1   Gaze-dependent evidence accumulation predicts multi-alternative risky choice

2   behaviour

3   Felix Molter[1,2,3,*], Armin W. Thomas[2,4], Scott A. Huettel[5,6], Hauke R. Heekeren[2,7], and Peter N.

4   C. Mohr[1,2,3]

5   [1] School of Business & Economics, Freie Universität Berlin, Berlin, Germany
6   [2] Center for Cognitive Neuroscience, Freie Universität Berlin, Berlin, Germany
7   [3] WZB Berlin Social Science Center, Berlin, Germany
8   [4] Department of Electrical Engineering and Computer Science, Technische Universität Berlin, Berlin, Germany
9   [5] Center for Cognitive Neuroscience, Duke University, Durham, NC, USA
10   [6] Department for Psychology and Neuroscience, Duke University, Durham, NC, USA
11   [7] Department for Education and Psychology, Freie Universität Berlin, Berlin, Germany

12

13   [*] Correspondence concerning this article should be addressed to Felix Molter, School of Business & Economics,
14   Freie Universität Berlin, Garystr. 21, 14195 Berlin (E-mail: felixmolter@gmail.com)

Abstract

Choices are influenced by gaze allocation during deliberation, so that fixating an alternative longer leads to increased probability of choosing it. Gaze-dependent evidence accumulation provides a parsimonious account of choices, response times and gaze-behaviour in many simple decision scenarios. Here, we test whether this framework can also predict more complex context-dependent patterns of choice in a three-alternative risky choice task, where choices and eye movements were subject to attraction and compromise effects. Choices were best described by a gaze-dependent evidence accumulation model, where subjective values of alternatives are discounted while not fixated. Finally, we performed a systematic search over a large model space, allowing us to evaluate the relative contribution of different forms of gaze-dependence and additional mechanisms previously not considered by gaze-dependent accumulation models. Gaze-dependence remained the most important mechanism, but participants with strong attraction effects employed an additional similarity-dependent inhibition mechanism found in other models of multi-alternative multi-attribute choice.

Introduction

Imagine you inherited money from a distant relative and you need to decide how to invest it. You reach out to your trusted investment advisor who swiftly sends you a brochure with different investment alternatives. After reading the brochure, you look at the overview table provided on the final page and deliberate between two alternatives that particularly interest you, visually inspecting them, one at a time, comparing expected returns, associated risks, fees, and other attributes. Previous work has established that the role of visual attention in decision making under risk exceeds mere information sampling[1–3]. Instead, as in other forms of preferential and

38   perceptual decision making, visual fixations have a constructive role in the decision process, so

39   that alternatives that are looked at for a longer time are generally more likely to be chosen[4–10].

40   This association between gaze and choice has been formalized in gaze-dependent evidence

41   accumulation models[1,2,5,6,10–13]. The attentional Drift Diffusion Model (aDDM)[6], Gaze-weighted

42   Linear Accumulator Model[12,13], and others that apply explicitly to risky choice[1] assume that

43   decision makers accumulate evidence in favour of each alternative until evidence for one

44   alternative reaches a threshold and a decision is made. Crucially, accumulation rates are assumed

45   to depend on gaze allocation, so that evidence for an alternative is discounted while it is not

46   fixated. Prior work has tested these assumptions in simple decisions under risk, involving two

47   risky gambles with two equally probable outcomes[3] or two risky gambles described by outcome

48   and probability[1]. Like the example of choosing an investment plan, however, many real-life

49   decisions are more complex than simple binary choice. For example, investment decisions can

50   involve more than two alternatives that vary on multiple attributes (such as expected return,

51   associated risk, and fees), including both sets of options with similarly low expected returns (e.g.,

52   government bonds, fixed deposits) and sets of other riskier options promising larger gains at

53   higher risk (e.g., stocks or derivatives).

54   Crucially, choices in these multi-alternative, multi-attribute settings pose a challenge to

55   many traditional models of risky choice, including Expected Utility Theory (EU)[14] and Prospect

56   Theory[15,16], which obey basic axioms of rational choice like independence of irrelevant

57   alternatives (IIA)[17]. Briefly, IIA asserts that the preference between two alternatives should not

58   depend on other alternatives. At least three context effects – the attraction[18], compromise[19], and

59   similarity[20] effects – show that IIA is frequently violated: The attraction effect describes an

60   increase in preference for an alternative following the addition of a similar, but slightly inferior

61   alternative. The compromise effect describes an increase in preference for an alternative after a

62   third alternative was added that makes it appear as intermediate. The similarity effect predicts

63   that adding a third option that is similar to one of the original options, and equally appealing, will

64   increase relative preference for the other, dissimilar alternative. While these effects are

65   predominantly investigated using consumer goods, some studies also found them to affect

66   choices between risky gambles[18,20–23].

67   Generally, models that assume that each alternative can independently be assigned a single

68   scale value denoting its utility (fixed utility or simple scalability models[24]) cannot account for the

69   observed violations. In contrast, models that can predict the effects often assume that preferences

70   are constructed by means of comparisons between alternatives on different attribute

71   dimensions[20,25–28], and employ additional psychological or neurobiologically-inspired

72   mechanisms, such as loss-aversion or inhibition between alternatives.

73   While current gaze-dependent accumulation models of risky choice assume that each

74   alternative can be assigned a singular underlying scale value (e.g., expected utilities for each

75   alternative), choices are also assumed to be constructed in an accumulation process that remains

76   malleable due to its dependence on gaze. Therefore, gaze-dependent accumulation models

77   potentially could account for violations of IIA if the distribution of gaze itself changed depending

78   on the decision context (e.g., shift towards dominant alternatives in the attraction effect,

79   compromise alternatives in the attraction effect, or dissimilar alternatives in the similarity effect).

80   Initial evidence for such an attention-shift in the attraction effect comes from a study of

81   intertemporal choice[29]. Furthermore, recent work has suggested gaze-dependent evidence

82   accumulation as a model of multi-alternative, multi-attribute choice in the absence of context

83   effects[2,30].

84   Here, we test to what extent gaze-dependent accumulation can explain risky choices in the

85   presence of context effects. We replicate the attraction and compromise effects in a task where

86   participants made repeated choices in a multi-alternative, multi-attribute setting involving risky

87   gambles. We compare a model of gaze-dependent accumulation adapted from previous work on

88   binary risky choice[1] with an established model of context-dependent multi-alternative, multi-

89   attribute choice without gaze-dependence and with other traditional models of risky choice.

90   Notably, the gaze-dependent accumulation model did not include any dedicated mechanisms to

91   produce context effects. Our task induced a set of context effects across the group, with

92   substantial variability between individuals, providing a complex testing scenario to compare

93   gaze-dependent accumulation models with competing theoretical accounts. We found the gaze-

94   dependent model to give the overall best account of the choice data, while underestimating strong

95   attraction effects of some individuals. A systematic search over a large model space combining

96   features across model classes confirmed that all participants' behaviour was best described by a

97   gaze-dependent accumulation process, but that individual differences in attraction effect strengths

98   are predicted best by variants integrating mechanisms from other models.

Results
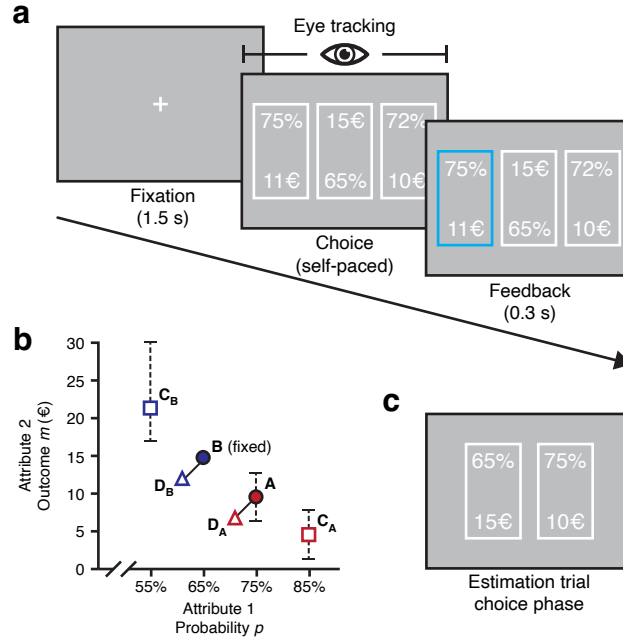
Choice task

Figure 1: **Choice task and attribute space of stimuli. (a)** Choice task. Trials started with a
central fixation cross for 1.5 s. Next, three gambles were presented. Participants made a choice by
button press, without time limit. Eye movements were recorded during the choice phase. Finally,
brief feedback on the choice (but not on the gamble outcome) was displayed. **(b)** Attribute space.
Each gamble is described by two attributes: Probability p and outcome m. The core options A
and B were presented in every trial, along with one of four decoy alternatives: Compromise
decoys (C_A or C_B) and asymmetrically dominated decoys (D_A or D_B) are expected to elicit
compromise and attraction effects, respectively. Dashed lines indicate possible stimulus
placements. **(c)** Exact positions for C_B, A and C_A were calibrated for pairwise indifference in
separate binary choice estimation blocks for each individual. Differences in outcome between
neighbouring alternatives was not less than 2 EUR. Dominated decoys D_A and D_B were always
2% and 1 EUR worse than A and B respectively.

In our experiment, 40 participants made repeated decisions between three all-or-nothing gambles, each described by a probability p to win an outcome m and nothing otherwise (Fig. 1a). We recorded participants' eye movements during the task with an eye tracker (see Methods for details). Participants were instructed that after completing the task, their chosen gamble from one randomly determined trial would be played out for an additional bonus payment. In contrast to hypothetical choices between consumer goods described on attribute dimensions like quality and price, risky gamble stimuli offer a high amount of control over their attributes and a straightforward way to incentivise choices. The gamble stimuli were designed to elicit attraction and compromise effects and were individually tailored to account for each participant's risk preferences (Fig. 1b). Participants performed three pairs of indifference-estimation and experimental blocks, for a total of 225 experimental trials (see Methods for additional details on the experimental procedure and gamble stimuli). In indifference-estimation blocks, participants made repeated choices between pairs of gambles with different probabilities (Fig. 1c). Gamble outcomes m were adjusted according to participants' choices so that four approximately equally preferred gambles C_B, B, A, and C_A were constructed with winning probabilities $p = 55\%$, 65%, 75%, and 85%. Additionally, two asymmetrically dominated decoys D_A and D_B were defined to be 2% and 1 EUR worse than gambles A and B, respectively. Following this indifference estimation, participants performed an experimental block of 75 ternary choice trials, 32 of which were compromise trials balanced with respect to the target option (i.e., 16 choice sets $\{A, B, C_A\}$, 16 choice sets $\{A, B, C_B\}$), 32 attraction trials (16 choice sets $\{A, B, D_A\}$, 16 choice sets $\{A, B, D_B\}$) and 11 distractor trials showing randomly created options with expected value of 10 EUR and low (5–33%), medium (34–64%) and high (65–95%) probability p. Uniform noise was added to all outcomes m (± 0 EUR, $+1$ EUR) and probabilities p (-3%, $\pm 0\%$, $+3\%$) in each trial.

137 Asymmetrically dominated decoys D_A and D_B received the same noise as the target option to

138 preserve the dominance relation.

139 **Context effects**

Trial type	Target	Competitor	Decoy
D_A	0.60	0.39	0.01
D_B	0.53	0.45	0.02
C_A	0.46	0.37	0.16
C_B	0.32	0.31	0.37

140 Table 1: **Relative choice frequencies across participants in the four trial types.** Across the
141 group, target options were chosen more frequently than competitors in both types of attraction
142 (D_A and D_B) and compromise trials (C_A, C_B). Dominated decoys were almost never chosen. In
143 compromise trials including the high-outcome decoy C_B , the decoy was chosen more frequently
144 than core options.

