Supplementary Material for Attention biases preferential choice by enhancing stimulus value

Timothy J. Pleskac[1,2], Shuli Yu[2], Sergej Grunevski[1] and Taosheng Liu[3]

[1]University of Kansas
[2]Judge.me
[3]Michigan State University

September 9, 2022

Contents

1 Study 1

1.1 Psychometric function

1.1.1 Parameter estimates

Table S1: Mean and 95% HDI posterior estimates of the group-level threshold (μ), slope (θ) and sensitivity parameters of the psychometric function fit to Study 1.

	Condition	Preferential	Perceptual	Pref. vs. Perc.
Threshold (/*mu*)	Left	4.16 [0.54, 7.74]	3.45 [-0.2, 6.71]	0.71 [-4.27, 5.7]
	Neutral	-1.02 [-4.48, 2.69]	-1.26 [-4.86, 2.07]	0.25 [-4.84, 5.07]
	Right	-6.77 [-10.3, -3.22]	-4.51 [-8.05, -1.11]	-2.26 [-7.48, 2.43]
	L v N	5.18 [3.9, 6.43]	4.71 [3.63, 5.87]	0.47 [-1.22, 2.19]
	L v R	10.93 [9.67, 12.24]	7.96 [6.88, 9.12]	2.97 [1.17, 4.58]
	N v R	5.75 [4.47, 6.99]	3.25 [2.13, 4.33]	2.51 [0.73, 4.09]
Slope /*theta*	Left	17.91 [14.66, 21.13]	12.05 [9.61, 14.56]	5.85 [1.68, 9.85]
	Neutral	16.83 [13.81, 20.03]	12.2 [9.77, 14.76]	4.63 [0.64, 8.64]
	Right	17.44 [14.22, 20.59]	12.07 [9.6, 14.53]	5.36 [1.23, 9.31]
	L v N	1.08 [-0.08, 2.26]	-0.15 [-0.96, 0.67]	1.23 [-0.17, 2.68]
	L v R	0.47 [-0.74, 1.74]	-0.02 [-0.87, 0.83]	0.49 [-1.04, 1.95]
	N v R	-0.61 [-1.74, 0.53]	0.13 [-0.68, 0.96]	-0.73 [-2.08, 0.74]
Sensitivity	Left	0.0139 [0.0115, 0.0165]	0.0206 [0.0166, 0.025]	-0.0066 [-0.0116, -0.0018]
	Neutral	0.0148 [0.0122, 0.0176]	0.0203 [0.0163, 0.0245]	-0.0055 [-0.0107, -9e-04]
	Right	0.0143 [0.0117, 0.0169]	0.0205 [0.0166, 0.0249]	-0.0062 [-0.0112, -0.0014]
	L v N	-9e-04 [-0.0019, 1e-04]	3e-04 [-0.0011, 0.0017]	-0.0011 [-0.0029, 6e-04]
	L v R	-4e-04 [-0.0014, 6e-04]	0 [-0.0014, 0.0016]	-4e-04 [-0.0021, 0.0014]
	N v R	5e-04 [-5e-04, 0.0015]	-2e-04 [-0.0016, 0.0012]	7e-04 [-0.001, 0.0025]
Lower Asymptote Increment /*gamma*		.0046 [.0001, .0161]	0.0069 [.0001, .0153]	0.0069 [.0001, .0152]
Upper Asymptote Decrement /*lambda*		0.0054 [.0001, .0131]	.0121 [.0019, .0235]	-0.0066 [-0.0207, 0.0067]

A 4 parameter logistic function was used a psychometric function. The threshold parameter corresponds to the point of subjective equality (the halfway point between the lower and upper asymptotes of the psychometric function) in terms of the relative value d. The parameter (θ) is a slope parameter that determines how the probability of choosing the right option changes with the relative value. We allowed both parameters to vary between the decision frame and cue conditions. To estimate sensitivity or the ability to discriminate one option from another in terms of the relative value, we estimated the psychometric function's slope at the threshold. The model was estimated as a Bayesian hierarchical model.

1.1.2 Plot with subject level variability

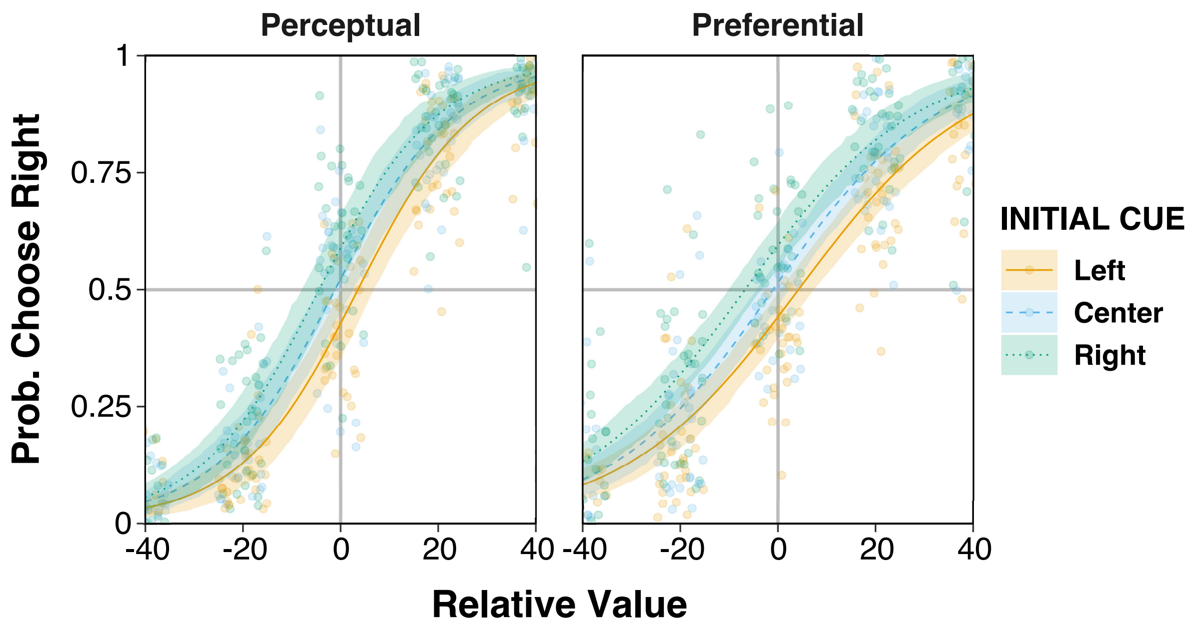

Figure S1: The probability of choosing the right option is plotted against the relative value (difference in mean dots between options, right-left). The data was conditioned by the decision frame (panels) and cue location (colors). The lines represent the posterior predicted choice proportions with the error regions indicating the 95% HDIs and the pale points showing individual participant data (jittered for better visibility) .

1.2 Response times

1.2.1 Regression coefficients

	Coefficient	**95% HDI**
Intercept	-0.11	[-0.36, 0.14]
Task frame	**0.37**	**[0.02, 0.72]**
Right cue	**-0.08**	**[-0.13,-0.04]**
Center cue	**-0.09**	**[-0.14,-0.05]**
Choice	**-0.06**	**[-0.11,-0.01]**
Relative value	**0**	**[-0.03,0.02]**
Relative value^2	**18.00**	**[16.27,19.72]**
Task frame x Right cue	**-0.14**	**[-0.20,-0.07]**
Task frame x Center cue	-0.05	[-0.11,0.01]
Task frame x Relative value	0.04	[0,0.07]
Task frame x Choice	-0.04	[-0.11,0.02]
Right cue x Relative value	0.03	[-0.01,0.07]
Center cue x Relative value	0	[-0.04,0.04]
Right cue x Choice	**0.23**	**[0.16,0.29]**
Center cue x Choice	**0.14**	**[0.07,0.21]**
Relative value x Choice	-0.01	[-0.05,0.03]
Task frame x Right cue x Relative value	-0.02	[-0.07,0.04]
Task frame x Center cue x Relative value	-0.01	[-0.06,0.05]
Task frame x Right cue x Choice	**0.13**	**[0.03,0.22]**
Task frame x Center cue x Choice	0.01	[-0.08,0.10]
Task frame x Relative value x Choice	**-0.13**	**[-0.19,-0.08]**
Right cue x Relative value x Choice	-0.03	[-0.08,0.03]
Center cue x Relative value x Choice	0	[-0.06,0.06]
Task frame x Right cue x Relative value x Choice	**0.08**	**[0,0.16]**
Task frame x Center cue x Relative value x Choice	0.05	[-0.02,0.13]

Table S2: Parameter means and 95% HDIs for the response time regression parameters for Study 1. Estimates from a hierarchical Bayesian linear regression model with random intercepts fit with *stan* (Goodrich et al., 2020). Decision frame, cue, and choice were dummy coded against the Perceptual frame, Left cue, and Left choice, respectively.

1.2.2 Plot with subject level variability

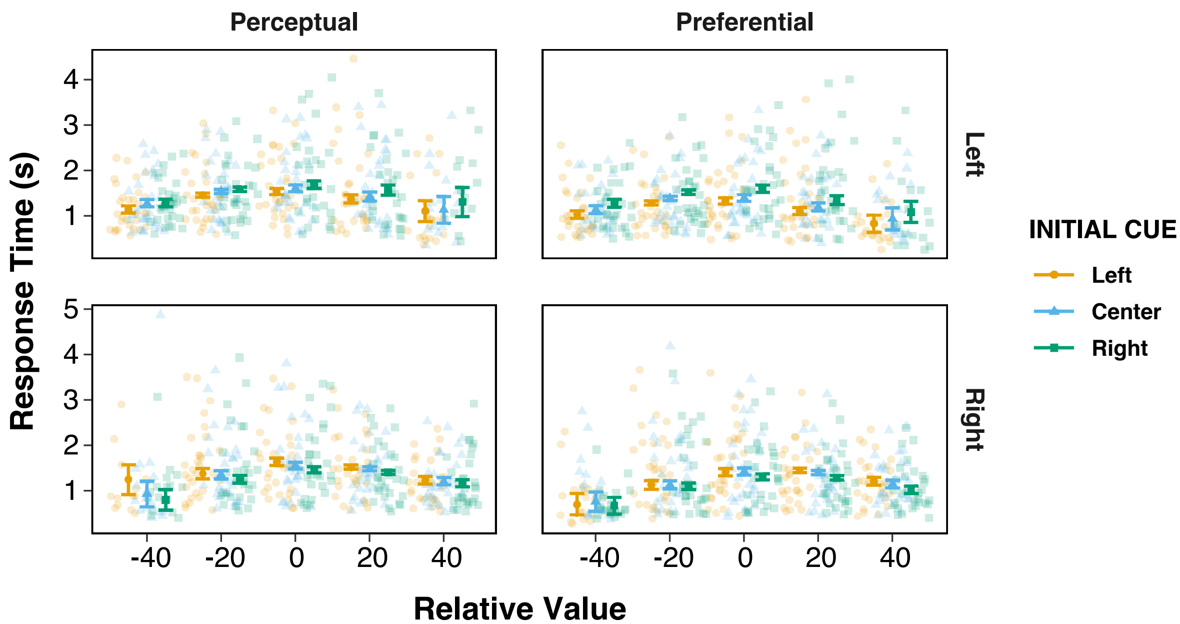

Figure S2: The posterior predicted means of response times are plotted against the relative value (difference in mean dots between options, right-left). The plot points are conditioned by the decision frame (panels) and cue location (colors). Error bars indicate 95% HDIs of the posterior predicted means and pale points show individual participant data (jittered for better visibility).

