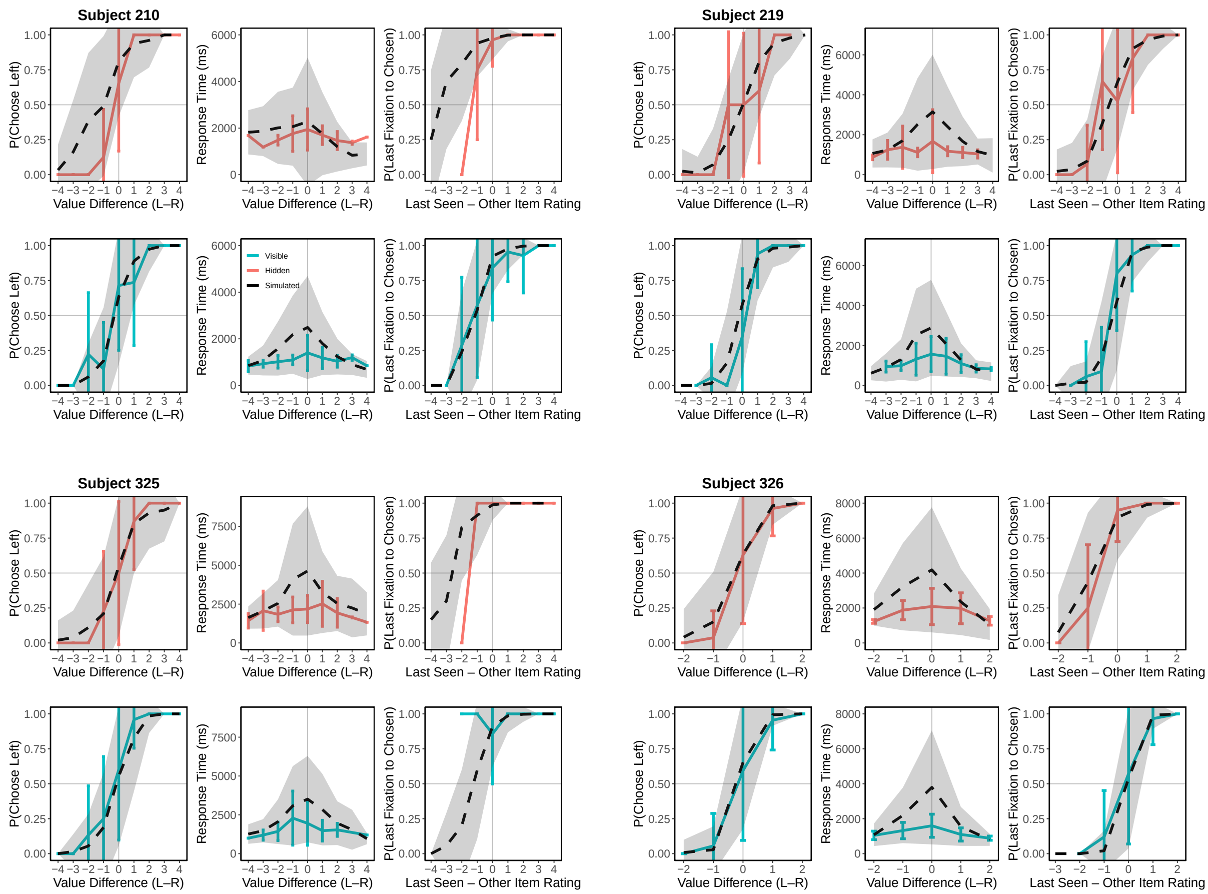
Supplementary Fig. 3. Additional fixation properties. (Top) Probability of first fixation to the left item in the two conditions. (Bottom) Latency to first fixation as a function of the relative rating of the fixated item. Columns indicate which dataset generated the figures. Black error bars show standard errors of the mean across participants. Color error bars show standard deviation across participants.