145

146   We first analysed participants' choice behaviour to test whether their choices could be
147 influenced by the set of available alternatives (see Table 1 and Fig. 2). Participants rarely chose
148 the dominated decoys in attraction trials (mean $\pm$ s.d. $= 0.02 \pm 0.02$ of attraction trials),
149 indicating that participants understood the dominance relationships among the stimuli. In
150 compromise trials, participants chose (non-dominated) decoy alternatives more frequently (mean
151 $\pm$ s.d. $= 0.27 \pm 0.16$). In particular, the decoy with the highest possible outcome C_B ($p \approx 55\%$,
152 $m \approx 18$ EUR) was chosen most frequently when it was available. Note that decoy choices in
153 compromise trials are expected, as extreme decoys were specifically calibrated to be
154 approximately equally preferred to the core options.

155   We tested the presence of the attraction and compromise effects by first computing the

156   relative choice share of the target alternative (RST)[31] for each individual, and separately for

157   attraction and compromise trials. The RST is computed as

$$RST = \frac{N_{\text{target}}}{N_{\text{target}} + N_{\text{competitor}}} \tag{1}$$

158

159   where N is the number of times an alternative was chosen. Given the balanced design,

160   where each core alternative acts as a target and competitor equally often, the RST should be close

161   to 0.5 if no context effects are present. If, however, the RST is different from 0.5, a systematic

162   context effect is indicated. We tested whether the mean RST significantly differed from 0.5

163   across participants by computing its 95% highest posterior density interval (HDI_{95}) using

164   Bayesian estimation (BEST)[32,33].

165   We find evidence for the attraction effect: The mean RST in attraction trials differed

166   meaningfully from 0.5 (mean RST = 0.56, HDI_{95} = [0.51, 0.60], mean d = 0.49, HDI_{95} = [0.10,

167   0.80]). 25 of 40 (62.5%) participants had an RST above 0.5 in attraction trials. Notably, similar to

168   previous work[34] a subgroup of participants (9 of 40, 22.5%) showed particularly strong attraction

169   effects with individual RSTs above 0.7. We could not, however, find evidence that these

170   individuals used the dominance relationship as a simplifying choice rule (see Supplementary

171   Note 2). To allow comparisons with other studies that quantified context effects as the difference

172   between choice shares between targets and competitors, we also report these differences: In

173   attraction trials, the average difference was 0.12 (HDI_{95} = [0.01, 0.20], mean d = 0.48, HDI_{95} =

174   [0.10, 0.85], Fig. 2b).

175   We only found weak evidence for the compromise effect using the gamble stimuli: The

176   mean RST in compromise trials was 0.53, but its estimated HDI_{95} did not exclude 0.5 (HDI_{95} =

177   [0.49, 0.57], 91.1% of posterior mass above 0.5, mean $d = 0.23$, $HDI_{95} = [-0.11, 0.59]$). 26 of 40

178   (65%) participants showed an RST above 0.5 in compromise trials. The mean difference between

179   choice shares for targets and competitors was 0.05 ($HDI_{95} = [-0.01, 0.11]$, mean $d = 0.29$, HDI_{95}

180   $= [-0.04, 0.62]$, 95.8 % of posterior mass above 0; Fig. 2e). These results are similar to the

181   marginal effects obtained using perceptual stimuli in previous work[35,36].

182   Similar to other work[26,31], we found a positive correlation between the effects across

183   participants, even though it was weaker in our task ($r = 0.24$, $HDI_{95} = [-0.05, 0.51]$, 94.5% of

184   posterior mass above 0).

185   Taken together, we successfully induced context effects within our participant sample,

186   with non-trivial variability in the strength of those effects across individuals. These data provide a

187   complex testing scenario to investigate gaze-bias effects in multi-alternative multi-attribute

188   choice and to compare gaze-dependent accumulation models with competing theories.
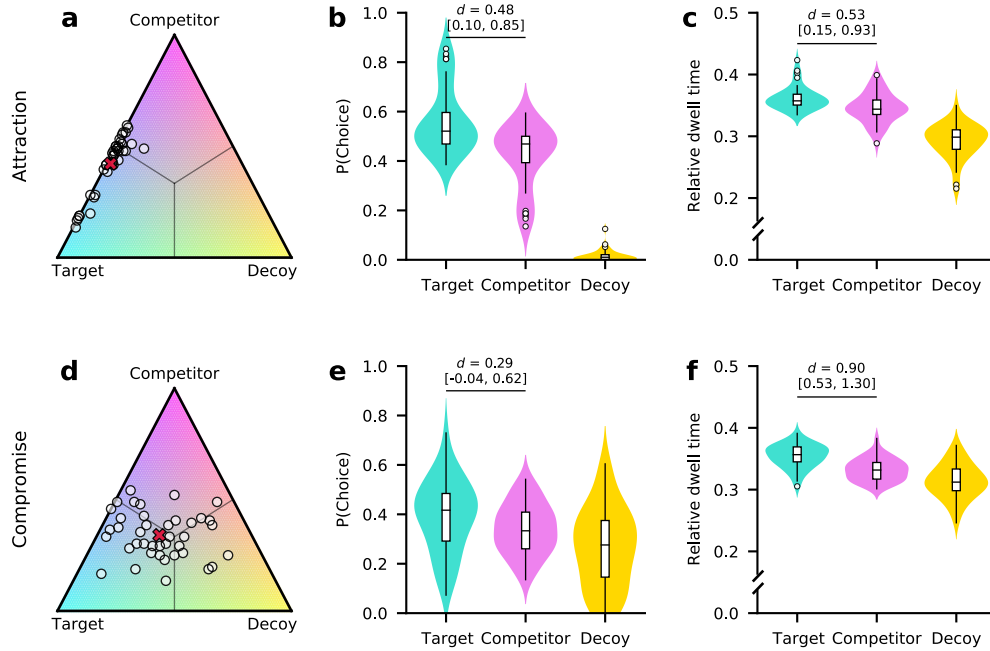
Figure 2: **Context effects are present in choices and relative dwell time.** Participants choices were influenced by asymmetrically dominated decoys and, to a lesser extent, by extreme compromise-making decoys. **(a, d)** Ternary plots of individual relative choice frequencies for target (lower left, teal), competitor (top, pink), and decoy (lower right, yellow) alternatives in attraction **(a)** and compromise **(d)** trials. Each dot represents one participant. The position on the simplex indicates relative choice frequencies for alternatives. Straight lines from the centre indicate equal frequencies for two alternatives. The red "x" indicates the group average. **(b, e)** Relative choice frequencies in attraction **(b)** and compromise **(e)** trials. In attraction trials, some participants strongly favoured the target alternative and almost no decoy choices were made. While target alternatives are still chosen more frequently than competitors in compromise trials, the effect is less pronounced, and extreme decoys are still chosen frequently. **(c, f)** Relative dwell time towards alternatives. In both, attraction **(c)** and compromise **(f)** trials, target alternatives received greater relative dwell times than competitors. d denotes Cohen's d from paired BEST analysis with HDI_{95} given in brackets. Violin plots show kernel density estimates of distributions of individual values. Box plots mark lower and upper quartiles and median. Whiskers extend from first and last datum within 1.5 times the interquartile range from lower and upper quartiles, respectively. Values outside this range are indicated by open circles.

11

Context effects are present in relative dwell times

207

208   Next, we tested whether participants' eye movements were affected by the set of available

209   alternatives, similar to their choices. We therefore computed each alternative's relative dwell

210   time in a trial by summing the duration of all fixations durations towards it, and normalizing to

211   the sum of all fixations in the trial. Both patterns of context effects of choice behaviour were

212   present in the relative dwell data: Target options received greater relative dwell time than

213   competitors in attraction (mean difference = 0.014, HDI_{95} = [0.004, 0.024], mean d = 0.53,

214   HDI_{95} = [0.15, 0.93], Fig. 2c) and compromise trials (mean difference = 0.022, HDI_{95} = [0.014,

215   0.030], mean d = 0.90, HDI_{95} = [0.53, 1.30], Fig. 2f). Targets and competitors also received

216   greater relative dwell time than decoys in both trial types.

217   To better understand participants' eye movements over the course of the trial, we

218   performed additional exploratory analyses of gaze location and transitions. We found an

219   increasing association between gaze and choice over the trial, and longer gaze towards targets,

220   even accounting for choice. Information search occurred both within and between alternatives,

221   with slightly more transitions within alternatives. We refer to Supplementary Note 1 for

222   additional details on these analyses.

Behavioural modelling and model comparison
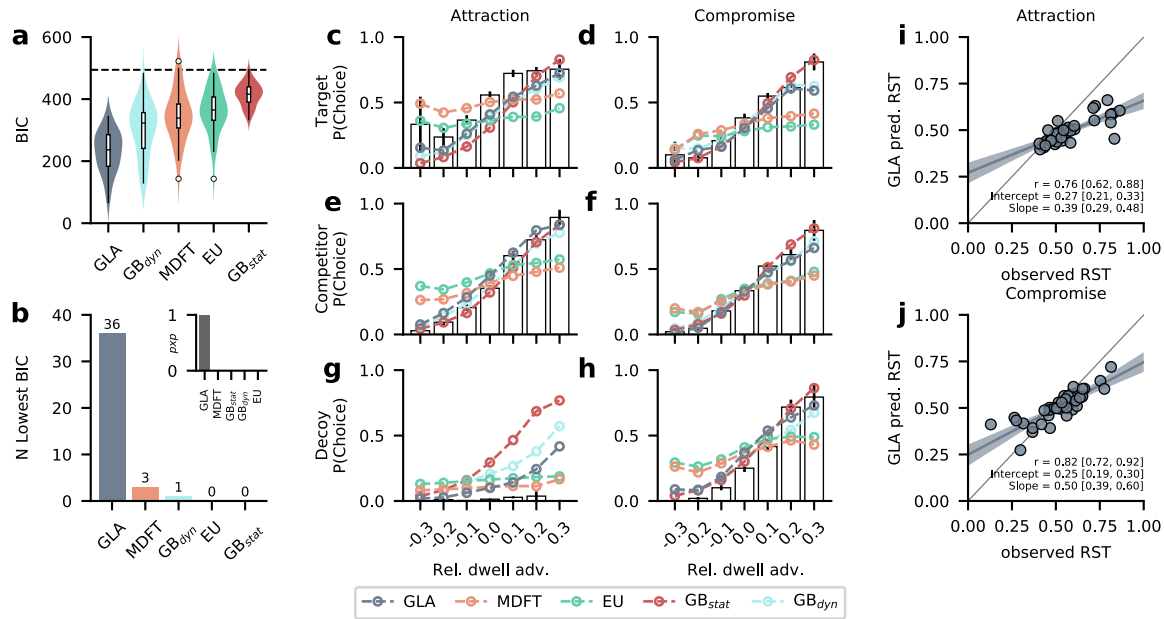
Figure 3: **Model comparison and predictions. (a)** The gaze-dependent leaky accumulator (GLA) provided the best fit (lowest mean BIC) across participants, followed by the dynamic gaze baseline model, MDFT, EU, and the static gaze baseline model. The dashed line indicates the BIC of the random choice baseline model. Violin plots show kernel density estimates of distributions of individual values. Box plots mark lower and upper quartiles and median. Whiskers extend from first and last datum within 1.5 times the interquartile range from lower and upper quartiles, respectively. Values outside this range are indicated by open circles. **(b)** The GLA fitted most (36 of 40, 90%) participants best, with a protected exceedance probability of 1 (inset). **(c-h)** Observed and model-predicted probability of choosing the target **(c, d)**, competitor **(e, f)**, or decoy **(g, h)** alternatives in attraction **(c, e, g)** and compromise trials **(d, f, h)** as a function of relative dwell time advantage. Relative dwell time advantage was computed as relative dwell time towards an alternative minus the mean relative dwell time to all other alternatives. White bars and error bars show mean ± s.e. observed data from even-numbered trials. Model predictions (coloured lines) are based on 50 simulations of each odd-numbered trial. **(i, j)** Observed and predicted RST of the best-fitting GLA for attraction **(i)** and compromise **(j)** trials. Each circle represents one participant. The winning model's predicted context effect sizes correlated significantly with the observed ones. Strong context effects, however, were underestimated, as indicated by the reduced slopes.