1.3 Sensitivity and Response Time Differences

1.3.1 Individual Level Analyses

As reported in the main paper, there were credible differences between decision frames regarding sensitivity and response times. To better understand these differences we examined sensitivity and response times at the individual level. Figure S3 plots the posterior mean estimates of sensitivity for each individual from the psychometric function (fit as a hierarchical model) by the mean response times. The plot shows the common positive relationship between response times and sensitivity. The plot further shows a cluster of participants in the preferential frame with distinctly low sensitivities and fast response times.

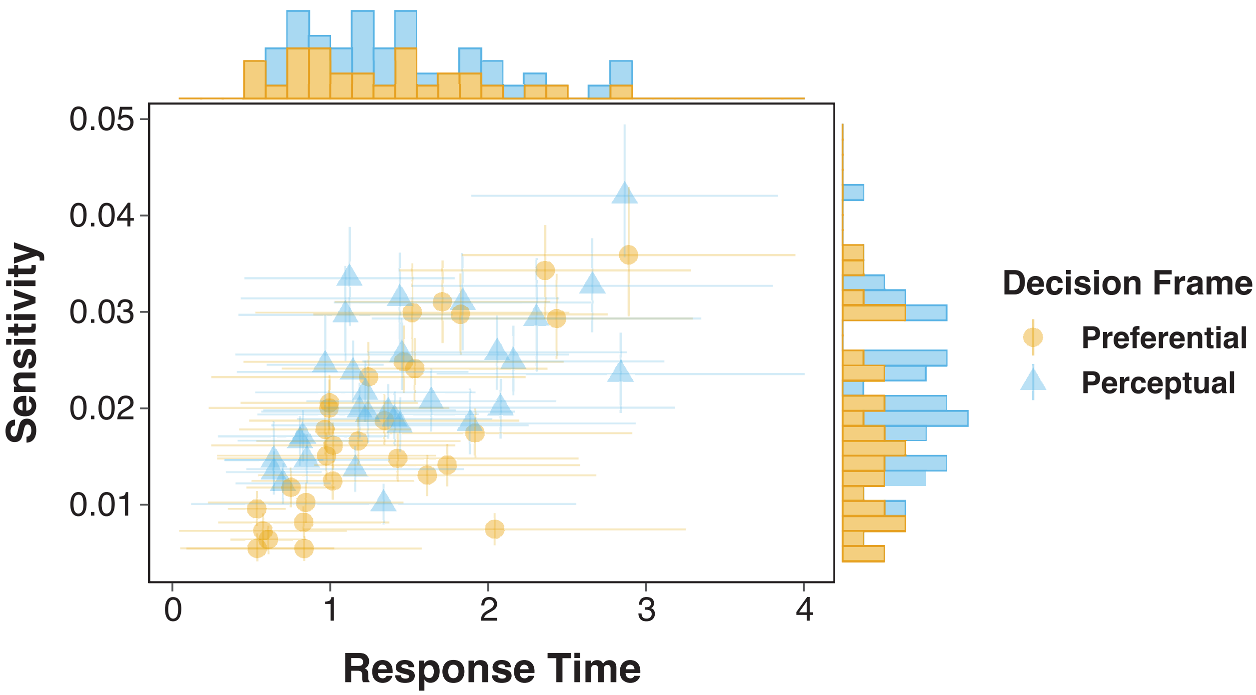

Figure S3: The mean response time for each individual (error bars are ±1 standard deviation) is plotted against the posterior mean sensitivity (error bars are the 95% HDI) across the cue conditions (Left, Center, or Right). The data was conditioned by the decision frame (colors and shapes). The histograms depict marginal distributions of response time and sensitivity for each decision frame.

Table S3 lists the mean posterior sensitivity, mean response times, and the mean posterior estimates of the difference in the threshold for all the participants. The table is sorted in terms of sensitivity. We have bolded the seven participants with the lowest sensitivity estimates, all of which were from the preferential frame. These were the participants in the cluster in Figure S3. The response times also show that they also tended to record the faster responses (participants 4, 18, 15, and 26 were the fastest responders). Recall that the average response times were $M = 1.43s$ $[1.21, 1.65]$ in the preferential frame and $M = 1.61s$ $[1.39, 1.84]$ in the perceptual frame. Examining the mean posterior estimates of the difference in threshold between the left and right cue conditions for the seven participants with the lowest sensitivity reveals that they also tended to exhibit a similar level or even smaller cueing effect than the overall group. Recall the overall shift in threshold for the preferential frame was $M = 10.21$ $[8.88, 11.52]$ and for the perceptual frame was $M = 8.17$ $[6.8, 9.37]$. This speaks against the hypothesis that participants with the lowest sensitivity and hence greater uncertainty would show a larger cue effect because any additional information could, in principle, have a greater effect on choice. We explore this issue further with analyses below that equate the preferential and perceptual conditions regarding sensitivity and response times.

Participant	Task	Sensitivity [95% HDI]	Response Time (SD)	Diff. in Threshold [95% HDI]
13	**Preferential**	**0.0055 [0.0042,0.0068]**	**0.83 (0.74)**	**10.36 [8.97,11.74]**
18	**Preferential**	**0.0055 [0.0041,0.007]**	**0.54 (0.49)**	**8.2 [6.11,10.19]**
26	**Preferential**	**0.0064 [0.0049,0.0078]**	**0.61 (0.24)**	**10.91 [9.67,12.24]**
15	**Preferential**	**0.0073 [0.0056,0.0091]**	**0.57 (0.53)**	**9.17 [7.65,10.78]**
3	**Preferential**	**0.0075 [0.0058,0.0091]**	**2.04 (1.21)**	**10.91 [9.61,12.18]**
19	**Preferential**	**0.0082 [0.0065,0.0099]**	**0.83 (0.54)**	**10.85 [9.56,12.14]**
4	**Preferential**	**0.0096 [0.0078,0.0114]**	**0.54 (0.18)**	**8.94 [7.45,10.34]**
51	Perceptual	0.0101 [0.008,0.0122]	1.34 (1.22)	7.79 [6.63,8.86]
21	Preferential	0.0102 [0.0081,0.0123]	0.84 (0.62)	10.74 [9.48,12.04]
7	Preferential	0.0118 [0.0097,0.0139]	0.75 (0.28)	10.93 [9.65,12.23]
58	Perceptual	0.0122 [0.0101,0.0143]	0.7 (0.3)	7.71 [6.59,8.79]
29	Preferential	0.0125 [0.0105,0.0146]	1.01 (0.51)	10.89 [9.54,12.11]
25	Preferential	0.0131 [0.0109,0.0152]	1.61 (1.07)	10.87 [9.62,12.17]
46	Perceptual	0.0134 [0.0111,0.0155]	0.64 (0.3)	7.67 [6.57,8.79]
39	Perceptual	0.0137 [0.0113,0.0161]	1.16 (0.69)	7.96 [6.87,9.12]
9	Preferential	0.0141 [0.0119,0.0163]	1.74 (0.84)	10.86 [9.6,12.16]
32	Perceptual	0.0147 [0.0114,0.0187]	0.64 (0.31)	6.63 [5.61,7.66]
34	Perceptual	0.0147 [0.0122,0.0173]	0.85 (0.52)	7.87 [6.73,8.97]
20	Preferential	0.0148 [0.0124,0.0174]	1.43 (1.14)	10.93 [9.64,12.22]
1	Preferential	0.0151 [0.0126,0.0175]	0.98 (0.69)	10.6 [9.27,11.8]
2	Preferential	0.0162 [0.0136,0.0186]	1.02 (0.77)	10.89 [9.65,12.22]
10	Preferential	0.0166 [0.0142,0.0191]	1.18 (0.64)	10.87 [9.62,12.19]
54	Perceptual	0.0167 [0.0141,0.0192]	0.81 (0.39)	7.86 [6.77,9]
59	Perceptual	0.0171 [0.0145,0.0198]	0.83 (0.54)	7.65 [6.58,8.77]
24	Preferential	0.0174 [0.0149,0.02]	1.92 (0.99)	10.87 [9.62,12.19]
28	Preferential	0.0178 [0.0152,0.0203]	0.97 (0.54)	10.91 [9.58,12.15]
61	Perceptual	0.0182 [0.0155,0.0212]	1.44 (0.81)	7.91 [6.79,9.03]
47	Perceptual	0.0184 [0.0152,0.022]	1.88 (1.05)	7.93 [6.84,9.08]
17	Preferential	0.0187 [0.0162,0.0215]	1.34 (0.85)	10.88 [9.55,12.12]
57	Perceptual	0.0189 [0.0161,0.0217]	1.41 (0.62)	7.93 [6.83,9.07]
33	Perceptual	0.0194 [0.0164,0.0223]	1.24 (0.7)	7.91 [6.82,9.05]
50	Perceptual	0.0197 [0.0169,0.0225]	1.37 (0.8)	7.95 [6.86,9.11]
36	Perceptual	0.0198 [0.0169,0.0227]	1.19 (0.61)	7.87 [6.75,8.99]
12	Preferential	0.02 [0.0172,0.023]	0.99 (0.76)	10.79 [9.48,12.02]
55	Perceptual	0.02 [0.0169,0.0231]	2.08 (1.1)	7.87 [6.78,9.01]
27	Preferential	0.0206 [0.0178,0.0234]	0.99 (0.56)	10.92 [9.6,12.18]
52	Perceptual	0.0207 [0.0176,0.0241]	1.64 (0.79)	7.69 [6.65,8.82]
48	Perceptual	0.0217 [0.0186,0.025]	1.22 (0.69)	7.83 [6.72,8.93]
22	Preferential	0.0233 [0.0198,0.0269]	1.24 (0.99)	10.75 [9.42,11.95]
35	Perceptual	0.0236 [0.0195,0.0278]	2.84 (1.16)	7.84 [6.71,8.93]
44	Perceptual	0.0238 [0.0206,0.027]	1.14 (0.73)	7.95 [6.85,9.09]
5	Preferential	0.0241 [0.0207,0.0273]	1.53 (0.84)	10.94 [9.63,12.21]
49	Perceptual	0.0245 [0.02,0.0297]	0.97 (0.37)	7.59 [6.57,8.74]
16	Preferential	0.0248 [0.0216,0.0286]	1.46 (1.01)	10.86 [9.55,12.11]
60	Perceptual	0.0249 [0.0213,0.0285]	2.16 (0.95)	7.89 [6.82,9.05]
56	Perceptual	0.0256 [0.0219,0.0294]	1.45 (1.05)	7.95 [6.87,9.12]
40	Perceptual	0.0258 [0.0219,0.0296]	2.06 (0.82)	7.65 [6.59,8.77]
23	Preferential	0.0293 [0.0251,0.034]	2.43 (0.86)	10.93 [9.69,12.26]
42	Perceptual	0.0293 [0.0238,0.0356]	2.31 (1.04)	7.91 [6.83,9.06]
38	Perceptual	0.0297 [0.0249,0.0348]	1.1 (0.68)	7.92 [6.83,9.07]
8	Preferential	0.0298 [0.0256,0.034]	1.82 (0.93)	10.92 [9.65,12.22]
14	Preferential	0.0299 [0.0251,0.035]	1.52 (0.99)	10.94 [9.68,12.26]
53	Perceptual	0.031 [0.026,0.0361]	1.84 (0.81)	7.95 [6.86,9.1]
30	Preferential	0.0311 [0.0268,0.0353]	1.71 (0.68)	10.94 [9.68,12.26]
41	Perceptual	0.0314 [0.0267,0.0362]	1.44 (1)	7.93 [6.84,9.07]
37	Perceptual	0.0327 [0.0279,0.0376]	2.66 (1.14)	7.96 [6.87,9.12]
43	Perceptual	0.0335 [0.0285,0.0388]	1.12 (0.66)	7.96 [6.86,9.11]
6	Preferential	0.0343 [0.0293,0.039]	2.36 (0.92)	10.93 [9.69,12.26]
31	Preferential	0.0359 [0.0295,0.043]	2.89 (1.05)	10.93 [9.69,12.27]
45	Perceptual	0.042 [0.0357,0.0494]	2.86 (0.97)	7.93 [6.83,9.08]

