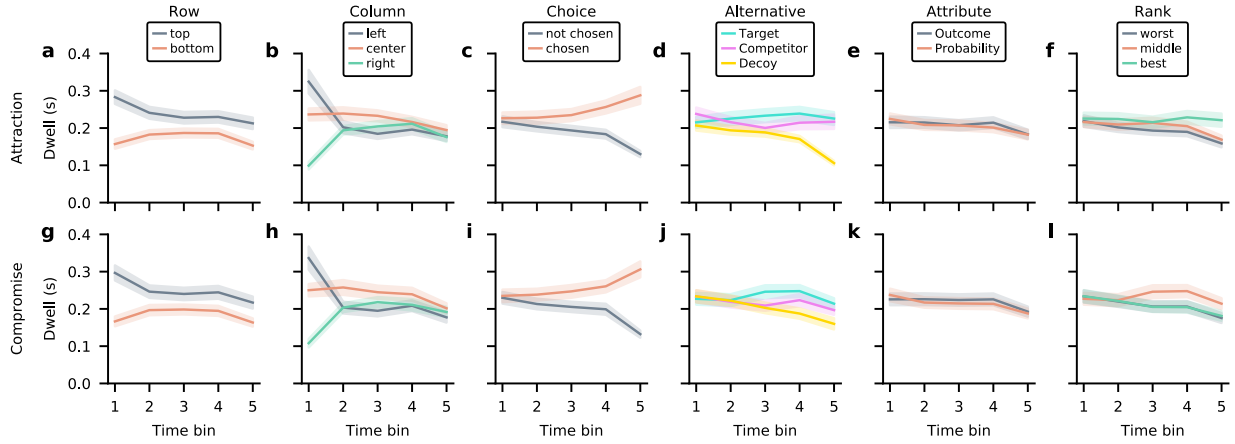
Supplementary Information



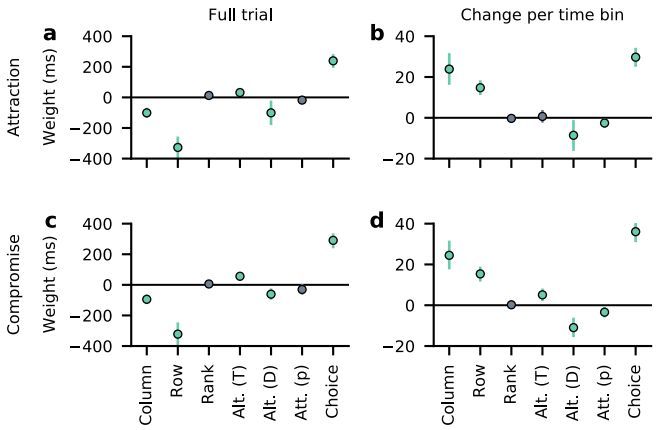
Supplementary Figure 1: **Distribution of gaze over the course of the trial depending on stimulus characteristics.** Each panel shows the average dwell time towards AOIs for a given stimulus feature (e.g., horizontal and vertical position) across five time-bins. Data is shown separately for attraction **(a-f)** and compromise **(g-l)** trials.

Supplementary Note 1

Regression analyses of gaze behaviour We performed two linear mixed effects regressions of total dwell time towards an AOI in each trial, separate for attraction and compromise trials (Supplementary Fig. 1). In the first ("full trial") model, the dependent variable was total dwell time towards an AOI across the full trial. This model used the following predictors: Vertical position (row, centred: -0.5 = top row, +0.5, bottom row), horizontal position (column, centred: -1 = left, 0 = centre, +1 = right), attribute (dummy coding probability attribute *p*), within dimension attribute rank (centred, -1 = worst, 0 = intermediate, +1 = best on attribute), target (dummy coded), decoy (dummy coded) and a dummy coded predictor for the ultimately chosen alternative. For the second model, we partitioned the dwell-data into five equally sized

time-bins. The dependent variable in this model then was total dwell time towards an AOI within a time-bin. This model included the same predictors as the "full trial" model. Crucially, it also included interaction terms for each predictor with the time-bin variable, and time-bin as additional predictor. Both models included random intercepts and slopes for each participant. Bayesian posterior distributions of the regression weights were estimated using the bambi Python library[1], with default priors[2], sampling four chains with 2000 samples each, after a tuning phase of 500 samples. Convergence was diagnosed visually and by means of the Gelman-Rubin statistic ($|1 - \hat{R}| \leq 0.05$ for all chains).