13

243

Given that participants' choices and eye movements exhibited modulation by the context

of available options, we next tested whether their behaviour could be described using a gaze-

dependent accumulation model, and how such a model performs in comparison to extant theories

of multi-alternative multi-attribute choice. To this end, we adapted a recently developed, gaze-

dependent leaky accumulator model of two-alternative risky choice[1] to the three-alternative

scenario. We refer to this model as the Gaze-dependent Leaky Accumulator (GLA). It assumes

that subjective utilities for each alternative i are constructed as in Cumulative Prospect Theory[16]

by first applying a probability weighting function, that transforms objective probabilities p into

subjective decision weights. Outcomes m are transformed into utilities using a power function.

Next, this model assumes that for each alternative i subjective utilities x_i^{GLA} are accumulated with

leakage over the time course of each trial and that utilities are discounted while they are not

fixated. At each fixation n, accumulators evolve according to

$$X_i^{GLA}(n) = \begin{cases} (1 - \lambda) \cdot X_i^{GLA}(n-1) + 1 \cdot x_i^{GLA} & \text{if } i \text{ fixated} \\ (1 - \lambda) \cdot X_i^{GLA}(n-1) + \theta \cdot x_i^{GLA} & \text{otherwise} \end{cases} \qquad (2)$$

where all $X_i(0) = 0$. Similar to the aDDM, the θ parameter $(0 \leq \theta \leq 1)$ controls

discounting of unattended alternative utilities. The λ parameter $(0 \leq \lambda \leq 1)$ controls the strength

of the accumulation leak. We computed choice probabilities from this model by applying the

soft-max choice rule (Eq. (4)) over the final accumulator values X_i^{GLA}.

Additionally, we fitted an implementation of Multialternative Decision Field Theory

(MDFT)[25], representing the class of dynamic multi-attribute, multi-alternative choice theories

that have been designed to explain context effects. In contrast to GLA, MDFT assumes that

preferences evolve dynamically over time by accumulation of attribute-wise comparisons

14

265   between alternatives, and alternatives inhibit each other depending on their distance (or

266   similarity). MDFT does not assume any influence of gaze on the decision process. We also

267   included three baseline models in our comparison. First was a standard Expected Utility model

268   (EU)[14] that did not consider gaze data. Second, a static gaze baseline model GB_{stat} predicted

269   choices only from trial-aggregated gaze towards each alternative, ignoring outcome values and

270   outcome probabilities. Third, we constructed a dynamic gaze baseline model GB_{dyn} that, like

271   GLA, assumed leaky evidence accumulation over fixations, but ignored attribute values, using

272   only the sequence of fixations in each trial to predict choices. See Methods for details on these

273   models' implementation and parameter estimation.

274   All models were fit individually to the choice data of each participant. We compared the

275   models using the Bayesian Information Criterion (BIC)[37], which penalizes more complex models

276   and therefore accounts for differences in complexity between models.

277   Across the group, all models performed better than the random baseline model (Fig. 3a).

278   The GLA had the lowest BIC (mean $\pm$ s.d. = 230.63 $\pm$ 66.80), followed by the dynamic gaze

279   baseline model GB_{dyn} (mean $\pm$ s.d. = 302.46 $\pm$ 86.15), MDFT (mean $\pm$ s.d. = 345.77 $\pm$ 82.44),

280   EU (mean $\pm$ s.d. = 360.57 $\pm$ 72.56) and the static gaze baseline GB_{stat} (mean $\pm$ s.d. = 414.81 $\pm$

281   36.81). Simulating choices from the models using maximum-likelihood estimates, the proportion

282   of correctly predicted choices was 74.2% for GLA, 62.8% for GB_{dyn}, 57.9% for MDFT, 53.8%

283   for EU and 45.7% for GB_{stat} (see Supplementary Fig. 4 for distributions of model-predicted

284   choice probabilities). Note, that target and competitor options (and decoys in compromise trials)

285   were designed to be closely matched in value, resulting in trials with high choice difficulty,

286   limiting overall model performance. On the individual level, based on lowest BIC scores, the

287   majority (36 of 40, 90%) of participants were best described by GLA. Three participants (7.5%)

288   were best described by MDFT, and one by the dynamic gaze-baseline model (Fig. 3b). Protected

289   exceedance probabilities[38], which measure the likelihood that model is more frequent than all

290   others, unambiguously identified GLA as the most likely model ($pXP_{GLA} = 1$, Fig. 3b inset).

291   The estimates for GLA's gaze-discount parameter θ, which describes the degree to which

292   alternative's values are attenuated while not fixated, indicate that participants exhibited a

293   moderate gaze-discount on average: θ estimates ranged from 0.13 (strong gaze-discount) to 0.95

294   (almost no gaze-discount), with mean $\pm$ s.d. $= 0.69 \pm 0.18$. Estimates for the leak parameter

295   ranged from 0.08 (almost no leak) to 0.65 (moderately strong leak), with mean $\pm$ s.d. $= 0.29 \pm$

296   0.20. Individual parameter estimates of all GLA parameters are plotted in Supplementary Fig. 3

297   and summarised in Supplementary Table 1.

298   Given that MDFT outperforms utility-based models when choices are influenced by

299   context[31], we tested whether model fit of MDFT (relative to GLA) was higher for participants

300   with higher RST. In other words: Does MDFT perform better for stronger context effects?

301   Overall, 7 out of 9 participants with an RST above 0.7 in attraction trials were clearly best

302   described by GLA. While we found that the relative superiority of GLA over MDFT decreased

303   with strong attraction effects, this was not the case for compromise trials (Supplementary Fig. 5),

304   suggesting that some features of MDFT might help capture strong attraction effects.

305   As an additional indicator of model fit, we tested whether the models could quantitatively

306   reproduce the observed positive association of gaze and choice (Supplementary Figs. 1d, j, 2).

307   Specifically, following previous work[6,11], we inspected the model-predicted probability of

308   choosing an alternative as a function of its relative dwell time advantage: the difference in

309   relative dwell time towards an alternative minus the mean relative dwell time to other

310   alternatives. The probability of choosing an alternative generally increased with its relative dwell

311   time advantage (Fig. 3c-h), except for dominated decoys in attraction trials, which were not

312   chosen even if looked at longer than other alternatives (Fig. 3g).

313   All gaze-dependent models were able to capture this positive association. Note, however,

314   that they also predicted choices of dominated decoys in attraction trials, if decoys had a large

315   dwell time advantage (Fig. 3g). In this case, GLA performed better than GB_{dyn} and GB_{stat}, as it

316   predicted fewer decoy choices. However, MDFT and EU generally could not capture the

317   empirical association of gaze and choice; they predicted too many choices of alternatives with

318   negative dwell time advantage, and too few choices of alternatives with positive dwell time

319   advantage. Nor did they predict choices of dominated decoys, even if the decoy was looked at

320   longer.

321   Finally, we investigated the ability of the best-performing model to predict individual

322   differences in context effect strengths. Therefore, we predicted choices from the fitted GLA

323   model and correlated the resulting RST with the observed data. Predicted RST correlated

324   significantly with observed ones in attraction ($r = 0.76$, $HDI_{95} = [0.62, 0.90]$, Fig. 3i) and

325   compromise ($r = 0.82$, $HDI_{95} = [0.72, 0.92]$, Fig. 3j) trials. However, the model underestimates

326   large deviations from $RST = 0.5$, suggesting that its gaze-discount mechanism can capture the

327   qualitative pattern of context effects but not their expression in participants with extreme RSTs.

328   Inspection of predicted choice probabilities (Supplementary Fig. 4) shows that, on

329   average, GLA predicted high probabilities for the empirically chosen alternative (indicating good

330   overall fit), but comparable proportions of target and competitor choices (resulting in reduced

331   RST). Other models, like MDFT, predicted larger differences between target and competitor

332   choices for some participants, but assigned lower probability to empirically chosen alternatives

333   (resulting in overall inferior fit).

Systematic exploration of a large model space

The model comparison identified the advantage of the gaze-dependent accumulation model over its competitors. To better understand the contribution of individual model mechanisms (such as accumulation leak, inhibition between alternatives, or the gaze-dependent discount) to model performance, we performed a search across a large, systematically designed model space, in a so-called "switchboard analysis"[39]. Here, the idea is to isolate, group and exhaustively combine mechanistic assumptions of different models. Each group of mechanisms is considered a switch that can take different levels (e.g., an "inhibition" switch could take the levels "distance-dependent" as in MDFT, "constant" or "none"). Ultimately, this approach can be used to gauge the contribution of individual model mechanisms (opposed to evaluating whole models or model classes as in the more traditional model comparison presented above). In addition, it provides a systematic way to generate novel hybrid models, combining mechanisms from different *a priori* defined models.

This analysis approach further allowed us to test different assumptions about the functional forms of the gaze bias mechanism (e.g., as discounting non-fixated alternatives' values, controlling accumulation leak, among others[40]). We therefore expanded the range of gaze-dependent mechanisms from the original set of models to include additional eye-movement related switches, like attribute- and alternative-wise gaze-discounts, gaze-dependent leakage and inhibition. This allowed us to test if gaze-bias implementations different from the ones usually used in gaze-dependent accumulation[3,6,10–13,41,42].
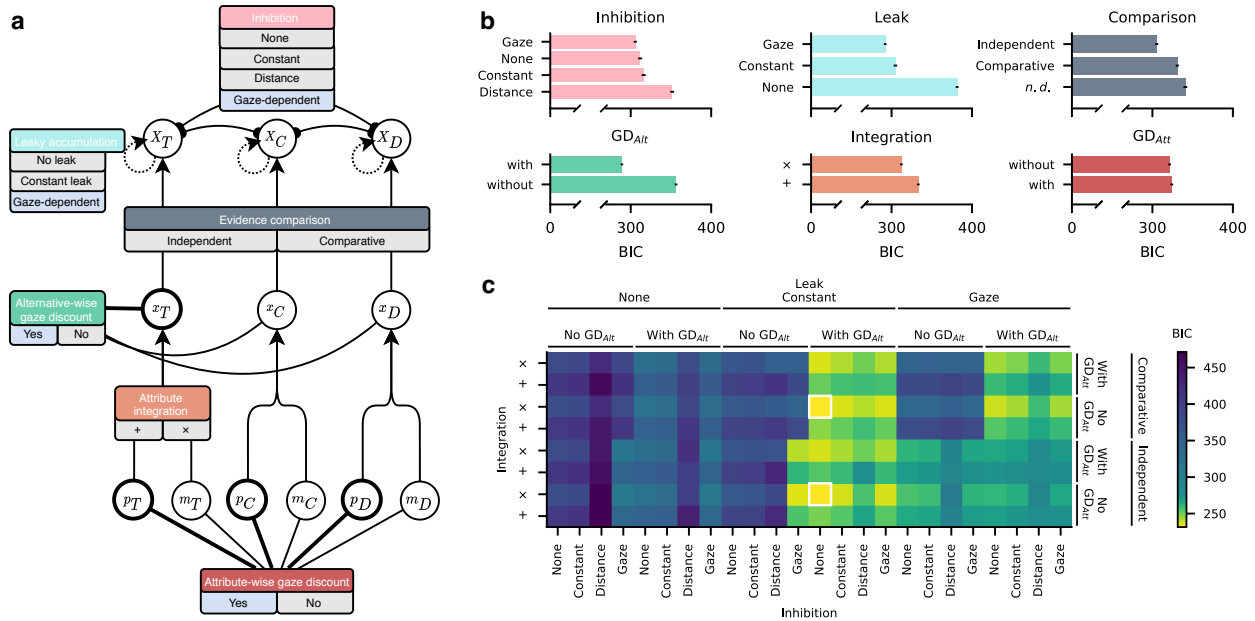
Figure 4: **Switchboard analysis. (a)** Overview of general switchboard framework and individual switches from which individual variants are constructed by setting the switches to a set of levels. Gaze-dependent switch levels are shaded blue. Attributes can be discounted based on gaze (lower level) and are integrated into alternative values (middle level). Alternative values can be discounted based on gaze, compared, and integrated (upper level) with leak and inhibition. **(b)** Average model fit associated with each switch's levels. Each bar shows the average BIC for all model variants that had the respective switch set to this level (e.g., first panel, top bar: average BIC of all variants with gaze-dependent inhibition). Gaze-dependent inhibition and leak, independent evidence accumulation, alternative-wise gaze-discount, multiplicative attribute integration, and no attribute-wise gaze-discount yielded lower BIC on average. **(c)** Overview of mean BIC for each of 192 model variants. More yellow colours indicate lower BIC and better model fit. The variant with the lowest BIC is identical to the GLA (alternative-wise, no attribute-wise gaze-discount, multiplicative attribute integration, constant leak, and no inhibition) and is outlined in white. Note that some variants were mathematically equivalent (see main text and Methods) including the variant with lowest BIC, which is therefore highlighted twice.