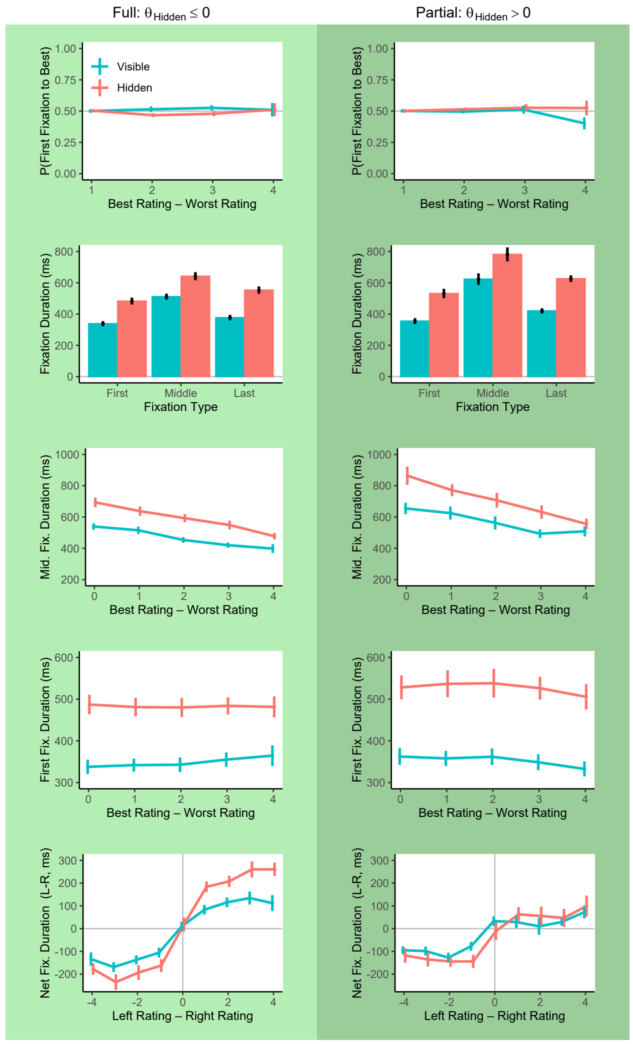
Supplementary Fig. 4. Fixation durations. (Row 1) Middle fixation duration as a function of the fixated item rating. (Row 2) Middle fixation duration as a function of the nonfixated item rating. (Row 3) First fixation duration as a function of the fixated item rating. (Row 4) First fixation duration as a function the nonfixated item rating. Columns indicate which dataset generated the figures. Error bars show standard errors of the mean across participants.



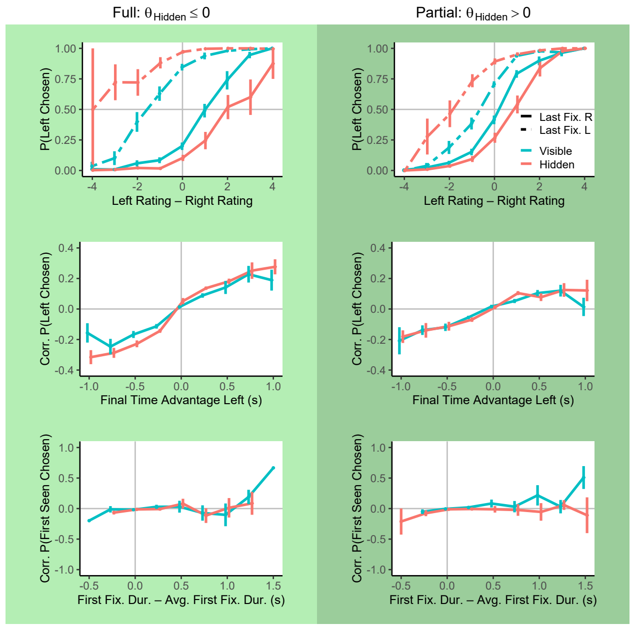
Supplementary Fig. 5. Group level predictions in the joint dataset. We use the estimates of the hierarchical aDDM in the odd trials to make predictions out-of-sample, in the even trials, separately for each subject and condition. For each subject, we simulate 10 observations per trial, and compare the simulated and observed data. *Blue lines:* Behavior in the visible condition. *Red lines:* Behavior in the hidden condition. *Black dashed lines and grey areas:* Simulated behavior for the respective condition (dash = mean, grey = SD). Error bars show standard deviations across subjects.