Table S3: Individual-level posterior means and 95% HDIs for sensitivity and difference in threshold between left and right cues, and mean and standard deviation of response times. Data are sorted in ascending order of sensitivity; the seven participants who had the lowest sensitivity are shown in a bold font.

1.3.2 Matching Analyses

Matching decision frames in terms of sensitivity and threshold As reported in the main paper, we found that the preferential frame had credibly lower sensitivity and response times. This difference suggests there may be some information processing differences between these two decision frames, which may explain the differential cue effect on the threshold. For instance, perhaps there was more uncertainty in the preferential frame (thus lower sensitivity) or there was more uncertainty due to the faster responses, making participants more susceptible to the cue. The individual-level estimates reported above suggest that these differences between decision frames were driven by a few individuals. Nevertheless, we carried out an additional set of analyses to match the preferential and perceptual conditions in terms of sensitivity and response time. To do so, we isolated the neutral condition (i.e., the cue was in the center) for each participant in each decision frame, and fit the psychometric functions to estimate a threshold and sensitivity for each participant in the neutral condition.

We then created a group of participants in the preferential frame and a group of participants in the perceptual frame that had matched levels of sensitivity and thresholds. To do so, we treated the standardized sensitivity and threshold estimates as coordinates in a multidimensional space. Then for each participant in the perceptual frame, we found their nearest neighbor in terms of Euclidean distance. If that distance was within one standard deviation, then the perceptual participant and the nearest preferential neighbor were placed into the respective conditions. We repeated this procedure for each perceptual participant, which ultimately created 20 pairs of matched participants from the two decision frames. We then repeated our analyses with these two matched groups fitting the psychometric function hierarchically across all the conditions with Bayesian estimation techniques. Doing so replicated the results we reported with the full set of participants. There was a credible shift in threshold for the preferential ($M = 10.21$ $[8.88, 11.52]$) and perceptual frames ($M = 8.17$ $[6.8, 9.37]$). Moreover, consistent with the AIV hypothesis, the cueing effect in the threshold was credibly larger in the preferential frame than in the perceptual frame ($M = 2.05$ $[0.26, 3.84]$). Crucially, because the conditions were matched in terms of sensitivity, there was no credible difference in sensitivity between decision frames ($M = -0.0026$ $[-0.0078, 0.0026]$).

Matching decision frames in terms of sensitivity, threshold, and response times We repeated the matching analyses creating a group of participants in the preferential frame and a group of participants in the perceptual frame that matched sensitivity levels, thresholds, and response times. In this case, we used the individual estimates of threshold and sensitivity from the neutral conditions and included the mean response times. We treated the three standardized values of these statistics as coordinates in a three-dimensional space and repeated our procedure of selecting a participant in the perceptual frame and identifying their nearest neighbor in the preferential frame. If that distance was within one standard deviation, then the perceptual participant and the nearest preferential neighbor were placed into the respective conditions. We repeated this for each perceptual participant, which ultimately created a 18 pairs of matched participants. We then repeated our analyses with these two matched groups fitting the psychometric function hierarchically across all the conditions with Bayesian estimation techniques. We again replicated the results we reported with the full set of participants. There was a a credible shift in threshold for the preferential ($M = 10.84$ $[9.42, 12.5]$) and perceptual frames ($M = 8.48$ $[7.04, 9.84]$). Moreover, consistent with the AIV hypothesis, the cueing effect in the threshold was credibly larger in the preferential frame than in the perceptual frame ($M = 2.36$ $[0.28, 4.4]$). Crucially, because the conditions were matched in terms of sensitivity and response time, there was neither a credible difference in sensitivity ($M = -0.0040$ $[-0.0101, 0.0024]$), nor a credible difference in response times ($M = -0.12s[-0.42, 0.14]$) between decision frames.

Summary In sum, our additional analyses help rule out the hypothesis that the cueing effect in Study 1 is due to different levels of information processing between the decision frames where the cue more impacts participants with lower sensitivity levels. Overall, the two decision frames are fairly well matched in terms of sensitivity and response times. The credible group-level differences we observed were largely driven by a few extreme individuals with low sensitivity and fast responses in the preferential frame. Those same individuals tended to exhibit a similar or smaller cueing effect. Moreover, the differential cueing effect between decision frames remains when we further subsampled participants such that the two frames were matched in terms of sensitivity and response times.

1.4 Gaze variability

1.4.1 Regression coefficients

	Coefficient	95% HDI
Intercept	**4.17**	**[3.74,4.61]**
Task frame	-0.33	[-0.95,0.29]
Center cue	-0.07	[-0.16,0.03]
Right cue	**-0.16**	**[-0.26,-0.07]**
Unsigned relative value	**-0.21**	**[-0.27,-0.16]**
Task frame x Center cue	-0.06	[-0.20,0.07]
Task frame x Right cue	0.10	[-0.04,0.23]
Task frame x Unsigned relative value	**0.08**	**[0.002,0.17]**
Center cue x Unsigned relative value	-0.01	[-0.09,0.07]
Right cue x Unsigned relative value	**0.10**	**[0.02,0.18]**
Task frame x Center cue x Unsigned relative value	0.08	[-0.04,0.19]
Task frame x Right cue x Unsigned relative value	-0.10	[-0.22,0.02]

Table S4: Parameter means and 95% HDIs for the gaze deviation regression parameters for Study 1. Estimates from a hierarchical Bayesian linear regression model with random intercepts fit with *stan* (Goodrich et al., 2020). Decision frame and cue were dummy coded against the Perceptual frame and Left cue respectively. Credible effects in bold face.

1.4.2 Plot with subject level variability

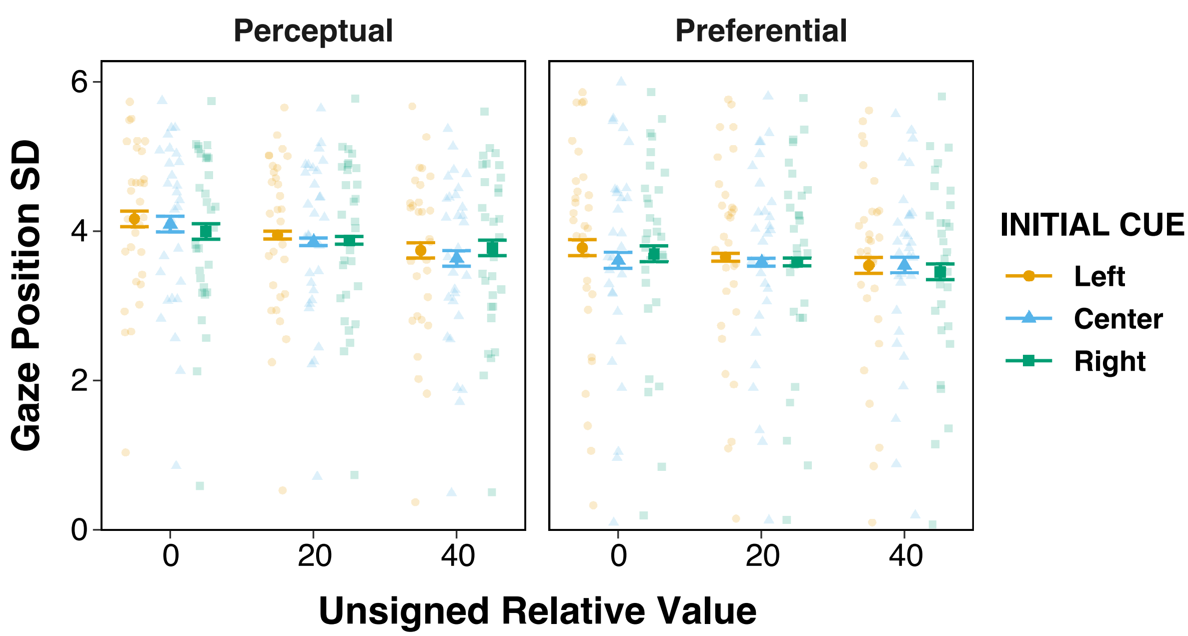

Figure S4: The posterior predicted means of gaze position standard deviation are plotted against the unsigned relative value (absolute difference in mean dots between options, right-left). The plot points are conditioned by the decision frame (panels) and initial cue location (colors). Error bars indicate 95% HDIs of the posterior predicted means and pale points show individual participant data (jittered for better visibility).