Regression weight estimates are shown in Supplementary Fig. 2. Across the trial, we find strong effects of position on dwell time, so that dwell times towards the top and left were longer. Significant negative interaction effects of the row and column predictors with time showed that these effects diminish across the trial. We also find a gaze-cascade effect[3,4], where dwell times to AOIs belonging to the ultimately chosen alternative are longer across the trial (and increasingly so throughout the trial, indicated by the positive interaction term with time). Dwell times to decoys decreased significantly during the trial, and across the full trial, dwell times to decoys were shorter than other alternatives. Similarly, dwell times towards probability attributes *p* shortened across the trial. Across the full trial, however, dwell times towards probability AOIs were not shorter than those to outcome AOIs. Finally, dwell times to target alternatives were longer across the trial in both compromise and attraction trials. In addition, this effect increased throughout the trial in compromise, but not in attraction trials. Note that these effects are independent of the effect of choice, as choice is a separate predictor in the model. We could not find an association between the attribute rank (being the worst, best, or intermediate item on an attribute) and dwell time.



Supplementary Figure 2: **Weight estimates from regression analyses on absolute dwell times.** We performed two mixed-effects regression analyses of dwell time towards each AOI onto stimulus properties: **(a, c)** Regressing the total dwell time towards an AOI over a full trial onto AOI column, row, attribute rank (best, middle or worst value on the attribute), two dummy predictors coding alternative, attribute (probability or outcome) and whether the AOI belonged to the subsequently chosen alternative. **(b, d)** Second, we binned dwell times in each trial into five time-bins and added an interaction term with time-bin for each predictor. The panels show the interaction term weights. Analyses were carried out separately for attraction **(a, b)** and compromise **(c, d)** trials. Regression models had random intercepts and slopes across participants. Points and intervals mark posterior mean estimates and associated HDI_{95} (coloured green if the interval excluded 0).

Direction of information search We further analysed participants' direction of information search. Therefore, we counted the number of vertical (transitions within the same alternative), horizontal (within the same row, between alternatives) and diagonal (between rows and alternatives) transitions. On average, participants made over 7 horizontal transitions in attraction (mean ± s.d. = 7.28 ± 2.36) and compromise (mean ± s.d. = 7.29 ± 2.63) trials, with no meaningful difference between effects. Participants made, however, more vertical transitions in compromise trials (mean ± s.d. = 7.55 ± 3.11) than attraction trials (mean ± s.d. = 7.17 ± 3.06; mean difference = 0.39, HDI_{95} = [0.11, 0.64], d = 0.49, HDI_{95} = [0.13, 0.82]). The number

of diagonal transitions was lower overall, but higher in attraction (mean $\pm$ s.d. = 2.88 $\pm$ 1.18) than compromise trials (mean $\pm$ s.d. = 2.72 $\pm$ 1.31; mean difference = 0.19, HDI_{95} = [0.06, 0.32], d = 0.62, HDI_{95} = [0.14, 1.17]).