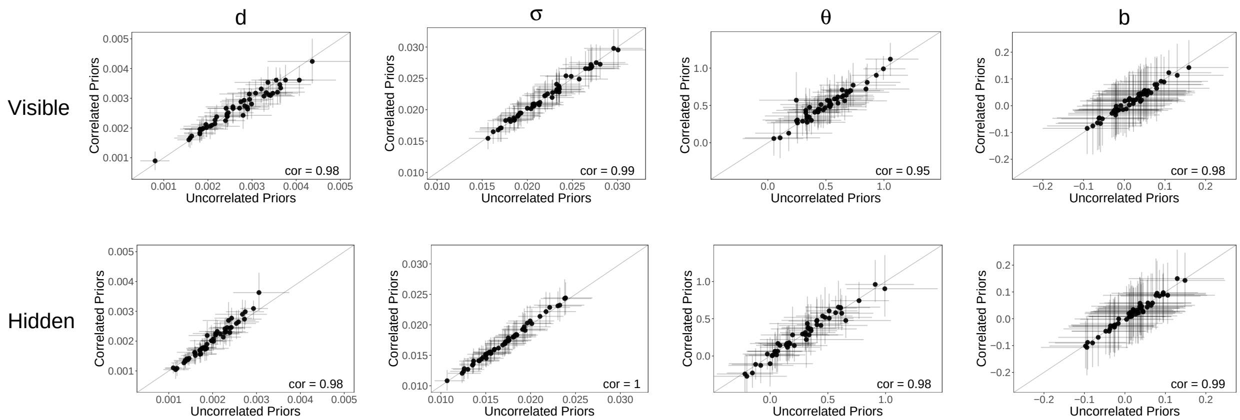
Supplementary Fig. 6. Subject-level simulations. Out-of-sample predictions versus data for four randomly selected subjects. See Figure S5 and Supplementary Methods for details. *Blue lines:* Behavior in the visible condition. *Red lines:* Behavior in the hidden condition. *Black dashed lines and grey areas:* Simulated behavior for the respective condition (dash = mean, grey = SD). Error bars show standard deviations across subjects.



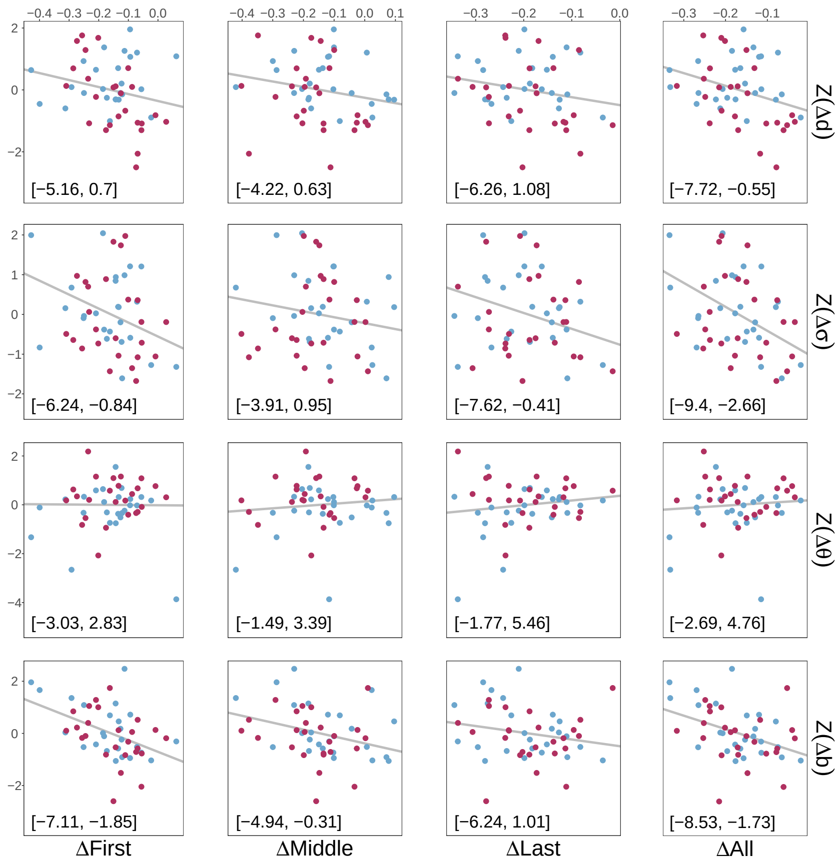
Supplementary Fig. 7. Fixation properties by attentional discounting group. (Row 1) The probability that the first fixation is to the best item as a function of choice difficulty. (Row 2) Fixation durations by fixation type. (Row 3) Middle fixation duration as a function of choice difficulty. (Row 4) First fixation duration as a function of choice difficulty. (Row 5) Net fixation duration to the left item as a function of its relative value. Columns indicate which dataset generated the figures: "$\theta_{Hidden} \leq 0$" indicates subjects with full attentional discounting ($N = 7$), "$\theta_{Hidden} > 0$" indicates subjects with partial attentional discounting ($N = 43$). Data is from the joint dataset. Error bars show standard errors of the mean across participants.