19

370   Our switchboard comprised a total of 192 model variants, derived from six switches that

371   were combined exhaustively: Attribute integration (additive vs. multiplicative), evidence

372   comparison (independent accumulation for each alternative or comparative accumulation of

373   evidence relative to other alternatives), alternative-wise and attribute-wise gaze-discount

374   (included or not), accumulation leak (none, constant or gaze-dependent) and inhibition between

375   alternatives (none, constant, gaze-dependent or distance-dependent). The models generally

376   resembled the form of the *a priori* defined GLA, but with substantial differences depending on

377   the specific set of switch levels (Fig. 4a; see Methods and Supplementary Table 2 for details on

378   the framework and switch levels). As before, each model variant was fit to the individual data of

379   each participant and model performance was evaluated based on the BIC score. Note that the

380   GLA coincides with one of the 192 variants (variant A in Table 2; multiplicative attribute

381   integration, alternative-wise gaze-discount, constant leak, no inhibition, no comparison).

382   Similarly, one variant (not in Table 2) conceptually resembles MDFT in some, but not all aspects

383   (additive attribute integration, comparative evidence accumulation, constant leak, distance-based

384   inhibition, strong attribute-wise gaze-discount).

385   **Best model mechanisms**   On the level of model mechanisms, multiplicative attribute

386   integration outperformed additive integration (mean $BIC_{mul.} = 312.25$, mean $BIC_{add.} = 333.20$;

387   Fig. 4b), inclusion of an alternative-wise gaze-discount (mean $BIC_{GD\ alt.} = 289.13$, mean

388    $BIC_{no\ GD\ alt.} = 356.32$), but not attribute-wise gaze-discount yielded lower BIC scores (mean

389    $BIC_{GD\ att.} = 324.11$, mean $BIC_{no\ GD\ att.} = 321.33$). Other gaze-dependent mechanisms also

390   improved model fit: Variants with gaze-dependent leak yielded lower BIC scores than variants

391   with constant or no leak ($BIC_{gaze} = 292.06$, $BIC_{constant} = 304.84$, $BIC_{none} = 381.49$). Gaze-

392   dependent inhibition performed better than constant, none or distance-dependent inhibition

393   ($BIC_{gaze} = 305.64$, $BIC_{constant} = 316.06$, $BIC_{none} = 311.75$, $BIC_{distance} = 351.57$). In summary, all

394 mechanisms that allowed a model to exploit the positive association between gaze and choice

395 improved model fit on average.

396   Counting switch-values of individually best fitting variants, most participants were best

397 described by model variants with multiplicative integration, with alternative-wise and no

398 attribute-wise gaze-discount, with constant accumulation leak parameter and no inhibition

399 (Supplementary Fig. 6).

	N	GD_{Alt}	GD_{Att}	Leak	Inhibition	Integration	Comparison
A	17	**Yes**	No	Constant	None	Multiplicative	*n.d.*
B	9	**Yes**	No	Constant	Distance	Multiplicative	Comparative
C	4	**Yes**	No	Constant	None	Additive	*n.d.*
D	3	No	No	**Gaze**	None	Multiplicative	Independent
E	2	No	No	Constant	**Gaze**	Multiplicative	Independent
F	1	**Yes**	No	**Gaze**	Distance	Multiplicative	Comparative
G	1	**Yes**	No	**Gaze**	Constant	Multiplicative	Comparative
H	1	**Yes**	**Yes**	Constant	Distance	Multiplicative	Independent
I	1	**Yes**	No	Constant	**Gaze**	Multiplicative	Comparative
J	1	**Yes**	No	**Gaze**	None	Multiplicative	Comparative

400 Table 2: **Overview of individually best fitting model variants.** N indicates the number of

401 participants best described by the variant described in the row. The top variant (A) coincided with

402 the GLA model. Note that all individually best fitting models had some form of gaze-dependence

403 (blue shaded cells, mostly alternative-wise gaze-discount). *"n.d."* denotes variants where

404 comparison mechanisms were not distinguishable by the analysis.

Best model variants   The overall best-fitting model variant was the variant identical to the GLA (Fig. 4c, Supplementary Table 3): It included multiplicative attribute integration, an alternative-wise gaze-discount, and constant leak. Note that our analysis did not allow us to distinguish independent and comparative accumulation for this variant, because they yield equivalent results. All of the ten best performing models used multiplicative attribute integration, and most used an alternative-wise but no attribute-wise gaze-discount and constant leak (Supplementary Table 3). Results were more ambiguous for the comparison and inhibition mechanisms.

Notably, one of the best ten variants employed a distance-dependent inhibition mechanism, and used other features resembling MDFT like constant leak, and accumulation of comparative values. Unlike MDFT, but like GLA, this variant used an alternative-wise gaze-discount, no attribute-wise mechanism of attention, and multiplicative integration of attributes. While not achieving the best overall fit, this hybrid variant performed significantly better than the original MDFT implementation (mean $BIC_{MDFT} = 345.77$; Fig. 3a).

The GLA variant also described 17 of 40 (42.5%) participants best (Table 2). Another similar variant with additive attribute integration described an additional four participants best. An additive relationship between attributes is typically assumed by models of multi-attribute multi-alternative choice[25,27]. Furthermore, additive integration of probability and outcome has recently been suggested as an alternative to multiplicative mechanisms and has been demonstrated to be equivalent for particular parameterisations and parameter values[43]. Five participants were best described by variants similar to the GLA, but using gaze-dependent leakage or inhibition instead of an alternative-wise gaze-discount. Note that gaze-dependent inhibition and leakage mechanisms can act similarly to an alternative-wise gaze-discount: All three mechanisms effectively reduce accumulated evidence of non-fixated alternatives. While the

429   alternative-wise gaze bias mechanism applies a constant discount to the accumulation inputs, the

430   gaze-dependent inhibition mechanism is proportional to the accumulated evidence of the

431   currently fixated alternative, and applies to the already accumulated evidence of other options,

432   not the inputs to the accumulation process. Gaze-dependent leakage similarly reduces already

433   accumulated evidence, proportional to the momentary accumulator value.

434   Notably, nine participants (22.5%) were best described by the previously described hybrid

435   variant using a distance-dependent inhibition mechanism (Table 2). Additional two participants

436   were best described by other variants using distance-dependent inhibition in conjunction with an

437   alternative-wise gaze-discount.

438   **Hybrid variant**   Finally, we analysed the hybrid model variant in more detail

439   (variant B in Table 2), which described 9 participants best. This variant performed especially well

440   for participants with large attraction effects (Fig. 5a), whereas GLA best described the majority

441   of participants with attraction RST around 0.5. In contrast, better-performing variants could not

442   be separated by compromise effect strength (Fig. 5b). The hybrid variant correctly predicted

443   individual differences in the attraction effect (correlation between observed and predicted RST: r

444   $= 0.92$, $HDI_{95} = [0.86\ 0.96]$; Fig. 5c). Here, it performed better than the GLA model (Fig. 3i), as

445   it was also able to capture the behaviour of participants with particularly strong attraction effects:

446   Using distance-dependent inhibition, it was able to predict high choice probabilities for target

447   alternatives in attraction trials for some participants, and fewer choices of competitor and

448   dominated decoy alternatives (Supplementary Fig. 4f). Predictions of individual RST in

449   compromise trials were almost identical between the two models ($r = 0.81$, $HDI_{95} = [0.70, 0.91]$;

450   Fig. 5d, see Fig. 3j for GLA).

23
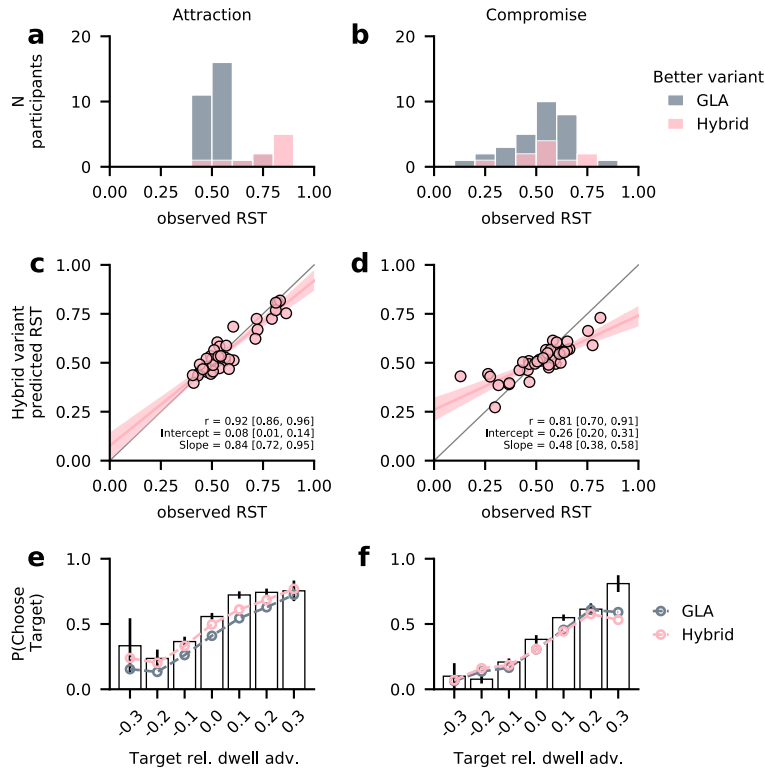
Figure 5: **Hybrid model variant details. (a, b)** Number of participants better described by the hybrid variant (pink) or the GLA (grey), dependent on strength of attraction **(a)** and compromise **(b)** effects. Participants with strong attraction effects were better described by the hybrid variant. **(c, d)** Individual observed and predicted RST in attraction **(c)** and compromise trials **(d)**. Compared to GLA (Fig. 3i, j), the hybrid model better predicted strong attraction effects for some participants. Predictions of compromise effects are similar. **(e, f)** Observed and model-predicted probability of choosing the target alternative, depending on the target's relative dwell time advantage. Like other gaze-dependent models (Fig. 3), the hybrid variant generally captured the positive association between gaze and choice. In contrast to GLA, however, it predicted an overall higher probability of choosing the target in attraction trials **(e)**. Predictions in compromise trials **(f)** are similar to GLA. White bars and error bars indicate mean $\pm$ s.e. observed data from even-numbered trials. Model predictions are based on 50 simulations for each odd-numbered trial.