1.5 Gaze dwell time

1.5.1 Regression coefficients

	Coefficient	95% HDI
Intercept	-0.10	[-0.78,0.59]
Task frame	-0.50	[-1.48,0.47]
Center cue	0.30	[-0.01,0.60]
Right cue	**0.48**	**[0.17,0.79]**
Unsigned relative value	-0.05	[-0.25,0.13]
Task frame x Center cue	0.17	[-0.28,0.63]
Task frame x Right cue	**0.76**	**[0.31,1.21]**
Task frame x Unsigned relative value	0.18	[-0.09,0.46]
Center cue x Unsigned relative value	-0.08	[-0.35,0.19]
Right cue x Unsigned relative value	0.11	[-0.15,0.38]
Task frame x Center cue x Unsigned relative value	0.03	[-0.36,0.42]
Task frame x Right cue x Unsigned relative value	-0.28	[-0.68,0.10]

Table S5: Parameter means and 95% HDIs for the gaze dwell time regression parameters for Study 1. Estimates from a hierarchical Bayesian linear regression model with random intercepts fit with *stan* (Goodrich et al., 2020). Decision frame and cue were dummy coded against the Perceptual frame and Left cue respectively. Credible effects in bold face.

1.5.2 Plot with subject level variability

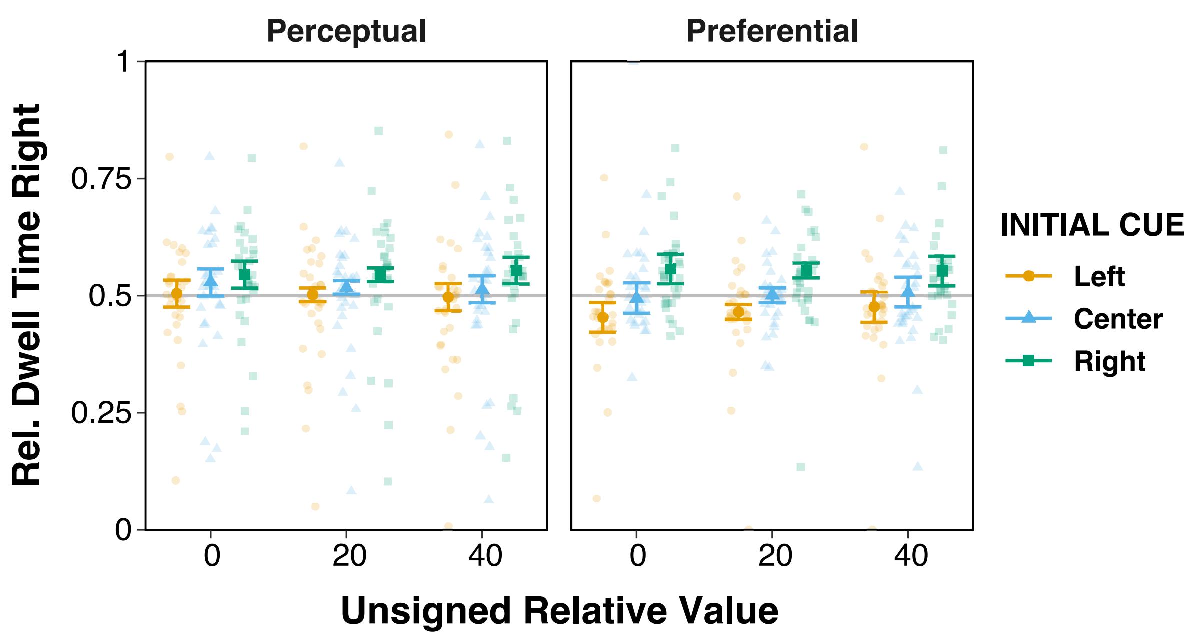

Figure S5: The posterior predicted means of relative dwell time for the right option are plotted against the unsigned relative value (absolute difference in mean dots between options, right-left). The plot points are conditioned by the decision frame (panels) and initial cue location (color). Error bars indicate 95% HDIs of the posterior predicted means and pale points are individual participant data (jittered for better visibility).

1.6 Gaze cascade regression table

	Coefficient	95% HDI
Intercept	**0.80**	**[0.53,1.06]**
Task frame	**0.43**	**[0.06,0.81]**
Right cue	**0.07**	**[0.02,0.12]**
Center cue	-0.01	[-0.06,0.04]
Unsigned relative value	**0.06**	**[0.03,0.09]**
Task frame x Right cue	**-0.16**	**[-0.23,-0.08]**
Task frame x Center cue	0.05	[-0.02,0.13
Task frame x Unsigned relative value	**-0.11**	**[-0.16,-0.07]**
Right cue x Unsigned relative value	-0.03	[-0.07,0.02]
Center cue x Unsigned relative value	0.03	[-0.01,0.08]
Task frame x Right cue x Unsigned relative value	**0.10**	**[0.03,0.16]**
Task frame x Center cue x Unsigned relative value	-0.02	[-0.09,0.04]

Table S6: Parameter means and 95% HDIs for the gaze cascade regression parameters for Study 1. Estimates from a hierarchical Bayesian logistic regression model with random intercepts fit with *stan* (Goodrich et al., 2020). Decision frame and cue were dummy coded against the Perceptual frame and Left cue, respectively.

1.6.1 Probability of fixating on chosen option averaged across the last 250 ms

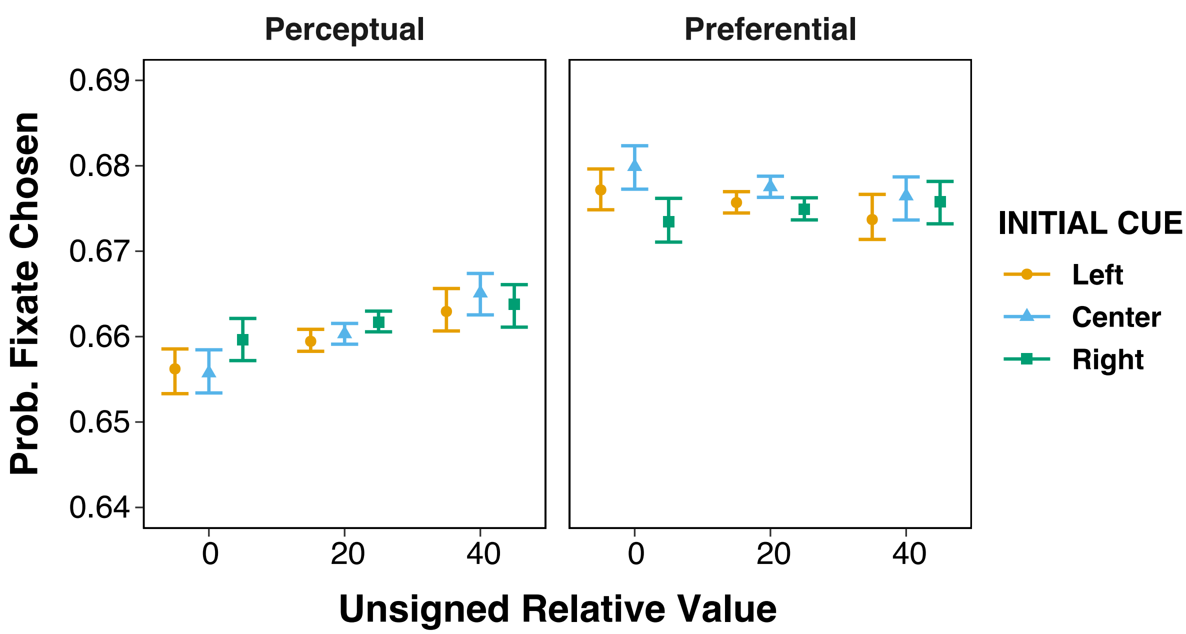

Figure S6: The posterior predicted means of the probability of fixating on the chosen option are plotted against the unsigned relative value. The plot points are conditioned by the decision frame (panels) and cue location (colors). Error bars indicate 95% HDIs of the posterior predicted means.

2 Study 2

2.1 Design of the stimuli

The study had a 2 (frame) × 5 (stimulus duration) × 5 (mean value difference) mixed design. The frame (preferential vs. perceptual) varied between participants, whereas the other two factors varied within participants. The relative value was the difference in the mean number of dots between options (right minus the left option: -50, -25, 0, 25, 50), made up from six combinations of option pairs. The stimulus duration was manipulated such that the first option was visible 33%, 50%, 67%, or 75%, over eight different presentation sequences of the two options outlined in Table S7.

These eight presentation sequences enabled us to control for several other factors during the trial. First, we systematically varied the number of switches between options so that half the trials had 1 switch and the other half had 2 switches. Second, we varied the number of updates of the flash stimulus (referred to as updates below and in the table) that the option was visible during each presentation so that during a presentation an option was shown for either 8 or 16 updates. These two factors meant that the total number of updates in each trial was also manipulated (16, 24 or 32 updates).