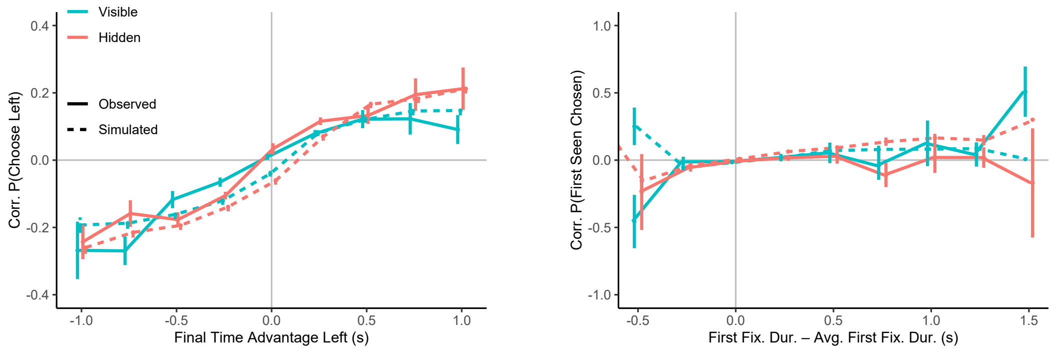
Supplementary Fig. 8. Choice biases by attentional discounting group. (Top) The probability of choosing the left item as a function of its relative value, conditional on last fixation location. (Middle) The corrected probability of choosing the left item as a function of the net fixation time to the left item. (Bottom) Corrected probability that the first seen item is chosen as a function of the excess first fixation duration, defined as first fixation duration minus mean first fixation duration (computed for each subject). Columns indicate which dataset generated the figures: "$\theta_{Hidden} \leq 0$" indicates subjects with full attentional discounting ($N = 7$), "$\theta_{Hidden} > 0$" indicates subjects with partial attentional discounting ($N = 43$). Data is from the joint dataset. Error bars show standard errors of the mean across participants.



Supplementary Fig. 9. Comparison of subject-level aDDM MAP parameter estimates between models with uncorrelated and correlated priors. Parameter estimate from the model with correlated priors as a function of its respective estimate from the model with uncorrelated priors. Rows separate by condition, columns separate by parameter. Data from joint data set. Black lines denote 95% HDIs. "cor" denotes Pearson correlation coefficient.



Supplementary Fig. 10. Linking fixation patterns with aDDM parameters. Each point represents a subject (blue = exploratory, red = confirmatory). ΔFirst is the mean first fixation duration in the visible condition minus the mean in the hidden condition. ΔMiddle, ΔLast, and ΔAll are the same, except for middle, last, and all fixations, respectively. Δd is the MAP estimate of drift rate in the visible condition minus the MAP estimate in the hidden condition. $\Delta\sigma$, $\Delta\theta$, and Δb are the same, except for the noise, attentional discounting, and bias parameters, respectively. Δparameters have been Z-scored. Data is from the joint dataset. Grey lines are univariate linear regression predictions. Black text presents the 95% HDI for the slope of the regression line.