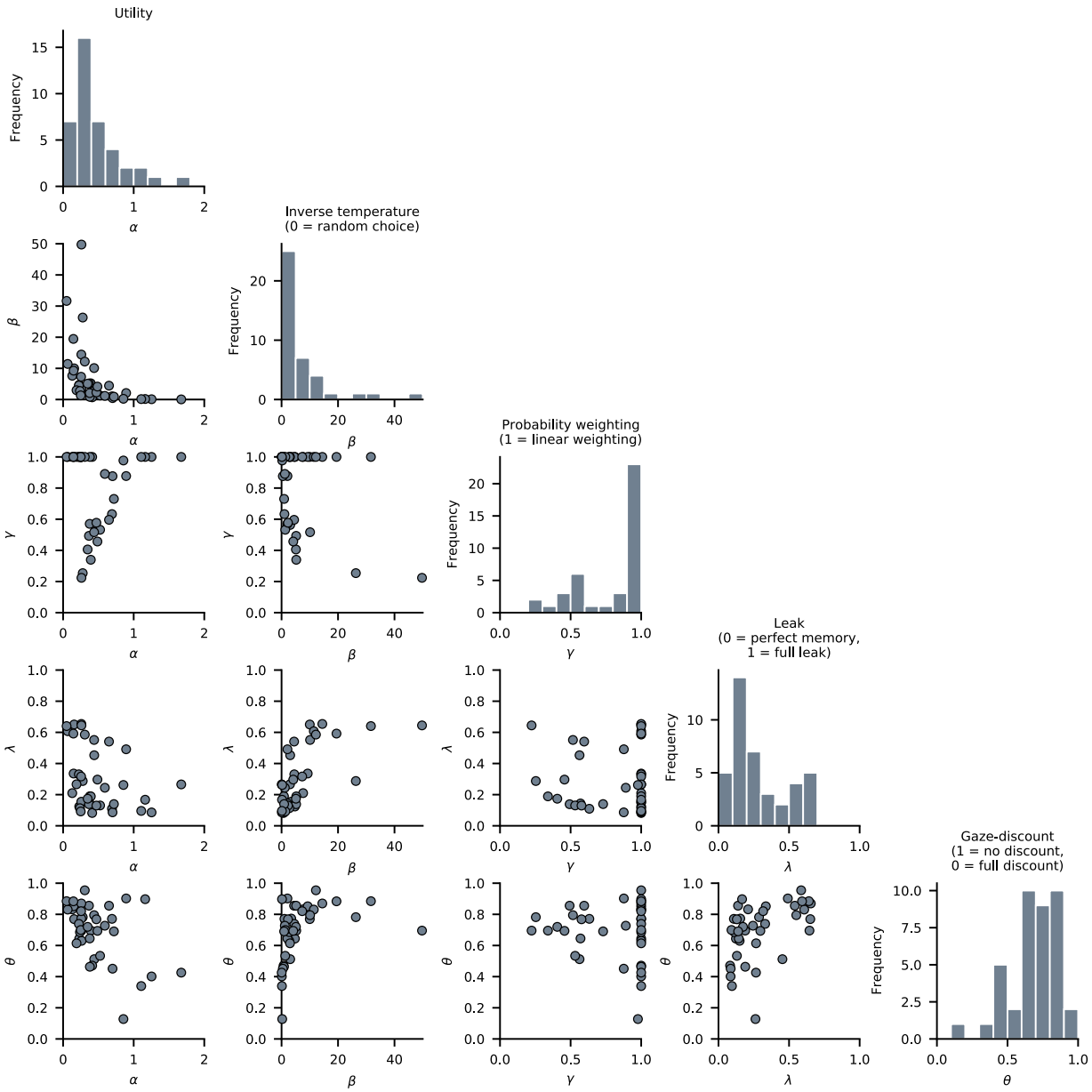
While vertical transitions always translate to transitions "within alternative", horizontal transitions are not necessarily always "within attribute", since the attribute positions of each alternative were random in the task. We therefore recoded transitions as "within alternative", "within attribute" and "between alternatives and attributes" and computed the Payne Index[5] for each trial as:

$$\text{Payne Index} = \frac{N_{\text{within alt.}} - N_{\text{within att.}}}{N_{\text{within alt.}} + N_{\text{within att.}}} \quad (18)$$

A more positive value on the index indicates more processing within alternatives, whereas more negative values indicate more processing between alternatives, within the same attribute dimension. Overall, the average Payne Index was slightly positive for both attraction (mean $\pm$ s.d. = 0.10 $\pm$ 0.16) and compromise trials (mean $\pm$ s.d. = 0.14 $\pm$ 0.18), suggesting a mixture of within-alternative and within-attribute processing, with slightly more processing within alternatives. It was, however, lower in attraction trials (mean difference = 0.04, HDI_{95} = [0.01, 0.06], d = 0.51, HDI_{95} = [0.16, 0.87]), implying comparably more processing between alternatives in attraction trials.