Table S7: Stimulus sequence design in the second study

	Number of Updates				
Number of Switches	Total	Set 1	Set 2	Set 3	Relative Duration of First Option
1	16	8	8	0	50%
1	24	8	16	0	33%
1	24	16	8	0	67%
1	32	16	16	0	50%
2	24	8	8	8	67%
2	32	8	8	16	75%
2	32	8	16	8	50%
2	32	16	8	8	75%

There were eight possible sequences, with different patterns of switches between the two options. The first four sequences had one switch and the last four sequences had two switches. Column "Total" shows the total duration of each stimulus sequence; columns "Set 1/2/3" shows the duration of each flash stimulus. The unit is in updates of the flash stimulus (20 Hz, 50 ms per stimulus update) . The last column shows the percentage of time Option 1 was shown.

2.2 Psychometric function

2.2.1 Parameter estimates

Table S8: Mean and 95% HDI posterior estimates of the group-level threshold, slope, sensitivity, and asymptote parameters of the psychometric function fit to Study 2.

	Condition	Preferential	Perceptual	Pref. vs Perc.
Threshold (μ)	33%	28.57 [24.62, 32.73]	24.11 [20.19, 27.95]	4.47 [-1.01, 9.97]
	50%	11.94 [7.95, 15.96]	10 [6.15, 14.16]	1.95 [-3.7, 7.57]
	67%	-3.6 [-7.71, 0.95]	-3.34 [-7.47, 1.13]	-0.26 [-6.34, 5.77]
	75%	-10.73 [-14.9, -6.65]	-8.78 [-13.07, -4.77]	-1.95 [-7.74, 3.87]
	33 vs 50%	16.63 [13.53, 19.67]	14.11 [11.67, 16.72]	2.52 [-1.55, 6.37]
	33 vs 67%	32.17 [28.83, 35.33]	27.44 [24.69, 30.3]	4.72 [0.34, 9.06]
	33 vs 75%	39.31 [36.17, 42.67]	32.89 [30.06, 35.46]	6.42 [2.24, 10.68]
	50 vs 67%	15.54 [13.4, 17.58]	13.33 [11.38, 15.2]	2.2 [-0.68, 5.04]
	50 vs 67%	22.68 [20.64, 24.74]	18.78 [16.85, 20.6]	3.9 [1.15, 6.73]
	67 vs 75%	7.14 [4.91, 9.55]	5.45 [3.43, 7.54]	1.69 [-1.34, 4.93]
Slope (θ)	33%	31.92 [24.18, 40.24]	26.72 [19.82, 34.13]	5.19 [-5.45, 16.15]
	50%	30.1 [22.79, 37.32]	26.54 [20.27, 33.77]	3.56 [-6.46, 13.63]
	67%	32.62 [25.42, 40.56]	28.84 [22.11, 36.27]	3.78 [-6.6, 14.49]
	75%	31.57 [23.94, 38.91]	27.7 [20.84, 34.83]	3.87 [-6.8, 13.92]
	33 vs 50%	1.82 [-1.35, 4.94]	0.18 [-2.36, 2.74]	1.63 [-2.69, 5.43]
	33 vs 67%	-0.7 [-3.91, 2.95]	-2.12 [-4.8, 0.79]	1.42 [-2.88, 5.97]
	33 vs 75%	0.34 [-3.05, 3.78]	-0.98 [-3.75, 1.94]	1.32 [-3.18, 5.67]
	50 vs 67%	-2.52 [-4.67, -0.33]	-2.3 [-4.24, -0.48]	-0.22 [-3.12, 2.52]
	50 vs 67%	-1.48 [-3.73, 0.57]	-1.16 [-3.04, 0.73]	-0.31 [-3.07, 2.63]
	67 vs 75%	1.04 [-1.39, 3.34]	1.14 [-0.84, 3.21]	-0.1 [-3.2, 3.08]
Sensitivity	33%	0.0077 [0.0059, 0.0096]	0.0089 [0.0066, 0.0112]	-0.0012 [-0.0044, 0.0017]
	50%	0.0081 [0.0063, 0.0101]	0.0089 [0.0067, 0.0111]	-8e-04 [-0.0038, 0.0023]
	67%	0.0075 [0.0058, 0.0092]	0.0082 [0.0063, 0.0102]	-7e-04 [-0.0033, 0.002]
	75%	0.0077 [0.006, 0.0096]	0.0086 [0.0066, 0.0108]	-8e-04 [-0.0036, 0.002]
	33 vs 50%	-5e-04 [-0.0013, 4e-04]	0 [-0.001, 8e-04]	-4e-04 [-0.0016, 8e-04]
	33 vs 67%	2e-04 [-7e-04, 0.001]	7e-04 [-3e-04, 0.0017]	-5e-04 [-0.0018, 7e-04]
	33 vs 75%	-1e-04 [-9e-04, 8e-04]	3e-04 [-6e-04, 0.0013]	-4e-04 [-0.0017, 9e-04]
	50 vs 67%	6e-04 [1e-04, 0.0012]	7e-04 [1e-04, 0.0014]	-1e-04 [-0.001, 8e-04]
	50 vs 67%	4e-04 [-2e-04, 0.001]	4e-04 [-3e-04, 0.001]	0 [-8e-04, 9e-04]
	67 vs 75%	-3e-04 [-9e-04, 3e-04]	-3e-04 [-0.001, 3e-04]	1e-04 [-8e-04, 0.001]
Lower Asymptote Increment (γ)		0.0262 [0.0049, 0.0502]	0.0332 [0.0122, 0.0564]	-0.0070 [-0.0387, 0.0264]
Upper Asymptote Decrement (δ)		0.0126 [0.0001, 0.0302]	0.0337 [0.0108, 0.0591]	-0.0211 [-0.0524, 0.0090]

A 4 parameter logistic function was used a psychometric function. The threshold parameter corresponds to the point of subjective equality (the halfway point between the lower and upper asymptotes of the psychometric function) in terms of the relative value d. The parameter (θ) is a slope parameter that determines the degree to which the probability of choosing the right option changes with the relative value. We allowed both parameters to vary between the decision frame and cue conditions. To estimate sensitivity or the ability to discriminate one option from another in terms of the relative value, we estimated the psychometric function's slope at the threshold. The model was estimated as a Bayesian hierarchical model.

2.2.2 Figure with subject-level variability

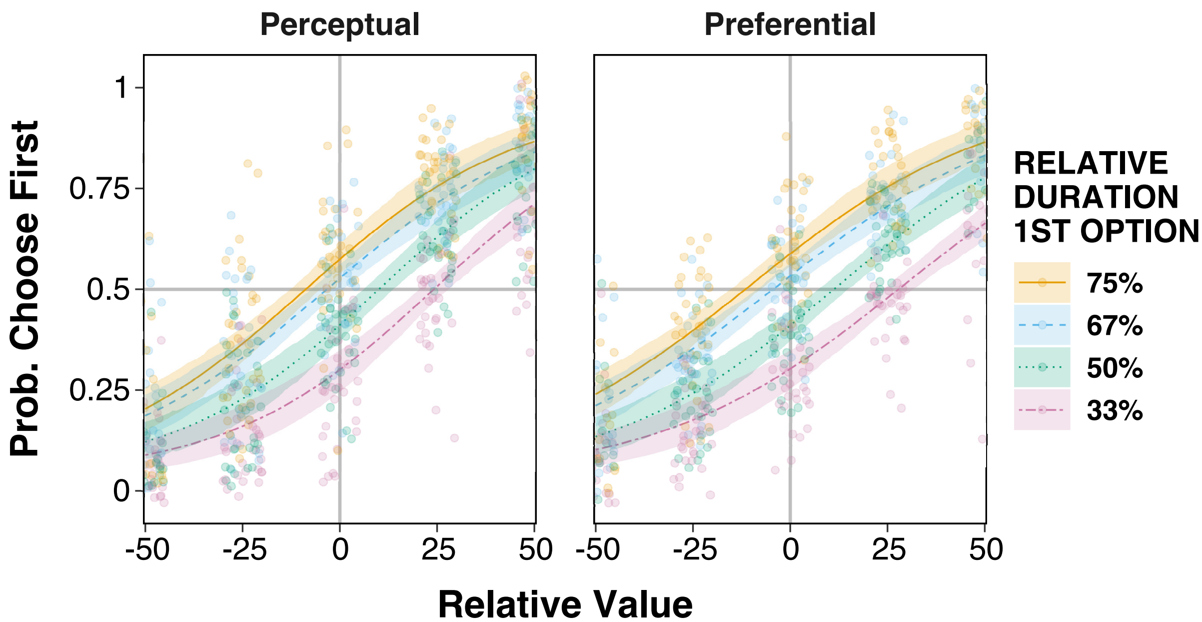

Figure S7: The probability of choosing the first option is plotted against the relative value between the two options. The data was conditioned by the decision frame (panels) and relative duration of the first option (colors). The lines represent the posterior predicted choice proportions with the error regions indicating the 95% HDIs of the proportions. The pale points represent individual participant data (jittered for better visibility).

2.3 Response times

	Coefficient	95% HDI
Intercept	**2.14**	**[1.96,2.31]**
Task frame	0.04	[-0.18,0.28]
Stimulus duration	**-0.08**	**[-0.09,-0.08]**
Unsigned relative value	0	[0,0]
Task frame x Stimulus duration	0	[-0.01,0]
Task frame x Unsigned relative value	0	[0,0.01]
Stimulus duration x Unsigned relative value	0	[0,0]
Task frame x Stimulus duration x Unsigned relative value	0	[0,0]

Table S9: Parameter means and 95% HDIs for the response time/manipulation check regression parameters for Study 2. Estimates from a hierarchical Bayesian linear regression model with random intercepts fit with *stan* (Goodrich et al., 2020). Decision frame was coded against the Perceptual frame.

The task used a cue-to-respond interrogation style procedure, therefore we do not expect any differences in response times. Nevertheless, we examined if varying stimulus exposure duration led to longer choice RTs. Response times were inversed and standardized before being regressed on the task frame (Perceptual vs. Preferential), unsigned relative value (absolute mean dot difference between options), and number of frames per trial. The analysis revealed that participants made their choices more slowly when more samples of the stimulus were displayed ($b = -0.08$ $[-0.09, -0.08]$), but this effect was quite small. There was not a credible difference in response times between between decisions frames ($b = 0.04$ $[-0.18, 0.28]$), or across differences in relative value ($b = 0$ $[0, 0]$).