Supplementary Fig. 11. Other choice bias simulations. (Left) Corrected probability of choosing the left item as a function of net fixation time to the left item, in observed (solid) and simulated (dashed) out-of-sample data. The corrected probability is computed by subtracting from each choice observation (coded as 1 if left chosen, and 0 otherwise) the proportion with which left is chosen at each relative value. (Right) Corrected probability that the first seen item is chosen as a function of the excess first fixation duration, defined as first fixation duration minus mean first fixation duration (computed for each subject), in observed and simulated out-of-sample data. Blue lines indicate data in the visible condition. Red lines indicate data in the hidden condition. Data is from the joint dataset. Error bars show standard errors of the mean across participants.

Supplementary Table 1. Regressions associated with the basic psychometric results in Fig. 3

		Exploratory			Confirmatory			Joint		
Dept. Var.	Indept. Var.	Est.	SE		Est.	SE		Est.	SE	
Left chosen	Intercept	0.11	0.07		0.03	0.04		0.07	0.04	
(Logistic)	L - R rating	1.58	0.13	*	1.69	0.15	*	1.64	0.10	*
(Top)	Hidden	-0.03	0.09		0.09	0.06		0.03	0.06	
	Interaction	0.19	0.10		0.08	0.08		0.12	0.06	*
RT	Intercept	1491.27	99.61	*	1759.54	121.52	*	1635.85	80.61	*
(Linear)	Best - worst rat.	-172.28	25.95	*	-213.22	26.02	*	-194.33	18.05	*
(Middle)	Hidden	542.61	76.61	*	497.94	96.90	*	520.08	61.05	*
	Interaction	-13.62	18.93		-33.38	30.68		-22.47	17.17	
# of fix.	Intercept	2.83	0.09	*	3.19	0.10	*	3.01	0.07	*
(Linear)	Best - worst rat.	-0.20	0.02	*	-0.27	0.02	*	-0.23	0.02	*
(Bottom)	Hidden	-0.02	0.07		-0.21	0.10	*	-0.11	0.06	
	Interaction	0.05	0.02	*	0.09	0.03	*	0.07	0.02	*

* indicates significance in all data sets at the 95% confidence level.
* indicates a significant effect that was not present in all three data sets.

Supplementary Table 2. Regressions associated with the fixation results in Fig. 4

		Exploratory			Confirmatory			Joint		
Dept. Var.	Indept. Var.	Est.	SE		Est.	SE		Est.	SE	
1st fix. best	Intercept	0.14	0.06	*	-0.10	0.05		0.01	0.04	
(Logistic)	Best - worst rat.	-0.04	0.03		0.04	0.03		0.00	0.02	
(Row 1)	Hidden	-0.16	0.09		0.13	0.08		0.00	0.06	
	Interaction	0.05	0.05		-0.06	0.04		-0.02	0.03	
Mid. fix. dur.	Intercept	618.35	36.49	*	595.78	36.27	*	612.92	23.59	*
(Linear)	Best - worst rat.	-55.23	8.09	*	-47.48	8.37	*	-51.20	5.84	*
(Row 3)	Hidden	103.82	27.53	*	129.57	33.42	*	141.50	22.48	*
	Interaction	-7.04	11.68		-7.87	15.14		-13.19	9.65	
1st fix. dur.	Intercept	341.64	21.31	*	353.30	16.29	*	348.37	13.23	*
(Linear)	Best - worst rat.	-0.83	2.43		1.50	2.38		0.32	1.66	
(Row 4)	Hidden	159.51	25.57	*	152.78	20.04	*	158.73	15.57	*
	Interaction	0.50	3.84		-4.78	3.39		-2.18	2.44	
Net fix. dur.	Intercept	14.45	20.80		-12.18	30.78		4.44	17.81	
(Linear)	Net. Val. > 0 (A)	-56.13	59.87		-38.46	69.75		-54.50	44.02	
(Row 5)	Net. Val. < 0 (B)	50.64	59.66		14.11	67.26		24.83	42.89	
	A : Net. Val. (C)	9.78	20.15		16.86	21.88		14.55	14.77	
	B : Net. Val. (D)	63.20	20.40	*	47.00	21.64	*	54.48	14.66	*
	Hidden (E)	-22.05	39.01		-21.88	42.16		-23.83	27.89	
	A:E	68.83	41.34		81.30	43.28		77.78	29.09	*
	B:E	-108.67	39.16	*	-40.40	43.83		-70.26	28.91	*
	C:E	3.08	12.73		8.90	14.65		5.49	9.84	
	D:E	-35.43	13.05	*	-6.01	13.85		-20.20	9.35	*

* indicates significance in all data sets at the 95% confidence level.
* indicates a significant effect that was not present in all three data sets.