465 The hybrid variant used an alternative-wise gaze-discount and could thus accurately

466 capture the relationship between relative dwell advantage and choice (Fig. 5e, f). Again, it

467 predicted an overall higher probability of target choices than GLA (Fig. 5e), and this was

468 primarily driven by the hybrid variant's predictions for individuals with strong attraction effects

469 (Supplementary Fig. 7). There was no meaningful difference between the two models in

470 compromise trials (Fig. 5f).

471   **Discussion**

472   In this study, we investigated whether risky choice behaviour could be characterized by a

473 gaze-dependent evidence accumulation framework, especially when choices are influenced by the

474 context of available alternatives. In line with previous findings, we found choice behaviour to be

475 context-dependent[44], but also subject to large interindividual differences. Importantly,

476 participants' gaze behaviour was also modulated by the context of available alternatives, allowing

477 a simple gaze-dependent evidence accumulation model derived from prior work on binary risky

478 choice[1] to provide the best description of their choices. Finally, in a systematic search across a

479 large space of possible model variants, we showed that gaze-dependent accumulation describes

480 all participants' behaviour best. Predicting data from participants with particularly strong

481 attraction effects, however, required inclusion of an additional similarity-based inhibition

482 mechanism.

483   Prior work could already demonstrate that choices between risky prospects can be

484 influenced by the set of available alternatives, producing attraction[18,21–23], similarity[20], and other

485 decoy effects[22]. These findings show that risky prospects, described by their winning probability

486 and outcome, are subject to the same context-dependent influences as other multi-attribute

487   stimuli, even though the natural (or normative) relationship between their attributes is

488   multiplicative and not additive.

489   Our findings add to this literature, replicating the attraction effect and providing novel

490   evidence for the compromise effect in risky choice. Notably, we find substantial individual

491   differences in the strength of the attraction effect, ranging from almost no to large effects. The

492   observed pattern includes a subgroup of participants predominantly choosing the target in

493   attraction trials, that is also present in previous work in inference[34] and risky choice tasks[21].

494   Importantly, we could not find any evidence that these participants used simplifying choice rules

495   based on the dominance relationship. The observed variability further emphasizes the importance

496   of focusing on individuals' behaviour instead of group aggregates, if the goal is to understand the

497   processes underlying individuals' choices[45].

498   Our study involved choices within sets of three risky gambles designed to elicit context

499   effects. Prior research makes contradicting predictions about the direction of information search

500   in this scenario: In the context of risky choice, empirical studies find a tendency towards within-

501   alternative processing, disagreeing with non-compensatory, heuristic approaches and providing

502   better support for compensatory strategies that assume integration of outcomes and

503   probabilities[46–48] (but see refs.[49,50], or in the related domain of intertemporal choice: ref.[51]). In the

504   domain of context effects, however, Noguchi & Stewart found pairwise comparisons between

505   alternatives on single attribute dimensions to be the dominant pattern of information search[52].

506   They argue that these comparisons should form the basis of choice models describing context

507   effects. Similarly, Marini *et al.* found that adding an asymmetrically dominated decoy to a choice

508   set shifts eye movements towards the target's dominant attribute, and results in more transitions

509   between target and decoy[29]. Cataldo & Cohen showed that the way information is displayed can

510   influence the size and direction of context effects[53]: Alternative-wise presentation yielded

similarity effects, whereas attribute-wise presentation, thought to induce comparisons between alternatives on single attributes, produced attraction and compromise effects. In line with the risky choice and context effects literature, we found participants to shift their gaze both within and between alternatives. While this does not constitute strong evidence for any particular process, this finding is compatible with current models of gaze-dependent accumulation in risky choice[1,3] and the GLA that assume within-alternative integration of probabilities and outcomes, and gaze-dependent accumulation and comparison processes to reach a decision.

Across decision making domains longer gaze towards an alternative is generally associated with a higher probability of choosing it[3–6,8,10,11,40,54–57]. This association is also present in choices between risky prospects[1,3,46,58]. While these results are mostly correlational, multiple studies found that manipulation of gaze towards an alternative increases its likelihood of being chosen, suggesting that gaze allocation influences choice[4,7–10,59].

The positive association between gaze and choice is also present in our data: Chosen gambles were looked at longer than others, and the effect increased over the course of a trial (see Supplementary Information). In addition, the probability of choosing an alternative increased with increasing relative dwell time. Gaze-dependent accumulation provides a formal account of the association between choice and gaze data, as unattended alternatives' accumulation is diminished, making them less likely to be chosen. Conversely, if context effects were present in participants' gaze, this would enable such models to predict context-dependent choice. Our data illustrate that this contextual modulation of gaze is indeed present.

Note that, in principle, GLA could produce a choice bias towards the alternative fixated last, by combining strong leakage with a strong gaze-discount: With a strong leak, predicted choices are influenced mainly by the information acquired in the final fixation. A strong gaze-

534   discount could then bias choice towards the fixated alternative. The obtained parameter estimates,

535   however suggest only moderate strengths of leak and discounting, indicating that the model's

536   good performance was not purely driven by effects of last fixations, which are often directed to

537   the chosen alternative.

538   Our results are closely related to recent work showing that another behavioural effect in

539   multi-alternative, multi-attribute choice is mediated by visual attention: Addition of a third

540   alternative to a choice set has been shown to affect choice accuracy through value-based

541   attentional capture in choices between risky prospects[2] and food items[55]. This mediation through

542   gaze, formalized by a gaze-dependent accumulation model, provided a better description of the

543   observed data than competing accounts. Adding to other work implicating mechanisms of visual

544   attention in the emergence of context effects[53,60], our work shows how gaze can mediate context

545   effects in a similar way: Choice sets affect the distribution of gaze, which in turn affects the

546   choice process.

547   Many traditional models of risky decision making assume that one scale value is assigned

548   to each alternative independent of the presence of others, and that choice probabilities are directly

549   derived from these values (e.g., Luce, 1959). These "simple scalability" theories include the most

550   influential models of risky decision making (e.g., Expected Utility Theory[14]; and Prospect

551   Theory,[15]). They obey rational axioms of choice like IIA and therefore cannot account for context

552   effects by design[24]. To explain context effects, multiple competing accounts have been

553   proposed[22,25–27,61–64] (see ref.[39], for a taxonomy of mechanisms and overview).

554   These models often assume that an alternative's value is computed in comparisons to other

555   alternatives on single attributes[25,27], that the considered attribute dimension switches

556   stochastically from moment to moment[25,27,61], and that choices result from accumulation (often

557 imperfect, i.e., leaky) of evidence until a threshold is reached[25–27,61]. Switching between attribute

558 dimensions can introduce correlations between accumulators for similar alternatives, generating

559 similarity effects[25,27,39,61]. In order to account for other context effects, these models employ

560 additional mechanisms: For example, loss aversion, that is, differential weighting of

561 advantageous and disadvantageous comparisons can produce attraction and compromise effects[27].

562 Distance-dependent inhibition between alternatives can yield similar results, by inhibiting similar

563 alternatives more strongly and bolstering alternatives that are similar but dominant[25]. Non-linear

564 value functions discounting alternatives with extreme attribute values can produce compromise

565 effects[26].

566   However, while they propose precise psychological processes leading up to decisions,

567 their relationship to observable process data, like eye movement recordings, remains implicit. For

568 example, the switching between attribute dimensions is often considered an attentional

569 mechanism[25,27,39,61], yet it is assumed to occur at every time-step (e.g., millisecond), and therefore

570 cannot be mapped to observable eye-movement data without additional assumptions. Notably,

571 thus far models of context-dependent choice do not include any gaze-dependency in the decision

572 process. This is in contrast to gaze-dependent accumulation models[1,5,6,11–13], which propose a

573 formal account of the association between gaze and choice.

574   In our study, context-dependent choices were best explained by a straightforward three-

575 alternative extension of a gaze-dependent accumulation model that was previously applied to

576 binary risky choices[1]. This model assumes that each alternative can be assigned a value by

577 multiplicative integration of probability and outcome attributes, independently of other

578 alternatives. Unlike simple scalable theories, however, it accumulates these values in a gaze-

579 dependent fashion until a choice is made. Through its gaze-dependence, this model was able to

580 predict individual differences in context effects. It performed best even across a large space of

581   models, which included variants using additive attribute integration, attribute-wise gaze-discount

582   and accumulation of comparative values. Such variants resemble extant models of context-

583   dependent choice (e.g., MDFT[25]), as they accumulate results from single attribute comparisons.

584   Yet they performed worse, even when they included gaze-dependency. Our results thus question

585   whether models of context-dependent choice must use attribute-wise comparisons over

586   alternative-wise integration of attributes. However, we also found that strong context effects

587   could be predicted best using an additional inhibition mechanism based on alternatives' similarity

588   (which is comparative in nature), while still using alternative-wise valuation at its core,

589   suggesting multiple parallel processes (i.e., within-alternative valuation *and* comparative

590   mechanisms).

591   More generally, our results suggest that extant models of context-dependent choice are

592   likely to benefit from implementing gaze-dependence. Even further, explicitly formalizing the

593   relationship between model variables and eye movements yields testable predictions that can help

594   distinguish and evaluate competing theories about the role of attention in context-dependent risky

595   choice. The identified class of models is compatible with observed transition data, quantitatively

596   captures the association of dwell time and choice probability, and uses the contextual modulation

597   of gaze in addition to a distance-dependent inhibition mechanism to predict context effects.

598   **Methods**

599   **Participants**

600   We recruited 44 participants for the experiment. All participants were required to have

601   normal or corrected to normal vision with soft contact lenses. Participants relying on glasses or

602   hard contact lenses were excluded from participation to ensure good eye tracking quality. Four

603   participants were excluded from the analyses: one due to a computer crash, two due to eye

604   tracking calibration worse than 1.0° of visual angle, and one because of misunderstanding task

605   instructions. The remaining 40 participants (25 female, 15 male) had mean $\pm$ s.d. age 27.2 $\pm$ 4.7.

606   All participants received a base compensation of 8 Euros per hour and could win an additional

607   bonus based on their choices during the experiment (see below). Written informed consent was

608   obtained from all participants prior to the experiment. The experimental procedures were

609   approved by Freie Universität's ethics committee.

610   **Task and stimuli**

611   Participants performed a value-based choice task with stimuli designed to elicit attraction

612   and compromise effects (Fig. 1). Each trial started with a 1.5 s fixation cross at the screen centre.

613   Then, three all-or-nothing gambles were presented next to each other on the screen. Gambles

614   were described by a probability p to win outcome m (and winning nothing otherwise). Each

615   gamble was enclosed by a rectangle. Gamble attributes p and m were arranged so that the vertical

616   distance between two attributes of one option was equal to the horizontal distance between the

617   centres of neighbouring alternatives. This distance was set to approximately 10.0° of visual angle.

618   Alternative positions and attribute positions within each gamble were random in each trial.

619   Participants were instructed to indicate their preference for one of the three gambles using their

620   right hand and the keyboard's arrow keys. There was no time limit. After their choice,

621   participants received a brief (0.3 s) feedback about their choice (but not about a gamble

622   outcome).

623   Participants were instructed that after completing the task, one of the trials would be

624   chosen randomly and the gamble chosen in this task would be played out for real money with a

625   virtual wheel of fortune, using a later to be disclosed payment multiplier. This multiplier was set

626   at 0.5 to scale winning bonuses to Freie Universität's payment standards.