2.4 Analysis of stimulus duration and stimulus switches

	Coefficient	**95% HDI**
Intercept	**0.15**	**[0.12,0.19]**
Task frame	0.01	[-0.05,0.06]
Relative value	**0.14**	**[0.12,0.17]**
Switch number	-0.03	[-0.08,0.03]
Relative duration	**0.56**	**[0.50,0.62]**
Task frame x Relative value	-0.01	[-0.05,0.02]
Task frame x Switch number	0	[-0.08,0.08]
Task frame x Relative duration	-0.02	[-0.10,0.07]
Relative value x Switch number	**0.05**	**[0.01,0.09]**
Relative value x Relative duration	**0.07**	**[0.02,0.11]**
Switch number x Relative duration	0.02	[-0.07,0.11]
Task frame x Relative value x Switch number	0.03	[-0.03,0.09]
Task frame x Relative value x Relative duration	0	[-0.06,0.07]
Task frame x Switch number x Relative duration	0.04	[-0.09,0.17]
Relative value x Switch number x Relative duration	**-0.09**	**[-0.16,-0.02]**
Task frame x Relative value x Switch number x Relative duration	-0.03	[-0.13,0.07]

Table S10: Parameter means and 95% HDIs for the order effects regression parameters for Study 2. Estimates from a hierarchical Bayesian logistic regression model with random intercepts fit with *stan* (Goodrich et al., 2020). Decision frame was coded against the Perceptual frame and number of switches was coded against the one-switch option.

Because the experimental design in Study 2 manipulated the order of stimuli presentation, we conducted a logistic regression to determine how the task framing (Perceptual vs. Preferential), relative value of the first option (-2, -1, 0, 1, and 2), number of switches (one vs. two), and relative duration of the first option for a given trial (33%, 50%, 67%, and 75%) affected choice for the first presented option. Note that in the two-switch condition the first option presented was also the last option seen. Table S10 lists the regression coefficients for the logistic regression. The odds of choosing the first option increased when the relative value and the relative duration of it increased. This effect of relative value was stronger when the first option was shown for a longer relative duration and also when it was shown a second time as indicated by the three-way interaction.

2.4.1 Order effects

We can also use this analysis and in particular the one-switch condition to assess the degree to which there were order effects in how the evidence was weighted in the choice. The posterior predicted likelihood of choosing the first option in this condition was 0.47 $[0.45, 0.50]$ in the preferential frame and 0.48 $[0.46, 0.50]$ in the perceptual condition. Thus, overall there was evidence of a small recency effect, but no difference between the decision frames (0.01 $[-0.03, 0.04]$).[1]

[1] The choice data from the two-switch condition are not diagnostic here as the first-shown option was shown first and last.

3 Computational modeling

3.1 Model description

We used a diffusion decision model (DDM; Ratcliff et al., 2016) to better investigate at the algorithmic level how attention impacted preferences. The DDM models decision making as a sequential sampling process, where participants sequentially sample information about the options and accumulate the information as evidence to make a choice. The key parameters of the DDM are outlined in Table S11.

3.1.1 Response Models

Free-viewing study There are different ways to model a response in the DDM and in our case the two studies had different response procedures. During the free-viewing study (Study 1), participants determined when to record a response. We modeled the decision as an optional stopping procedure where participants accumulated evidence to a threshold and responded with the respective choice. The parameter α indexes the separation between the choice thresholds (Table S11). The threshold separation measures the degree of response caution with lower thresholds permitting faster, but less consistent choices.

Fixed viewing study During the fixed-viewing study (Study 2), participants were cued to respond. We modeled this decision as an interrogation procedures where participants accumulated evidence and when they were cued to respond, they classified the evidence into a choice. If the evidence $L(t) > 0$ then the first option was chosen; otherwise the second option was chosen. In this case, the location of the evidence at any given point in time t is normally distributed with a mean of $\delta \times t$ and a standard deviation of $\sigma \times \sqrt{t}$, $L(t) \sim N(\delta t, \sigma\sqrt{t})$. The standard deviation,σ, is the variability in the accumulated evidence, and it was a function of the the pooled standard deviation of the stimulus and an additional parameter σ representing extraneous factors impacting the value representation, $s^2 = Var(\text{stimulus}) + \sigma^2$. Thus, the probability of choosing the first option at time t is,

$$Pr(\text{'First'}|t) = 1 - \Phi[L(t)|\delta, s]. \quad (1)$$

Where Φ is the normal cumulative distribution function. The probability of choosing the second option is $1 - Pr(\text{'First'}|t)$.

Note in principle the startpoint of the evidence accumulation could be allowed to free to vary in this task. But, we found this did not improve the fits of the models, as the models of the drift rate also had an additive factor (see below). We also explored whether the start point varied as a function of the duration of the stimulus analagous to how the cue impacted the start point in the free-viewing study. But, this also led to substantially worse fits. Thus, the startpoint was set to 0 for Study 2.

3.1.2 Value-based Model

Our first model of the drift rate posited that the drift rate was a linear function of the relative value so that,

$$\delta = \nu_0 + \nu_1 \times (\mu_{\text{right}} - \mu_{\text{left}}) \quad (2)$$

The parameter ν_0 is the baseline drift and ν_1 determines the contribution of the relative value. The variables μ_{right} and μ_{left} correspond to the mean number of dots in the right and left stimuli.

3.1.3 Additive Model

Following Cavanagh et al. (2014), we parameterized the drift rate to test different hypotheses about how gaze and value shaped the accumulating evidence. The additive model, tested for an independent influence of eye gaze and value on drift rate so that,

$$\delta = \nu_0 + \nu_1 \times (\mu_{\text{right}} - \mu_{\text{left}}) + \nu_2 \times (\text{gaze}_{\text{right}} - \text{gaze}_{\text{left}}) + \epsilon. \quad (3)$$

The parameter ν_0 is the baseline drift and ν_1, and ν_2 determine the contribution of the mean values of the options and the gaze, respectively. The variables $\text{gaze}_{\text{right}}$ and $\text{gaze}_{\text{left}}$ are the relative proportion of time

Table S11: Main Parameters of the Drift Diffusion Model and Their Substantive Interpretations

Parameter	Description
Threshold separation, α	The separation between the choice thresholds with $\alpha > 0$. With this parameterization the choice threshold for the right option is set at α, and the choice threshold for the left option set at 0. The threshold separation measures the degree of response caution with lower thresholds permitting faster, but less consistent choices.
Relative start point, β	The location of the starting point for evidence accumulation relative to the two thresholds, with $0 < \beta < 1$. The relative starting point indexes an initial bias for either response, with higher values of β indicating greater bias to choose the right option.
Drift rate, δ	The rate at which evidence accumulates in favor of the uncertain option. The sign of the drift rate indicates the average direction of the evolution, with negative values indicating evidence for the certain alternative and positive values indicating evidence for the uncertain alternative.
Non-decision time, τ	The proportion of the minimum observed response time that is allocated to other processes beyond the deliberation time specified by the DDM, with $0 < \tau < 1$. This non-decision time includes the time spent on encoding the stimulus, executing a response, and any other contaminant processes.

the participant fixated on the right and left option (Study 1) or the relative duration each option was visible (Study 2).

3.1.4 Interactive model

The interactive model formalized the hypothesis that the eye gaze directly influences the contribution of the values of the options to preference (Krajbich et al., 2010; Krajbich & Rangel, 2011; Smith & Krajbich, 2019). Specifically, the hypothesis is that the rate of evidence accumulation is a function of the difference in the values of the two options, but the value of the non-fixated option is discounted, so that,

$$d = f(\mu_{\text{fixated}} - \theta \times \mu_{\text{nonfixated}}). \tag{4}$$

Where the parameter $0 < \theta < 1$ discounts the value of the nonfixated option. Following Cavanagh et al. (2014), this hypothesis can be formalized with the following model of the drift rate,

$$\delta = \nu_0 + \nu_1 \times (\text{gaze}_{\text{right}} \times \mu_{\text{right}} - \text{gaze}_{\text{left}} \times \mu_{\text{left}}) + \nu_2 \times (\text{gaze}_{\text{let}} \times \mu_{\text{right}} - \text{gaze}_{\text{right}} \times \mu_{\text{left}}) + \epsilon. \tag{5}$$

The ratio ν_1/ν_2 corresponds to the degree to which the nonfixated option is discounted.

3.1.5 Full Model

The final model included the additive gaze component with the interactive model (Equation 6),

$$\begin{aligned}\delta =&\nu_0 + \\ &\nu_1 \times (\text{gaze}_{\text{right}} \times \mu_{\text{right}} - \text{gaze}_{\text{left}} \times \mu_{\text{left}}) + \\ &\nu_2 \times (\text{gaze}_{\text{let}} \times \mu_{\text{right}} - \text{gaze}_{\text{right}} \times \mu_{\text{left}}) + \\ &\nu_3 \times (\text{gaze}_{\text{right}} - \text{gaze}_{\text{left}}) + \epsilon.\end{aligned} \tag{6}$$

3.2 Fitting methods

The DDMs were embedded within a hierarchical framework. The hierarchical framework allows data from one subject to inform their own parameter estimates in different conditions as well as the parameter estimates of other subjects in the same conditions. It thus enabled us to acquire reliable and accurate estimates of the parameters of the decision process.

3.2.1 Trial-by-trial Error

As indicated by the *epsilon* in the drift models (Equations 3, 5, 6), we incorporated trial-by-trial variability into the drift rates (Ratcliff & Rouder, 1998). This follows the approach Cavanagh et al. (2014) took, and permits the models to account for so-called error rates where the response times for choices that are counter to the sign of the relative value (e.g., choosing right when the relative value favored the left option) are slower than the choices that are consistent with the relative value.

The response times shown in Figure 3 suggest the presence of fast errors where the response times for choices that are counter to the sign of the relative value (e.g., choosing right when the relative value favored the left option) are faster than the choices that are consistent with the relative value. This can be accounted for in the DDM with trial-by-trial variability in the startpoint. Therefore in the free-viewing study (Study 1) we modeled the relative start point from trial by trial as normally distributed to allow the start point to vary from trial to trial. In the fixed-viewing study (Study 2), we did not allow the start point to be different from 0, nor did we include trial-by-trial variability, as this model led to a substantially worse fit.