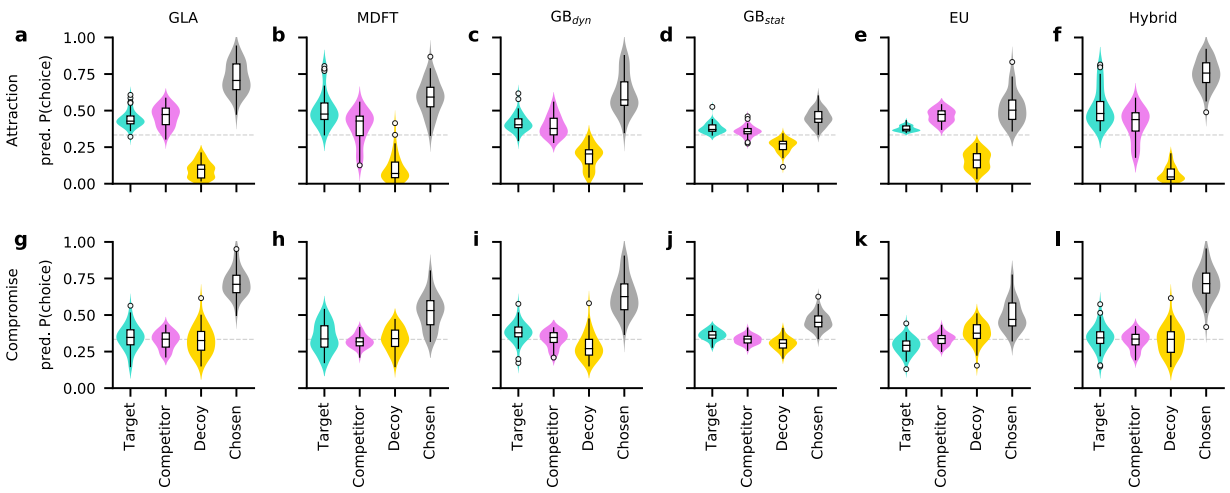
	mean	SD	min	25%	50%	75%	max
α	0.47	0.35	0.05	0.25	0.37	0.61	1.67
β	6.81	9.87	0.04	1.13	3.25	8.04	49.74
γ	0.81	0.25	0.22	0.58	1.0	1.0	1.0
λ	0.29	0.2	0.08	0.14	0.23	0.46	0.65
θ	0.69	0.18	0.13	0.63	0.72	0.83	0.95

Supplementary Table 1: **Summary of GLA estimates.** α is the utility parameter. β is the inverse temperature parameter of the choice rule (0 = random choice). γ is the probability weighting parameter (1 = objective probability weighting). λ is the leak parameter (0 = perfect memory, 1 = full leak of all previous information). θ is the gaze-discount parameter (1 = no gaze-discount, 0 = maximum gaze-discount).

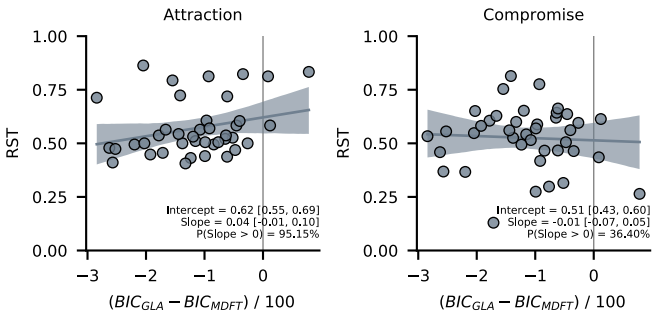


84

85 Supplementary Figure 3: **GLA maximum likelihood estimates and relationships between**
86 **parameters.** α is the utility parameter. β is the inverse temperature parameter of the choice rule
87 (0 = random choice). γ is the probability weighting parameter (1 = linear weighting). λ is the leak
88 parameter (0 = perfect memory, 1 = full leak of all previous information). θ is the gaze-discount
89 parameter (1 = no gaze-discount, 0 = maximum gaze-discount).