627   Participants first performed three pairs of estimation and experiment blocks. Estimation

628   consisted of a maximum of 30 trials with choices between two alternatives. These blocks served

629   the purpose of determining individual indifference points for stimuli with varying levels of

630   winning probability p in an adaptive and integrated fashion. Participants were asked to indicate

631   their preference between a fixed reference gamble B ($p_B = 65\%$, $m_B = 15$ EUR) and less risky

632   option A ($p_A = 75\%$, $m_A = 10$ EUR). The outcome m_A for option A is then either increased (if B

633   was preferred) or decreased (if A was preferred) and the procedure repeated. Indifference points

634   for options C_A and C_B with probabilities $p_{C_A} = 85\%$ and $p_{C_B} = 55\%$ were determined interleaved

635   and in the same fashion. A single estimation block yielded a stimulus set with four options A, B,

636   C_A and C_B designed in a way that option pairs $A - B$, $A - C_A$ and $B - C_B$ were approximately

637   equally preferred and their distance in outcome mi was not less than 2 EUR. Additionally,

638   asymmetrically dominated range-frequency decoy options D_A and D_B were introduced and

639   designed to be 2% and 1 EUR worse than options A and B, respectively.

640   The following experimental blocks had 75 ternary choice trials each, 32 of which were

641   compromise trials balanced with respect to the target option (i.e., 16 choice sets $\{A, B, C_A\}$, 16

642   choice sets $\{A, B, C_B\}$), 32 attraction trials (16 choice sets $\{A, B, D_A\}$, 16 choice sets $\{A, B, D_B\}$)

643   and 11 distractor trials showing randomly created options with expected value of 10 EUR and

644   low (5–33%), medium (34–64%) and high (65–95%) probability p. For each trial, uniform noise

645   was added to each option's outcome m (± 0 EUR, $+1$ EUR) and probability p (-3%, $\pm 0\%$, $+3\%$).

646   Asymmetrically dominated decoys D_A and D_B received the same noise as their focal option, to

647   preserve dominance relation.

648   Participants performed 25 practice trials not relevant for their payout under supervision of

649   the experimenter.

Eye tracking

650

651   Participants' eye movements were recorded at 60 Hz using a table-mounted SMI Red eye

652   tracker (SensoMotoric Instruments, Teltow, Germany). Participants were placed approximately

653   60 cm in front of the screen and instructed to minimize head movements during the task. Before

654   each block, the eye tracker was calibrated using a 5-point calibration and validation procedure

655   until a spatial resolution smaller than 1.0° visual angle was achieved horizontally and vertically.

656   Participants were instructed to re-centre their gaze on the central fixation cross between trials.

657   Eye tracking data was pre-processed according to the following procedures: First,

658   fixations, saccades and blinks were detected using SMI's Event-Detector software. Minimum

659   fixation duration for detection was left at the default setting (80 ms). Blinks and saccades were

660   discarded. Fixations were truncated when participants made a keyboard response. Next,

661   rectangular areas of interest (AOIs) were constructed around the six screen locations that

662   displayed stimulus attributes. Fixations towards non-AOI regions of the screen were discarded if

663   they were preceded and followed by fixations to different AOIs. If they were preceded and

664   followed by fixations towards the same AOI, the non-AOI fixation was re-coded to that AOI,

665   too[6,11]. Finally, the total dwell time towards each alternative and attribute in each trial was

666   computed by summing all fixation durations towards respective AOIs. Relative dwell time was

667   computed by normalisation to the sum of all dwells in the trial.

Behavioural modelling

668

669   **Baseline: Random choice**   The random choice model predicts equal choice

670   probabilities $p = \frac{1}{3}$ for all three alternatives and serves as a benchmark against which other

671   models can be compared.

672   **Expected Utility**   Expected Utility Theory (EU)[14] assumes that choice behaviour can

673   be described as maximization of expected subjective utility. We computed subjective utilities of

674   option outcomes m_i using a power function with one free parameter α:

$$U(m_i) = m_i^\alpha \tag{3}$$

675

676   Predicted choice probabilities were computed using a soft-max choice rule[65] over expected

677   utilities $x_i^{EU} = p_i \cdot U(m_i)$:

$$p(x_i) = \frac{e^{\beta x_i}}{\sum_{j \in J} e^{\beta x_j}} \tag{4}$$

678

679   Here, J is the set of all available alternatives. The inverse temperature parameter β

680   controls the degree of randomness in the choice (choices become more deterministic with larger

681   β).

682   **Multialternative Decision Field Theory**   Multialternative Decision Field Theory

683   (MDFT)[25] is a dynamic connectionist model for multi-attribute, multi-alternative choice. MDFT

684   can predict similarity, attraction and compromise effects. The core principle of MDFT is the

685   calculation of valences $V(t)$ at each point in time t. Valences are determined as

$$V(t) = CMW(t) + \varepsilon(t) \tag{5}$$

686

687   where M is a matrix containing all alternatives' attributes. $W(t)$ is a stochastic vector

688   indicating the momentary weight distribution between attributes, according to a weight parameter

689   w. C is a contrast matrix, designed to perform attribute-wise contrasts between one option and the

690   mean of other options' attributes. Finally, $\varepsilon(t)$ is a stochastic normally distributed noise

691   component. Preferences $P(t)$ are generated integrating all valences $V(t)$ up to time t:

692
$$P(t) = SP(t-1) + V(t) \tag{6}$$

693   Critically, S is a square feedback matrix, thought to reflect the neurobiological mechanism

694   of lateral inhibition between alternatives. The diagonal elements of S determine how much the

695   current preference state is influenced by the previous one, controlled by the decay parameter φ_2.

696   Off-diagonal elements represent the feedback connections between alternatives. MDFT assumes

697   that the connection strength between two alternatives depends on their perceived distance D in

698   attribute space, scaled by sensitivity parameter φ_1. Taken together, the feedback matrix S is given

699   by

700
$$S = I - \varphi_2 \exp(-\varphi_1 D^2). \tag{7}$$

701   Here, I is the identity matrix. D is a matrix containing pairwise distances between

702   alternatives. We used the distance function formalized in Hotaling et al.[66], where the distance

703   between two alternatives is expressed in dominance- and indifference-directions and additionally

704   scaled in dominance direction with an overweighting parameter w_d. The preference vector $P(t)$

705   is asymptotically normally distributed with mean ξ and covariance matrix Ω[67], from which choice

706   probabilities are derived as in the general Thurstone model[68].

707   In total, this MDFT implementation includes five free parameters: The attribute weight w,

708   the decay parameter φ_2 , the sensitivity parameter φ_1, the overweighting parameter w_d and the

709   variance σ of the noise component $\varepsilon(t)$.

710   Typically, MDFT assumes an additive relationship between attributes. To accommodate

711   for this, we log-transformed stimulus attributes probability p and outcome m[21]. Stimulus

712   attributes were also rescaled to a range between 0 and 1 as in previous studies[31].

713   **Gaze-based models**   We included three models that use gaze data to predict choices:

714   Two baseline models that ignore stimulus information and predict choices only based on gaze

715   data, and one model adapted from previous work[1] that combines stimulus and gaze information in

716   a leaky accumulation framework.

717   ***Static gaze baseline model***   The first gaze-based model predicts choices from

718   participants' cumulated dwell times towards each alternative. It assumes that preference strength

719    $x_i^{GB_{\text{stat}}}$ for an alternative increases when it is fixated, irrespective of the its attributes p_i and m_i:

$$x_i^{GB_{\text{stat}}} = d_i \qquad (8)$$

720

721   where d_i is the total dwell time (in s) towards alternative i in a trial. Preference strengths

722    $x_i^{GB_{\text{stat}}}$ are transformed into choice probabilities using the soft-max function (Eq. (4)).

723   ***Dynamic gaze baseline model***   The second gaze-based model uses the whole

724   sequence of fixations in a trial to predict choices. It assumes that at each fixation, evidence in

725   favour of the fixated alternative is accumulated, and that accumulated evidence is subject to leak.

726   Formally,

$$X_i^{GB_{\text{dyn}}}(n) = \begin{cases} (1-\lambda) \cdot X_i^{GB_{\text{dyn}}}(n-1) + 1 & \text{if } i \text{ fixated} \\ (1-\lambda) \cdot X_i^{GB_{\text{dyn}}}(n-1) + 0 & \text{otherwise} \end{cases} \qquad (9)$$

727

728   where all $X_i^{GB_{\text{dyn}}}(0) = 0$. The λ parameter ($0 \le \lambda \le 1$) controls the strength of the

729   accumulation leak. Choice probabilities are computed from the final accumulator values using the

730   soft-max function (Eq. (4)).

731   ***Gaze-biased leaky accumulator model (GLA)***   Gaze-biased leaky accumulator

732   model (GLA). Finally, following a recent study on binary risky choice[1], we adapted a leaky

733   accumulator model[69], where option values are discounted depending on eye movements as in the

734   aDDM[6,11] and the related GLAM[12,13].

735   Here, the subjective utility for each alternative is computed by first applying a probability

736   weighting function[16], that transforms objective probabilities into subjective decision weights:

737
$$w(p) = \frac{p^\gamma}{(p^\gamma + (1 - p)^\gamma)^{\frac{1}{\gamma}}} \tag{10}$$

738   where γ ($0.28 \leq \gamma \leq 1$) is a free parameter controlling the shape of the weighting

739   function. If $\gamma = 1$, the subjective weights equal the objective probabilities[70]. Outcome utilities are

740   assumed to obtained from a power function as in the EU model (Eq. (3)). Subjective expected

741   utilities are then given by $x_i^{GLA} = w(p_i) \cdot U(m_i)$.

742   Next, this model assumes that for each alternative subjective expected utilities are

743   accumulated with leak and gaze bias over the time course of each trial. At each fixation n,

744   accumulators evolve according to

745
$$X_i^{GLA}(n) = \begin{cases} (1 - \lambda) \cdot X_i(n - 1) + 1 \cdot x_i^{GLA} & \text{if } i \text{ fixated} \\ (1 - \lambda) \cdot X_i(n - 1) + \theta \cdot x_i^{GLA} & \text{otherwise} \end{cases} \tag{11}$$

746   where all $X_i(0) = 0$. The θ parameter ($0 \leq \theta \leq 1$) controls discounting of unattended

747   alternative values. The λ parameter ($0 \leq \lambda \leq 1$) controls the strength of the accumulation leak.

748   Predicted choice probabilities are again computed from the soft-max function (Eq. (4))

749   over the final accumulator values X_i^{GLA}.

750   **Parameter estimation**   All models' parameters were estimated by minimizing the

751   negative summed log-likelihood $-\ln(\hat{L})$ of observed choices under the model. Minimization was

752   performed by a differential evolution algorithm[71] implemented in the SciPy Python library[72]. The

753   algorithm was provided sensible *a priori* defined bounds for each parameter and initialized

754   randomly.

755   **Model comparison**   Individually best-fitting models were identified based on the

756   Bayesian Information Criterion (BIC)[37], computed for each model m as

$$BIC_m = -2\ln(\hat{L}) + \ln(n_{trials})k_m \tag{12}$$

757

758   where k_m is the number of free parameters of model m and $\ln(\hat{L})$ is the summed log-

759   likelihood of observed choices under model m.

760   **Switchboard analysis**

761   We performed a switchboard analysis, similar to the one performed by Turner *et al.*[39] to

762   further investigate which components of the cognitive model are particularly important in

763   predicting the data. In a switchboard analysis, a cognitive model of the decision process is built,

764   where individual model mechanisms can take different forms, or levels, which can be switched

765   and combined with each other. One switch, or node, could for example be the integration of

766   attributes to form item values. This integration could happen multiplicatively, so that expected

767   outcomes are computed by multiplying outcome value and probability. It could also occur in a

768   weighted additive fashion, so that both outcome value and probability make independent

769   contributions to overall item value[43]. In the switchboard analysis, model variants using both

770   implementations, and combinations with all other levels of other nodes, are constructed and fit to

771   the behavioural data.