3.2.2 MCMC Estimation

We used Bayesian estimation techniques to estimate the model parameters (Kruschke, 2014; Lee & Wagenmakers, 2014). We estimated the model using a Markov Chain Monte Carlo (MCMC) sampler JAGS (Plummer et al., 2003) with the Wiener distribution provided by Wabersich and Vandekerckhove (Wabersich & Vandekerckhove, 2014). JAGS was run via matjags (Steyvers, 2011) in MATLAB 2018b and all post-processing was done with MATLAB. JAGS and post-processing code are on the OSF website for this paper.

We took an estimation approach to make inferences in this framework (Gelman et al., 2013; Kruschke, 2014). We report both the most credible value and the 95% Highest Density Interval (HDI) in brackets to describe the posterior distribution around a parameter of interest. This informed the number of samples we collected via the MCMC sampler. Each of the reported estimates in the paper were based on an estimated sample size of at least 10,000 after accounting for autocorrelation. For the optional stopping models for Study 1, due to high autocorrelation among some parameters, we collected 64,000 samples using 32 chains with 2,000 samples each, a burn in of 2,000, and thinning set to keep every fifth sample. For the interrogation models for Study 2 we collected 96,000 samples using 32 chains with 3,000 samples each, a burn in of 3,000, and no thinning. To test differences between conditions, we compute the difference between the conditions and test whether the 95% HDI contains a null value of 0. If it does not, we describe the difference between conditions is credible.

3.3 Study 1

We fit value-based model (Equation 2), the additive model (Equation 3), the integrative model (Equation 5), and the full model (Equation 6) to the free-viewing dataset. In all the models, we allowed the relative start point β to vary by the location of the attentional cue, allowing us to measure how much the cue changed the initial bias.

3.3.1 Model Comparison

For each drift model we calculated Deviance Information Criterion (DIC; Spiegelhalter et al., 2014). The DIC's for each model in each decision frame are listed in Table S12. The lower the DIC the better the fit. These model comparisons show that in the preferential frame the Full Model that has an interactive component allowing for the value of the non-fixated option to be discounted and an additive component of relative gaze provided the best fit. In contrast, in the perceptual frame the best fitting model was the additive model where relative gaze contributed independently to the drift rate and hence the evidence being accumulated. Given these differences, we used the Full Model in the main paper where both the interactive and additive model are special cases. Consistent with the model comparisons presented here, in the Full Model, the degree to which the non-fixated option was discounted in the perceptual frame was credibly lower than in the preference frame (see Main Text).

Table S12: Comparisons of different models of drift in the DDM for Study 1.

Model	Preference	Perceptual
Value-based	33265	40798
Additive	32540	**38897**
Interactive	32457	39254
Full	**32274**	39415

The values in each cell are the respective DIC values, with lower values indicating better model fit.

3.3.2 Posterior predictive fits to choice behavior

The posterior predictive fits of the Full Model to the choice data are shown in Figure S8. The fits show that DDM recreates the data and the different effects quite well. For instance, they show the cue effect with the left cue decreasing the probability of choosing the right option and the right cue increasing the probability choosing the right option. They also show that the effect of the cue was greater in the preferential vs. perceptual frame. Finally, the change in the probability of choosing the right option as the relative value increased was smaller in the preferential frame compared to the perceptual frame. One thing to note, however, is that the posterior predicted choice probabilities for both decision frames were a bit more regressive to 0.5 than the observed data. This deviation is partly due to the hierarchical structure of the model. The other reason is that the DDM is simultaneously fit to the observed choices and response times as compared to, for instance, a psychometric function which is only fit to the observed choices.

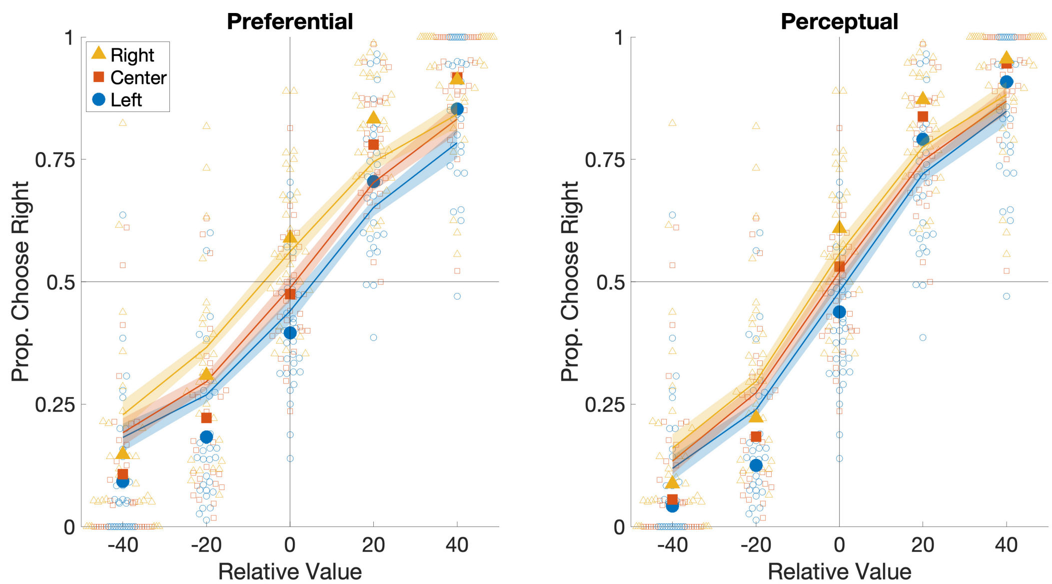

Figure S8: The probability of choosing the right option is plotted against the relative value (difference in mean dots between options, right-left). The data was conditioned by the decision frame (panels) and cue location (colors). The lines represent the posterior predicted choice proportions from the DDM with the full drift model, with the error regions indicating the 95% HDIs and the pale points showing individual participant data (jittered for better visibility).

3.3.3 Posterior predictive fits to response time distributions

Figures S9, S10, S11, and S12 show the posterior predictive fits of the Full Model to the response time distributions collapsed across participants. Each panel plots the .1, .3, .5, .7, and .9 quantiles of the response time distributions when either the right or left option was chosen for the preferential and perceptual decision frame. Across the columns of panels, the relative value is increasing. Across the rows, the location of the cue is changing. The observed quantile is marked by $\times$. A red dot marks the mean posterior predicted quantile, and the 95% HDI around each quantile is denoted by the black line. In general, the model does a good job recreating the response time distributions. The one small deviation is that when the right option was chosen, and the relative value favored the left option, the observed responses were slightly faster than predicted by the model. This difference may be due to the left bias that was apparent in the eye tracking data.

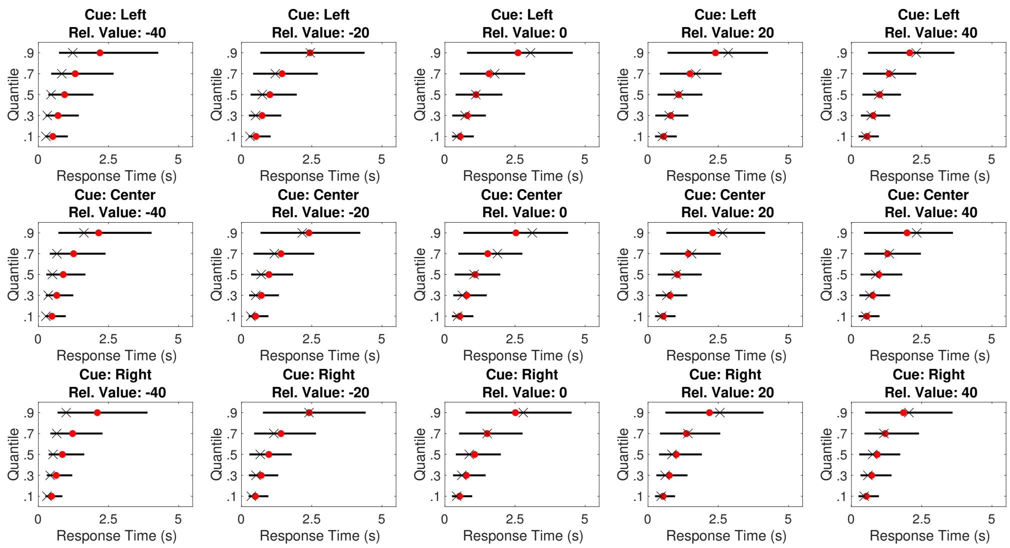

Figure S9: Response time quantiles in the preferential frame when the right option was chosen. Each panel plots the observed quantile (x) and the mean (red dot) and 95%HDI (black bar) posterior predicted quantile. Across the rows of the figure the cue was on the left, center, and right. Across the columns of the figure the relative value increased.

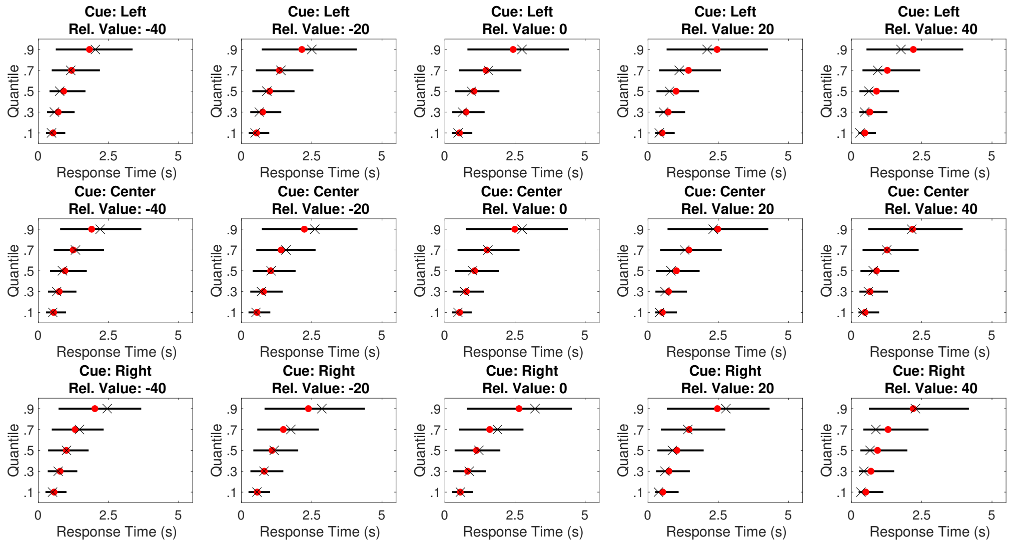

Figure S10: Response time quantiles in the preferential frame when the left option was chosen. Each panel plots the observed quantile (x) and the mean (red dot) and 95%HDI (black bar) posterior predicted quantile. Across the rows of the figures the cue was on the left, center, and right. Across the columns of the figures the relative value increased.