Supplementary Figure 4: **Model-predicted choice probabilities.** Each panel shows distributions of participant-level mean model-predicted choice probabilities for the target, competitor, decoy and ultimately chosen alternative. Predictions for attraction and compromise trials are displayed separately in the top **(a-f)** and bottom rows **(g-l)**. Predictions were computed using individual maximum likelihood estimates. The hybrid model **(f, l)** was derived from the switchboard analysis and combines an alternative-wise gaze-discount with a distance-dependent inhibition mechanism. 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.



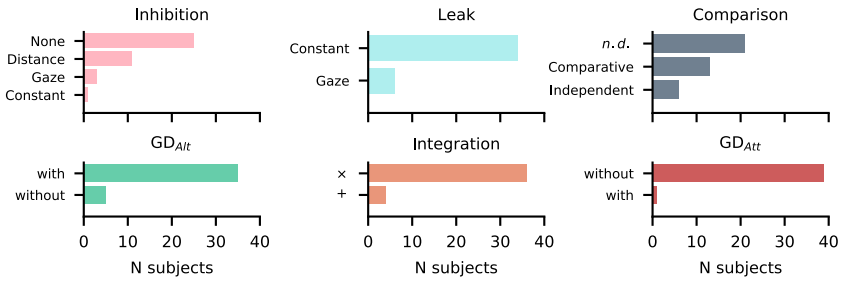
Supplementary Figure 5: **Relative model fits between GLA and MDFT in relation to RST**. Relative model fits of MDFT (indicated by BIC difference between GLA and MDFT) tended to be higher for participants with higher RST in attraction trials (left panel; slope = 0.04, HDI_{95} = [-0.01, 0.09] increase in RST per 100 unit increase in BIC difference, 93.6% of posterior mass above 0), but not compromise trials (right panel), even though 7 out of 9 participants with attraction RST above 0.7 were better described by GLA overall (participants left of dashed vertical line).

Switch	Gaze-independent levels			Gaze-dependent levels	N
Attribute integration	**× Multiplicative** Outcomes and probabilities combine multiplicatively into expected utilities[6,7].	**+ Weighted additive** Outcomes and probabilities are normalized, weighted and added[8].		*n.a.*	2
Comparison	**Independent** Accumulation of absolute item values per alternative[7,9].	**Comparative** Accumulation of relative values per alternative[10].		*n.a.*	2
Attribute-wise gaze-discount	**False** Fixated and non-fixated alternatives are processed equally.			**True** Non-fixated alternatives' values are discounted by a parameter θ[11].	2
Alternative-wise gaze-discount	**False** Fixated and non-fixated attributes are processed equally.			**True** Attributes on the non-fixated dimension are discounted by a parameter η[12,13].	2
Accumulation leak	**None** Perfect integration over fixations.	**Constant** With each fixation, accumulators leak information proportional to their current value, controlled by parameter λ[10,14].		**Gaze-dependent** With each fixation, accumulators of non-fixated alternatives leak information, as in Constant leak[15].	3
Accumulator inhibition	**None** No inhibition between accumulators.	**Constant** Accumulators inhibit each other proportional to their current value, controlled by parameter φ[16].	**Distance-dependent** Inhibition between accumulators depends on pairwise distance between alternatives in attribute space. Parameters w_d and φ[10].	**Gaze-dependent** With each fixation, accumulators of non-fixated alternatives are inhibited, proportional to the currently fixated items' accumulator value. Controlled by parameter φ.	4
Total variants					192