Supplementary Table 3. Regressions associated with the first fixation latency results in Fig. S3

		Exploratory			Confirmatory			Joint		
Dept. Var.	Indept. Var.	Est.	SE		Est.	SE		Est.	SE	
Latency to 1st fix.	Intercept	167.16	8.89	*	157.22	6.89	*	162.48	162.48	*
(Linear)	Fix. - nonfix. rating	-0.97	3.26		1.45	4.38		0.30	0.30	
	Hidden	3.74	4.42		5.06	4.75		8.20	8.20	
	Interaction	1.12	3.97		0.81	4.59		2.17	2.17	

* indicates significance in all data sets at the 95% confidence level.
* indicates a significant effect that was not present in all three data sets.

Supplementary Table 4. Regressions associated with the fixation duration results in Fig. S4

		Exploratory			Confirmatory			Joint		
Dept. Var.	Indept. Var.	Est.	SE		Est.	SE		Est.	SE	
Mid. fix. dur.	Intercept	587.88	36.06	*	567.63	27.92	*	576.42	21.60	*
(Linear)	Fix. rating	-11.83	7.66		1.43	8.44		-3.96	5.66	
(Row 1)	Hidden	61.91	28.76	*	85.31	29.15	*	78.08	20.22	*
	Interaction	24.21	10.49	*	29.37	11.11	*	26.19	7.51	*
Mid. fix. dur.	Intercept	672.53	38.31	*	654.00	28.25	*	661.13	23.87	*
(Linear)	Nonfix. rating	-42.07	6.73	*	-29.98	8.84	*	-34.24	5.74	*
(Row 2)	Hidden	143.89	30.69	*	218.65	38.68	*	193.83	25.95	*
	Interaction	-8.57	9.43		-21.26	9.77	*	-16.84	6.62	*
1st fix. dur.	Intercept	329.36	21.44	*	340.71	20.31	*	334.77	14.75	*
(Linear)	Fix. rating	4.17	2.23		5.71	3.47		5.19	2.00	*
(Row 3)	Hidden	189.71	30.63	*	156.43	23.03	*	176.84	19.61	*
	Interaction	-10.27	4.56	*	-3.45	3.56		-7.31	3.07	*
1st fix. dur.	Intercept	357.71	21.32	*	374.03	17.20	*	367.47	13.59	*
(Linear)	Nonfix. rating	-6.73	2.15	*	-6.63	2.01	*	-6.52	1.46	*
(Row 4)	Hidden	129.55	24.69	*	122.36	19.34	*	130.51	15.53	*
	Interaction	10.20	3.18	*	7.81	2.94	*	8.80	2.12	*

* indicates significance in all data sets at the 95% confidence level.
* indicates a significant effect that was not present in all three data sets.

Supplementary Table 5. Regressions associated with choice bias results in Fig. 5

		Exploratory			Confirmatory			Joint		
Dept. Var.	Indept. Var.	Est.	SE		Est.	SE		Est.	SE	
Left chosen	Intercept	-0.68	0.16	*	-1.59	0.26	*	-1.11	0.16	*
(Logistic)	Left - right rating (A)	1.54	0.14	*	1.75	0.18	*	1.64	0.11	*
(Top)	Last fix. loc. (0=R,1=L;B)	1.83	0.31	*	3.20	0.47	*	2.46	0.30	*
	Hidden (C)	-1.20	0.30	*	-1.46	0.36	*	-1.32	0.23	*
	A:B	0.15	0.12		-0.14	0.13		-0.01	0.08	
	A:C	0.28	0.15	*	-0.18	0.18		0.08	0.09	
	B:C	3.74	0.63	*	3.82	0.63	*	3.69	0.43	*
	A:B:C	-0.59	0.23	*	0.14	0.25		-0.26	0.15	
Corr. left chosen	Intercept	0.00	0.01		0.00	0.00		0.00	0.00	
(Linear)	Net fixation left (s)	0.25	0.04	*	0.24	0.04	*	0.24	0.03	*
(Middle)	Hidden	0.00	0.01		0.01	0.01		0.01	0.01	
	Interaction	0.04	0.04		0.04	0.03		0.04	0.02	
Corr. 1st seen chosen	Intercept	-0.01	0.01		-0.02	0.01	*	-0.01	0.00	*
(Linear)	Excess first fix. dur. (s)	0.21	0.07	*	0.21	0.08	*	0.22	0.05	*
(Bottom)	Hidden	-0.01	0.01		0.00	0.01		0.00	0.01	
	Interaction	-0.21	0.07	*	-0.13	0.06	*	-0.18	0.05	*