772   The switchboard analysis included different eye-movement related nodes, such as attribute

773   and alternative wise gaze biases or gaze-dependent leakage and inhibition, so that the

774   mechanisms that describe the data best can be identified and their relative contribution to model

775   fit can be measured. All switchboard models resembled the general form of the gaze-dependent

776   accumulation model presented in Glickman *et al.*[1] and the GLA adaptation to three items (see

777   above and Fig. 4a for a schematic). Here, evidence X_i in favour of each item is accumulated over

778   individual fixations. Accumulation can be subject to gaze-discount effects (so that non-fixated

779   items accumulate less evidence), leak and inhibition over time. Choice probabilities are computed

780   using a soft-max function (Eq. (4)) over the final accumulator values. The general accumulation

781   framework (in vector form, parallel for each item) is then given by

782
$$X(t) = S \times X(t-1) + \Theta Cx \tag{13}$$

783   where S is a square feedback matrix, combining the effects of accumulation decay (on its

784   diagonal elements) and inhibition between accumulators (on off diagonal elements). Θ is the

785   alternative-wise gaze discount vector (where the ith entry is set to 1 when item i is fixated, and

786   other entries are set to the discount value θ, $0 \leq \theta \leq 1$). C is a contrast matrix which, as in

787   MDFT, can perform comparisons between the entries of the item value vector x. We now

788   describe the different nodes and levels of the analysis, that are combined to generate the different

789   model variants:

790   **Attribute integration**   The attribute integration switch had two levels: First,

791   outcome probability and outcome value could be integrated *multiplicatively*, so that expected

792   outcome values per item are constructed. This level included the probability weighting function

793   w (Eq. (10)) using a free parameter γ ($0.28 \leq \gamma \leq 1$), and a utility function U (Eq. (3) free

794   parameter α ($0 \leq \alpha$). The item values are given as

795
$$x_i = w(p_i)U(m_i). \tag{14}$$

796   Alternatively, attribute integration could be implemented in a *weighted additive* fashion[43].

797   In this case, attributes were first normalized using divisive normalization[73] to make them

798   commensurable on a single scale:

799
$$p_i^{norm} = \frac{p_i}{\sum_i^n p_i} \tag{15}$$

800   and

801
$$m_i^{norm} = \frac{m_i}{\sum_i^n m_i}. \tag{16}$$

802   Next, the normalized attributes would be combined *additively*, with weighting parameter

803   w_p ($0 \leq w_p \leq 1$), controlling the relative contribution of the probability attribute:

804
$$x_i = w_p p_i^{norm} + (1 - w_p) m_i^{norm} \tag{17}$$

805   **Evidence comparison**   The evidence comparison switch had two levels: First, item

806   values x_i could be accumulated *independently* for each alternative, without comparison to other

807   alternatives. In this case the contrast matrix C is set to an identity matrix. Second, item values x_i

808   could be accumulated in a *comparative* fashion. Then, the contrast matrix C is set up to perform

809   comparisons between each input x_i and the mean of all other inputs $x_{j \neq i}$, as in MDFT. To this

810   end, diagonal entries of C are set to 1, and off-diagonal elements to $\frac{-1}{N-1}$, where N is the number

811   of alternatives (here $N = 3$).

812   **Alternative-wise gaze discount**   This switch could take the values *"on"* or *"off"*. If

813   switched on, the model included a free parameter θ ($0 \leq \theta \leq 1$) controlling the discount rate of

814   unattended alternatives during accumulation. If switched off, the gaze discount vector Θ was set

815   to one.

816   **Attribute-wise gaze discount**   The analysis also included the option of attribute-

817   wise gaze-dependent discounting (similar to the two-layer model from Glickman *et al.*[1] and the

818   model presented in Fisher[41]). If switched on, stimulus attributes of the currently unattended

819   dimension (e.g., probability, when an lottery outcome was fixated) were discounted by a free

820   parameter η ($0 \leq \eta \leq 1$). In combination with *additive* attribute integration, the attribute-wise

821   gaze discount was applied after attribute normalization, but prior to the weighted addition. For

822   *multiplicative* attribute integration, attributes were discounted before entering probability-

823   weighting and utility functions w and U.

824   **Accumulation leak**   We investigated three different forms of accumulation leak: First,

825   accumulation without leak. In these variants, the diagonal elements of the feedback matrix S were

826   set to 1, resulting in no leak. The second possibility was uniform *constant* leak, where we

827   estimated a parameter λ ($0 \leq \lambda \leq 1$, where 1 indicates perfect memory without leak, and 0

828   indicates leak of all prior information), occupying the diagonal elements of the feedback matrix

829   S. The third type of leak we investigated was *gaze-dependent*. Here, only accumulators of

830   unattended alternatives leak according to the λ parameter.

831   **Inhibition**   We considered four types of inhibition between accumulators: First,

832   independent accumulation without inhibition. In this case, all off-diagonal elements of S were set

833   to 0. Second, we considered uniform *constant* inhibition, where we estimated a parameter ϕ ($0 \leq$

834    $\phi \leq 1$) and set each off-diagonal element in S to $-\phi$, resulting in uniform inhibition

835   (proportional to the accumulators' activation level), across items. Thirdly, we considered *distance*

836   *dependent* inhibition, as implemented in MDFT (see Eq. (7)). Here, the inhibition between

837   accumulators is a function of the corresponding items' distance in attribute space. The distance is

838   expressed in indifference and dominance directions, and the dominance direction is overweighted

839   by a parameter w_d. As we did for MDFT, we log-transformed probability and outcome attributes

840   and rescaled them to a range between 0 and 1 before applying the distance function. This

841   implementation uses a sensitivity parameter ϕ, a parameter estimating the relative importance of

842   the probability attribute w_p (this parameter is already used if attribute integration is *additive*), and

843   the overweighting parameter of the dominance direction w_d. Note, we only computed off-

844   diagonal elements of S according to Eq. (7), as the diagonal entries were controlled by the

845   accumulation leak switch. Finally, we considered *gaze-dependent* inhibition. Here, the rationale

846   is that only the accumulator of the currently attended item exerts inhibition over all others. In this

847   type of inhibition, all off-diagonal elements of S in the column corresponding to the currently

848   attended item are set to $-\phi$, and others are set to 0.

849   **Total number of variants and parameter estimation**   Exhaustive combination of all

850   switch levels yields 192 model variants. The effective number of uniquely identifiable models

851   was, however, reduced to 160 because for some variants comparative and independent

852   accumulation versions cannot be distinguished when choice probabilities are derived from a soft-

853   max choice rule with a freely estimated inverse temperature parameter over final accumulator

854   values. This is the case for variants with no or constant inhibition and leak. Each variant was fit

855   individually to the data from each participant by maximum-likelihood estimation, using a

856   Differential Evolution optimization algorithm (see above). As the number of parameters differ

857   between model variants, we computed the BIC for each model and participant to obtain a

858   measure of model fit, corrected for model complexity.

859   The optimization algorithm failed to find a solution better than a random model for 108 of

860   6400 (1.69%) of estimation runs. Since all model variants used the soft-max choice rule (Eq. (4))

861   and therefore could predict random choices by setting the inverse temperature parameter β to 0,

862   this indicates non-convergence of the optimization algorithm. All but one non-converged

863   estimation run used distance-dependent inhibition. We set the maximum-likelihood estimates of

864   the failed runs to that of the nested random model for all analyses.

865   **Statistical modelling**

866   We used Bayesian estimation (BEST)[32,33] of differences, effect size d and associated 95%

867   highest posterior density intervals (HDI$_{95}$) for all reported pairwise comparisons. Reported

868   correlation coefficients and associated HDI$_{95}$ result from Bayesian correlation analyses[74].

869   Regression estimates and HDI$_{95}$ result from Bayesian regression analyses implemented in

870   PyMC3[75].

871   **Software**

872   The task was programmed in MATLAB (The Mathworks Inc., USA) using the

873   PsychToolBox[76]. Data processing and analyses were done in Python with numpy[77], scipy[72] and

874   pandas[78] libraries. Bayesian analyses were implemented in PyMC3[75], mixed models used

875   bambi[79]. Exceedance probabilities were computed in MATLAB using SPM12[80]. Figures were

876   created using matplotlib[81], seaborn[82] and python-ternary[83].

877   **Data and code availability**

878   All raw and preprocessed data, and scripts to reproduce all reported processing, analyses

879   and figures are available at https://github.com/moltaire/gda-context.

Author contributions

883

884   F.M., S.A.H. and P.N.C.M. designed the study. F.M. and A.W.T. conceptualized the

885   behavioural models. F.M. performed the experiment, wrote all software, analysed the data and

886   visualized results. P.N.C.M. and H.R.H. supervised the project. P.N.C.M. acquired funding. F.M.

887   wrote the original draft. All authors reviewed and edited the manuscript.

Competing interests

888

889   The authors declare no competing interests.

References

1. Glickman, M. *et al.* The formation of preference in risky choice. *PLOS Computational Biology* **15**, e1007201 (2019).

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

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

4. Armel, K. C., Beaumel, A. & Rangel, A. Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making* **3**, 8 (2008).

5. Cavanagh, J. F., Wiecki, T. V., Kochar, A. & Frank, M. J. Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology. General* **143**, 1476–1488 (2014).

6. Krajbich, I., Armel, C. & Rangel, A. Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience* **13**, 1292–1298 (2010).

7. Pärnamets, P. *et al.* Biasing moral decisions by exploiting the dynamics of eye gaze. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 4170–4175 (2015).

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

9. Sui, X.-Y., Liu, H.-Z. & Rao, L.-L. The timing of gaze-contingent decision prompts influences risky choice. *Cognition* **195**, 104077 (2020).

911   10.   Tavares, G., Perona, P. & Rangel, A. The Attentional Drift Diffusion Model of Simple

912   Perceptual Decision-Making. *Frontiers in Neuroscience* **11**, 468 (2017).

913   11.   Krajbich, I. & Rangel, A. Multialternative drift-diffusion model predicts the relationship

914   between visual fixations and choice in value-based decisions. *Proceedings of the National*

915   *Academy of Sciences of the United States of America* **108**, 13852–13857 (2011).

916   12.   Molter, F., Thomas, A. W., Heekeren, H. R. & Mohr, P. N. C. GLAMbox: A Python

917   toolbox for investigating the association between gaze allocation and decision behaviour.

918   *PLoS ONE* **14**, (2019).

919   13.   Thomas, A. W., Molter, F., Krajbich, I., Heekeren, H. R. & Mohr, P. N. C. Gaze bias

920   differences capture individual choice behaviour. *Nature Human Behaviour* **3**, 625–635

921   (2019).

922   14.   Von Neumann, J. & Morgenstern, O. *Theory of games and economic behavior*. (Princeton

923   University Press, 1947).

924   15.   Kahneman, D. & Tversky, A. Prospect Theory: An Analysis of Decision under Risk.

925   *Econometrica* **47**, 263–291 (1979).

926   16.   Tversky, A. & Kahneman, D. Advances in prospect theory: Cumulative representation of

927   uncertainty. *Journal of Risk and Uncertainty* **5**, 297–323 (1992).

928   17.   Luce, R. D. *Individual choice behavior*. (John Wiley, 1959).

929   18.   Huber, J., Payne, J. W. & Puto, C. Adding Asymmetrically Dominated Alternatives:

930   Violations of Regularity and the Similarity Hypothesis. *Journal of Consumer Research* **9**,

931   90–98 (1982).

932   19.   Simonson, I. Choice Based on Reasons: The Case of Attraction and Compromise Effects.

933   *Journal of Consumer Research* **16**, 158–174 (1989).

934   20.   Tversky, A. Elimination by aspects: A theory of choice. *Psychological Review* **79**, 281–

935   299 (1972).

936   21.   Mohr, P. N. C., Heekeren, H. R. & Rieskamp, J. Attraction Effect in Risky Choice Can Be

937   Explained by Subjective Distance Between Choice Alternatives. *Scientific Reports* **7**, 1–10

938   (2017).

939   22.   Soltani, A., De Martino, B. & Camerer, C. A Range-Normalization Model of Context-

940   Dependent Choice: A New Model and Evidence. *PLoS Computational Biology* **8**, (2012).

941   23.   Wedell, D. H. Distinguishing among models of contextually induced preference reversals.