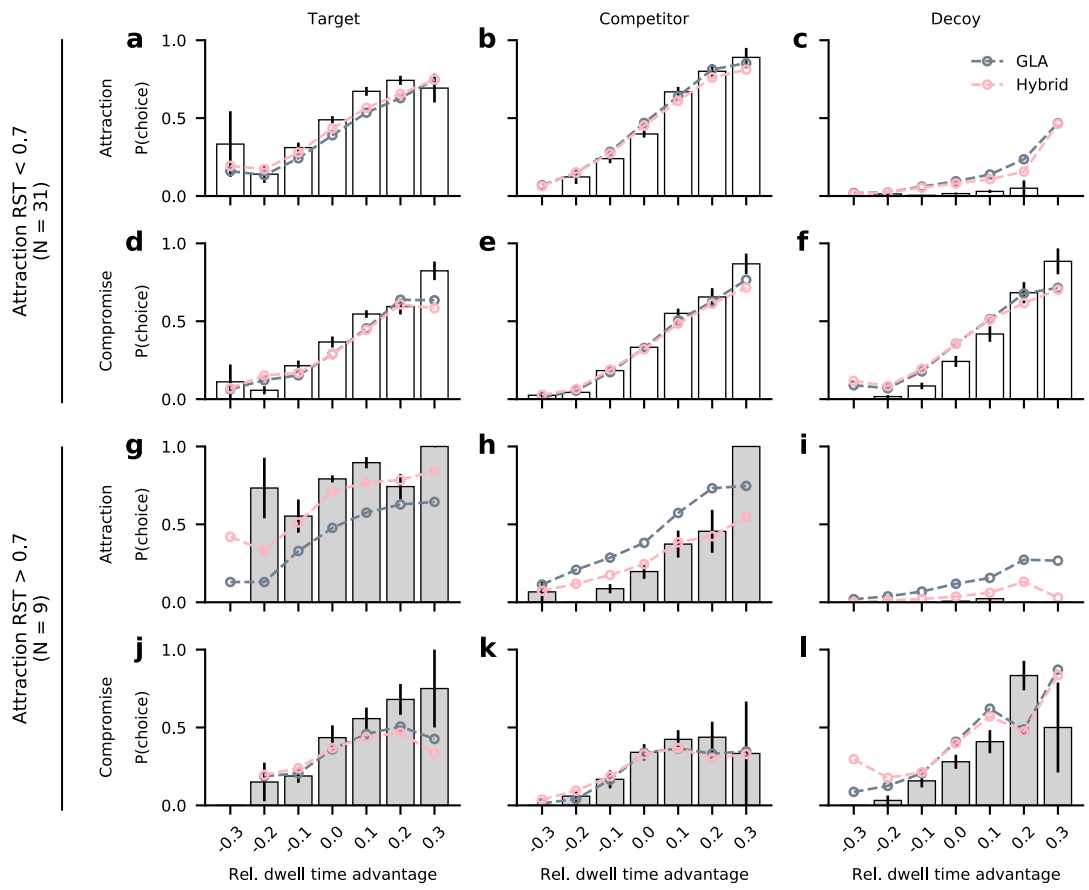
Supplementary Table 2. Overview of the nodes and switch-levels used in the switchboard analysis. Switch levels that depend on gaze-data are shaded blue. Note that due to the model fitting procedure, where model predicted choice probabilities are derived from a soft-max function over the final accumulator values, the comparison switch levels independent and comparative are not distinguishable for a subset of the model space, reducing the total number of unique variants to 160.

Rank	GD_{Alt}	GD_{Att}	Leak	Inhibition	Integration	Comparison	BIC
1	**Yes**	No	Constant	None	Multiplicative	*n.d.*	232.08
2	No	No	Constant	**Gaze**	Multiplicative	Independent	235.94
3	**Yes**	No	Constant	**Gaze**	Multiplicative	Comparative	236.75
4	**Yes**	**Yes**	Constant	None	Multiplicative	*n.d.*	237.21
5	**Yes**	No	Constant	**Gaze**	Multiplicative	Independent	237.31
6	**Yes**	No	Constant	Constant	Multiplicative	*n.d.*	237.40
7	**Yes**	No	**Gaze**	None	Multiplicative	Comparative	238.41
8	No	**Yes**	Constant	**Gaze**	Multiplicative	Independent	241.04
9	**Yes**	No	Constant	Distance	Multiplicative	Comparative	241.40
10	**Yes**	**Yes**	Constant	**Gaze**	Multiplicative	Comparative	241.78

Supplementary Table 3: **Overview of average best fitting model variants.** All ten model variants that fit the data best on average used some form of gaze-dependence (blue shaded cells), mostly an alternative-wise gaze discount. *"n.d."* denotes variants where comparison mechanisms were not distinguishable by the analysis.



Supplementary Figure 6: **Counts of individual best fitting switches.** Most participants were best described by model variants that included multiplicative attribute integration, with alternative-wise gaze discount, no attribute-wise gaze discount, constant leakage and no inhibition.