* indicates significance in all data sets at the 95% confidence level.

* indicates a significant effect that was not present in all three data sets.

Excess first fixation duration is defined as first fixation duration minus mean first fixation duration.

Supplementary Table 6. Group-level MAP parameter estimates for model with correlated priors across datasets and conditions

	Exploratory		**Confirmatory**		**Joint**	
	H	V	H	V	H	V
d_{group}	0.002	0.003	0.002	0.003	0.002	0.003
	[0.002, 0.002]	[0.002, 0.003]	[0.002, 0.002]	[0.002, 0.003]	[0.002, 0.002]	[0.002, 0.003]
σ_{group}	0.018	0.023	0.016	0.021	0.017	0.022
	[0.015, 0.020]	[0.021, 0.025]	[0.014, 0.019]	[0.019, 0.023]	[0.016, 0.019]	[0.021, 0.023]
θ_{group}	0.38	0.54	0.20	0.49	0.28	0.52
	[0.22, 0.54]	[0.42, 0.66]	[0.06, 0.35]	[0.37, 0.62]	[0.18, 0.39]	[0.44, 0.61]
b_{group}	0.02	0.03	0.01	0.00	0.02	0.02
	[-0.10, 0.14]	[-0.06, 0.11]	[-0.10, 0.14]	[-0.09, 0.09]	[-0.04, 0.08]	[-0.03, 0.06]

MAP estimate and 95% HDI of group-level mean.

Supplementary Table 7. Effects of early versus late fixations on corrected choice

		Exploratory			Confirmatory			Joint		
Dept. Var.	Indept. Var.	Est.	SE		Est.	SE		Est.	SE	
Corr. left chosen	Intercept	0.03	0.01	*	0.03	0.01	*	0.03	0.01	*
Observed	Net fix. L before 1 s (A)	0.17	0.03	*	0.12	0.03	*	0.16	0.02	*
	Net fix. L after 1 s (B)	0.24	0.05	*	0.23	0.03	*	0.23	0.03	*
	Hidden (C)	0.02	0.01		0.03	0.01		0.02	0.01	
	A:C	-0.03	0.04		0.04	0.03		-0.01	0.02	
	B:C	0.02	0.04		0.04	0.03		0.04	0.02	
Corr. left chosen	Intercept							-0.01	0.00	
Simulated	Net fix. L before 1 s (A)							0.05	0.01	*
	Net fix. L after 1 s (B)							0.13	0.02	*
	Hidden (C)							0.01	0.00	
	A:C							0.02	0.01	
	B:C							0.06	0.02	*

* indicates significance in all data sets at the 95% confidence level.

* indicates a significant effect that was not present in all three data sets.

"Prop.": proportion.

Trials shorter than 1 s dropped for this analysis.

Supplementary Table 8. Regressions assocaited with choice bias simulations in Fig. S11

Dept. Var.	Indept. Var.	Est.	SE	
Corr. left chosen	Intercept	-0.02	0.00	*
(Linear)	Net fixation left (s)	0.13	0.02	*
(Left)	Hidden	0.00	0.00	
	Interaction	0.08	0.02	*
Corr. 1st seen chosen	Intercept	-0.01	0.00	*
(Linear)	Excess first fix. dur. (s)	0.06	0.02	*
(Right)	Hidden	0.03	0.00	*
	Interaction	0.06	0.02	*

* indicates significance at the 95% confidence level.