942   *Journal of Experimental Psychology: Learning, Memory, and Cognition* **17**, 767–778

943   (1991).

944   24.   Rieskamp, J., Busemeyer, J. R. & Mellers, B. A. Extending the Bounds of Rationality:

945   Evidence and Theories of Preferential Choice. *Journal of Economic Literature* **44**, 631–

946   661 (2006).

947   25.   Roe, R. M., Busemeyer, J. R. & Townsend, J. T. Multialternative decision field theory: A

948   dynamic connectionst model of decision making. *Psychological Review* **108**, 370–392

949   (2001).

950   26.   Trueblood, J. S., Brown, S. D. & Heathcote, A. The multiattribute linear ballistic

951   accumulator model of context effects in multialternative choice. *Psychological Review*

952   **121**, 179–205 (2014).

953 27. Usher, M. & McClelland, J. L. Loss Aversion and Inhibition in Dynamical Models of

954   Multialternative Choice. *Psychological Review* **111**, 757–769 (2004).

955 28. Tversky, A. & Simonson, I. Context-Dependent Preferences. *Management Science* **39**,

956   1179–1189 (1993).

957 29. Marini, M., Ansani, A. & Paglieri, F. Attraction comes from many sources: Attentional

958   and comparative processes in decoy effects. *Judgment and Decision Making* **15**, 23

959   (2020).

960 30. Cohen, A. L., Kang, N. & Leise, T. L. Multi-attribute, multi-alternative models of choice:

961   Choice, reaction time, and process tracing. *Cognitive Psychology* **98**, 45–72 (2017).

962 31. Berkowitsch, N. A. J., Scheibehenne, B. & Rieskamp, J. Rigorously testing

963   multialternative decision field theory against random utility models. *Journal of*

964   *Experimental Psychology: General* **143**, 1331–1348 (2014).

965 32. Kruschke, J. K. Bayesian estimation supersedes the t test. *Journal of Experimental*

966   *Psychology: General* **142**, 573–603 (2013).

967 33. Kruschke, J. K. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*.

968   (Academic Press, 2014).

969 34. Trueblood, J. S. Multialternative context effects obtained using an inference task.

970   *Psychonomic Bulletin & Review* **19**, 962–968 (2012).

971 35. Trueblood, J. S., Brown, S. D. & Heathcote, A. The fragile nature of contextual preference

972   reversals: Reply to Tsetsos, Chater, and Usher (2015). *Psychological Review* **122**, 848–

973   853 (2015).

974   36.   Trueblood, J. S., Brown, S. D., Heathcote, A. & Busemeyer, J. R. Not Just for Consumers:

975   Context Effects Are Fundamental to Decision Making. *Psychological Science* **24**, 901–

976   908 (2013).

977   37.   Schwarz, G. Estimating the Dimension of a Model. *The Annals of Statistics* **6**, 461–464

978   (1978).

979   38.   Rigoux, L., Stephan, K. E., Friston, K. J. & Daunizeau, J. Bayesian model selection for

980   group studies Revisited. *NeuroImage* **84**, 971–985 (2014).

981   39.   Turner, B. M., Schley, D. R., Muller, C. & Tsetsos, K. Competing theories of

982   multialternative, multiattribute preferential choice. *Psychological Review* **125**, 329–362

983   (2018).

984   40.   Ashby, N. J. S., Jekel, M., Dickert, S. & Glöckner, A. Finding the right fit: A comparison

985   of process assumptions underlying popular drift-diffusion models. *Journal of*

986   *Experimental Psychology: Learning, Memory, and Cognition* **42**, 1982–1993 (2016).

987   41.   Fisher, G. An attentional drift diffusion model over binary-attribute choice. *Cognition* **168**,

988   34–45 (2017).

989   42.   Krajbich, I., Lu, D., Camerer, C. & Rangel, A. The attentional drift-diffusion model

990   extends to simple purchasing decisions. *Frontiers in Psychology* **3**, 193 (2012).

991   43.   Rouault, M., Drugowitsch, J. & Koechlin, E. Prefrontal mechanisms combining rewards

992   and beliefs in human decision-making. *Nature Communications* **10**, 1–16 (2019).

993   44.   Busemeyer, J. R., Gluth, S., Rieskamp, J. & Turner, B. M. Cognitive and Neural Bases of

994   Multi-Attribute, Multi-Alternative, Value-based Decisions. *Trends in Cognitive Sciences*

995   **23**, 251–263 (2019).

996   45.   Liew, S. X., Howe, P. D. L. & Little, D. R. The appropriacy of averaging in the study of

997   context effects. *Psychonomic Bulletin & Review* **23**, 1639–1646 (2016).

998   46.   Fiedler, S. & Glöckner, A. The Dynamics of Decision Making in Risky Choice: An Eye-

999   Tracking Analysis. *Frontiers in Psychology* **3**, (2012).

1000   47.   Glöckner, A. & Herbold, A.-K. An eye-tracking study on information processing in risky

1001   decisions: Evidence for compensatory strategies based on automatic processes. *Journal of*

1002   *Behavioral Decision Making* **24**, 71–98 (2011).

1003   48.   Johnson, E. J., Schulte-Mecklenbeck, M. & Willemsen, M. C. Process models deserve

1004   process data: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological*

1005   *Review* **115**, 263–272 (2008).

1006   49.   Russo, J. E. & Dosher, B. A. Strategies for multiattribute binary choice. *Journal of*

1007   *Experimental Psychology: Learning, Memory, and Cognition* **9**, 676–696 (1983).

1008   50.   Su, Y. *et al.* Is making a risky choice based on a weighting and adding process? An eye-

1009   tracking investigation. *Journal of Experimental Psychology: Learning, Memory, and*

1010   *Cognition* **39**, 1765–1780 (2013).

1011   51.   Amasino, D. R., Sullivan, N. J., Kranton, R. E. & Huettel, S. A. Amount and time exert

1012   independent influences on intertemporal choice. *Nature Human Behaviour* **3**, 383–392

1013   (2019).

1014   52.   Noguchi, T. & Stewart, N. In the attraction, compromise, and similarity effects,

1015   alternatives are repeatedly compared in pairs on single dimensions. *Cognition* **132**, 44–56

1016   (2014).

1017   53.   Cataldo, A. M. & Cohen, A. L. The comparison process as an account of variation in the

1018   attraction, compromise, and similarity effects. *Psychonomic Bulletin & Review* **26**, 934–

1019   942 (2019).

1020   54.   Folke, T., Jacobsen, C., Fleming, S. M. & Martino, B. D. Explicit representation of

1021   confidence informs future value-based decisions. *Nature Human Behaviour* **1**, 1–8 (2016).

1022   55.   Gluth, S., Kern, N., Kortmann, M. & Vitali, C. L. Value-based attention but not divisive

1023   normalization influences decisions with multiple alternatives. *Nature Human Behaviour*

1024   1–12 (2020) doi:10.1038/s41562-020-0822-0.

1025   56.   Towal, R. B., Mormann, M. & Koch, C. Simultaneous modeling of visual saliency and

1026   value computation improves predictions of economic choice. *Proceedings of the National*

1027   *Academy of Sciences* **110**, E3858–E3867 (2013).

1028   57.   Vaidya, A. R. & Fellows, L. K. Testing necessary regional frontal contributions to value

1029   assessment and fixation-based updating. *Nature Communications* **6**, 10120 (2015).

1030   58.   Stewart, N., Hermens, F. & Matthews, W. J. Eye Movements in Risky Choice: Eye

1031   Movements in Risky Choice. *Journal of Behavioral Decision Making* **29**, 116–136 (2016).

1032   59.   Liu, H.-Z. *et al.* Exploiting the dynamics of eye gaze to bias intertemporal choice. *Journal*

1033   *of Behavioral Decision Making* **n/a**, (2020).

1034   60.   Evans, N. J., Holmes, W. R., Dasari, A. & Trueblood, J. S. The impact of presentation
1035   order on attraction and repulsion effects in decision-making. *Decision* **8**, 36–54 (2021).

1036   61.   Bhatia, S. Associations and the accumulation of preference. *Psychological Review* **120**,
1037   522–543 (2013).

1038   62.   Noguchi, T. & Stewart, N. Multialternative decision by sampling: A model of decision
1039   making constrained by process data. *Psychological Review* **125**, 512–544 (2018).

1040   63.   Wollschläger, L. M. & Diederich, A. The 2N-ary Choice Tree Model for N-Alternative
1041   Preferential Choice. *Frontiers in Psychology* **3**, (2012).

1042   64.   Tsetsos, K., Chater, N. & Usher, M. Salience driven value integration explains decision
1043   biases and preference reversal. *Proceedings of the National Academy of Sciences* **109**,
1044   9659–9664 (2012).

1045   65.   Sutton, R. S. & Barto, A. G. *Reinforcement learning: An introduction*. (The MIT Press,
1046   2018).

1047   66.   Hotaling, J. M., Busemeyer, J. R. & Li, J. Theoretical developments in decision field
1048   theory: Comment on Tsetsos, Usher, and Chater (2010). *Psychological Review* **117**, 1294–
1049   1298 (2010).

1050   67.   Busemeyer, J. R. & Diederich, A. Survey of decision field theory. *Mathematical Social*
1051   *Sciences* **43**, 345–370 (2002).

1052   68.   Böckenholt, U. Multivariate models of preference and choice. in *Multidimensional models*
1053   *of perception and cognition* 89–113 (Lawrence Erlbaum Associates, Inc, 1992).

1054   69.   Usher, M. & McClelland, J. L. The time course of perceptual choice: The leaky,

1055   competing accumulator model. *Psychological Review* **108**, 550–592 (2001).

1056   70.   Cavagnaro, D. R., Pitt, M. A., Gonzalez, R. & Myung, J. I. Discriminating among

1057   probability weighting functions using adaptive design optimization. *Journal of Risk and*

1058   *Uncertainty* **47**, 255–289 (2013).

1059   71.   Storn, R. & Price, K. Differential Evolution A Simple and Efficient Heuristic for global

1060   Optimization over Continuous Spaces. *Journal of Global Optimization* **11**, 341–359

1061   (1997).

1062   72.   Jones, E., Oliphant, T. & Peterson, P. SciPy: Open Source Scientific Tools for Python.

1063   (2001).

1064   73.   Louie, K., Khaw, M. W. & Glimcher, P. W. Normalization is a general neural mechanism

1065   for context-dependent decision making. *Proceedings of the National Academy of Sciences*

1066   **110**, 6139–6144 (2013).

1067   74.   Lee, M. D. & Wagenmakers, E.-J. *Bayesian cognitive modeling: A practical course.*

1068   (Cambridge University Press, 2013).

1069   75.   Salvatier, J., Wiecki, T. V. & Fonnesbeck, C. Probabilistic programming in Python using

1070   PyMC3. *PeerJ Computer Science* **2**, e55 (2016).

1071   76.   Brainard, D. H. The psychophysics toolbox. *Spatial vision* **10**, 433–436 (1997).

1072   77.   Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).

1073   78.   McKinney, W. *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and*

1074   *IPython.* ("O'Reilly Media, Inc.", 2012).

1075   79.   Yarkoni, T. & Westfall, J. Bambi: A simple interface for fitting Bayesian mixed effects

1076   models. (2016) doi:10.31219/osf.io/rv7sn.

1077   80.   Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J. & Nichols, T. E. *Statistical*

1078   *parametric mapping: The analysis of functional brain images*. (Elsevier, 2011).

1079   81.   Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science &*

1080   *Engineering* **9**, 90–95 (2007).

1081   82.   Waskom, M. & team, the seaborn development. Mwaskom/seaborn. (2020)

1082   doi:10.5281/zenodo.592845.

1083   83.   Harper, M. & others. Python-ternary: Ternary plots in python. *Zenodo*

1084   *10.5281/zenodo.594435* (2015) doi:10.5281/zenodo.594435.