120

121 Supplementary Figure 7: **Observed and model-predicted association of dwell time advantage**
122 **and choice for participants with weaker and strong attraction effects. (a-f)** Data and model
123 predictions for participants with weaker attraction effects (RST $<$ 0.7). **(g-l)** Data and model
124 predictions for participants with strong attraction effects (RST $>$ 0.7) Each column refers to one
125 choice alternative: Target (first column; **a, d, g, j**); Competitor (second column; **b, e, h, k**);
126 Decoy (third column; **c, f, i, l**). Rows refer to trials in attraction **(a-c, g-i)** and compromise trials
127 **(d-f, j-l)**. White and grey bars and error bars show observed mean $\pm$ s.e. choice probabilities
128 computed from even-numbered trials, for participants with weaker and stronger attraction effects,
129 respectively. Coloured lines indicate model predictions derived from 50 simulations for each odd-
130 numbered trial.

Supplementary Note 2

No process evidence that strong attraction responders follow simple choice rule

Using process measures, we performed multiple tests of the hypothesis, that individuals with strong attraction effects follow a simple choice rule of choosing the dominant alternative. First, we tested whether the strength of individual attraction effects (individual RST in attraction trials) was related to differences in mean response times (RTs) in attraction trials. If individuals used a choice rule, their choices might be made faster, as they do not engage in multiple pairwise comparisons or calculations of expected outcomes. There was no correlation between the two measures (r = 0.06, HDI_{95} = [-0.24, 0.34]). Similarly, no relationship was found between individual RST and the number of fixations in attraction trials (r = 0.06, HDI_{95} = [-0.25, 0.35]). Mean RTs in attraction trials did not meaningfully differ between trials with target choices and trials with other choices (d = -0.2, HDI_{95} = [-0.67, 0.25]). Next, we tested whether individuals with strong attraction effects committed to a choice once they learned about the dominance relationship in the stimuli, as if using the dominance relationship as a stopping rule, or if they kept exploring the stimuli. There was, however, no relationship between individual RST and the mean number of fixations after all target and decoy attributes were fixated at least once (r = 0.04, HDI_{95} = [-0.28, 0.31]). The same analysis using fixation counts after target and decoy alternatives were both seen at least once on any attribute revealed no effect either (r = 0.11, HDI_{95} = [-0.20, 0.40]). Taken together, we did not find any evidence based on process data to support the hypothesis that strong attraction responders used a simple choice rule.

Supplementary Information References

1. Yarkoni, T. & Westfall, J. Bambi: A simple interface for fitting Bayesian mixed effects models. (2016) doi:10.31219/osf.io/rv7sn.

2. Westfall, J. Statistical details of the default priors in the Bambi library. *arXiv:1702.01201 [stat]* (2017).

3. Mullett, T. L. & Stewart, N. Implications of visual attention phenomena for models of preferential choice. *Decision* **3**, 231 (2016).

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

5. Payne, J. W. Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance* **16**, 366–387 (1976).

6. Tversky, A. & Kahneman, D. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* **5**, 297–323 (1992).

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

8. Rouault, M., Drugowitsch, J. & Koechlin, E. Prefrontal mechanisms combining rewards and beliefs in human decision-making. *Nature Communications* **10**, 1–16 (2019).

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

10. Roe, R. M., Busemeyer, J. R. & Townsend, J. T. Multialternative decision field theory: A dynamic connectionst model of decision making. *Psychological Review* **108**, 370–392 (2001).

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

12. Krajbich, I., Lu, D., Camerer, C. & Rangel, A. The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology* **3**, 193 (2012).

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

14. Usher, M. & McClelland, J. L. The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review* **108**, 550–592 (2001).

15. Ashby, N. J. S., Jekel, M., Dickert, S. & Glöckner, A. Finding the right fit: A comparison of process assumptions underlying popular drift-diffusion models. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **42**, 1982–1993 (2016).

16. Usher, M. & McClelland, J. L. Loss Aversion and Inhibition in Dynamical Models of Multialternative Choice. *Psychological Review* **111**, 757–769 (2004).