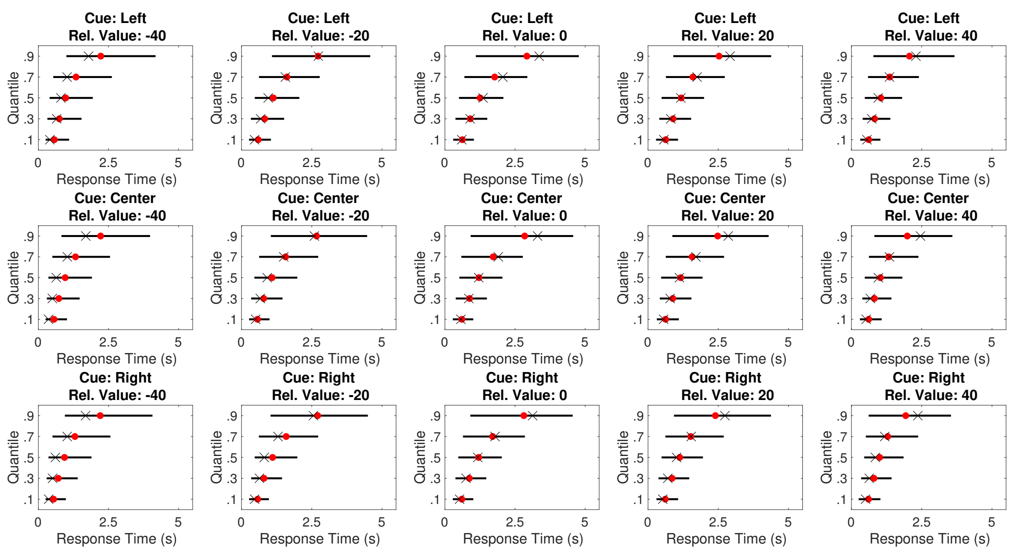

Figure S11: Response time quantiles in the perceptual frame when the right option was chosen. Each panel plots the observed quantile (x) and the mean (red dot) and 95%HDI (black bar) posterior predicted quantile. Across the rows of the figures the cue was on the left, center, and right. Across the columns of the figures the relative value increased.

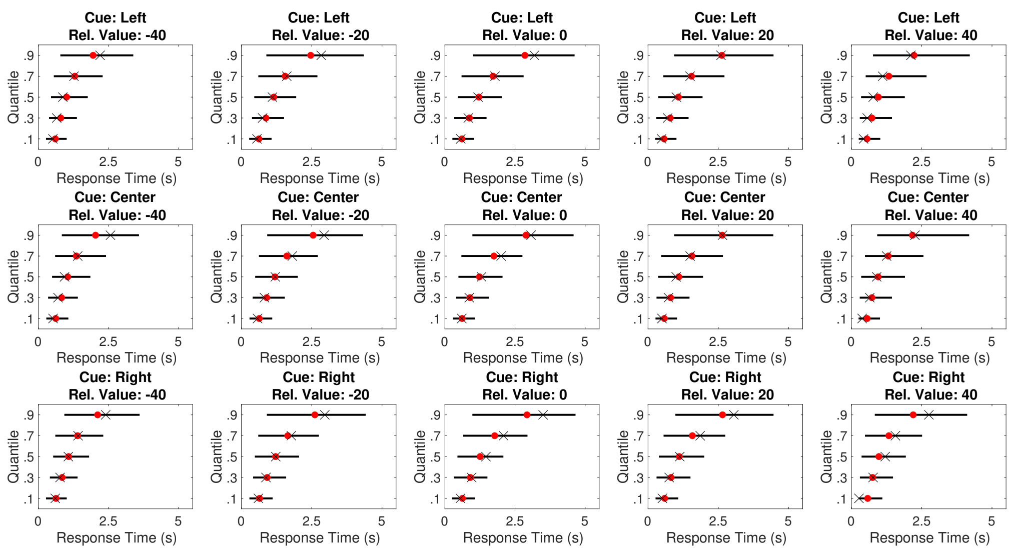

Figure S12: Response time quantiles in the perceptual frame when the left option was chosen. Each panel plots the observed quantile (x) and the mean (red dot) and 95%HDI (black bar) posterior predicted quantile. Across the rows of the figures, the cue was on the left, center, and right. Across the columns of the figures, the relative value increased.

Table S13: Model comparisons of different models of drift in the DDM for Study 2.

Model	Preference	Perceptual
Value-based	7188	6993
Additive	7154	6961
Interactive	**7075**	**6949**
Full	7195	7098

The values in each cell are the respective DIC values, with lower values indicating better model fit.

3.4 Study 2

We fit all four models to the fixed viewing study (Study 2), each as a Bayesian hierarchical model. Each of the DDMs modeled the response as an interrogation response procedure. We sought to make the models comparable to the models in Study 1. We explored whether the relative duration impacted the starting point of evidence accumulation, but those models provided a substantially worse fit. Thus, the starting point of evidence accumulation was set at $L(t) = 0$.

3.4.1 Model Comparisons

Table S13 lists the DIC's for each of the four drift rate models for both decision frames. Unlike Study 1, the interactive model provided the best fit to the data for both the preferential and perceptual frames. This suggests that there is no independent contribution of relative gaze to drift and hence evidence accumulation during fixed viewing. Consistent with this, the coefficient in the Full model for the additive contribution of gaze (ν_3) is not credibly different from 0 for both the preferential and perceptual frames (Table 3 in the main text). Although the interactive model provided the best fit, there was still credibly more discounting of the non-fixated option in the preferential frame (see Main Text). Notably, in the preferential frame, the degree of discounting was less in Study 2 compared to Study 1. In comparison, in the perceptual frame the degree of discounting was greater in Study 2 compared to Study 1. This suggests that some degree of the discounting is due to the free allocation of attention. When participants provide a more balanced allocation of attention (as in the perceptual frame of Study 1) the degree of discounting is decreased. For ease of comparison, we report the full model in the main paper.

3.4.2 Posterior predictive fits to choice behavior

Figure S13 displays the posterior predictive estimate of the probability of choosing the first option against the observed data. The model recreates the data well, capturing the effect of the relative duration on the probability of choosing the first option and the stronger effect of this duration in the preferential frame compared to the perceptual frame.

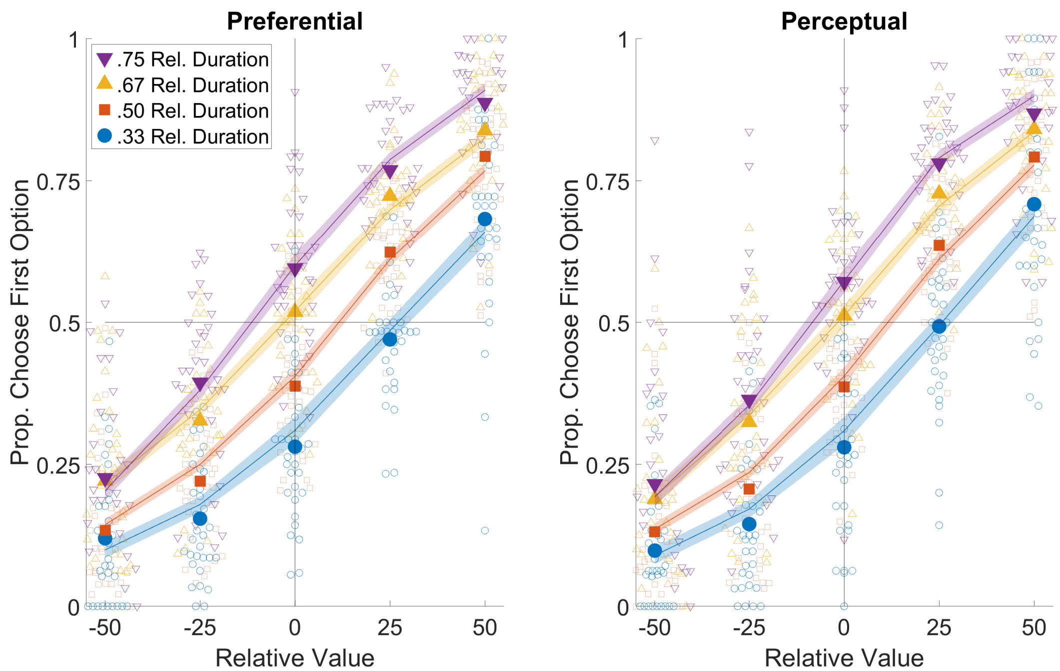

Figure S13: The probability of choosing the first option is plotted against the relative value between the two options. The data was conditioned by the decision frame (panels) and relative duration of the first option (colors). The lines represent the posterior predicted choice proportions of the DDM with the full drift model with the error regions indicating the 95% HDIs of the proportions. The pale points represent individual participant data (jittered for better visibility).

References

Cavanagh, J. F., Wiecki, T. V., Kochar, A., & Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology: General*, *143*(4), 1476–1488. doi: 10.1037/a0035813

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis.* CRC press.

Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2020). *rstanarm: Bayesian applied regression modeling via Stan.* Retrieved from `https://mc-stan.org/rstanarm` (R package version 2.21.1)

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298. doi: 10.1038/nn.2635

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, *108*(33), 13852–13857. doi: 10.1073/pnas.1101328108

Kruschke, J. (2014). *Doing bayesian data analysis: A tutorial with r, jags, and stan.* Academic Press.

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course.* Cambridge university press.

Plummer, M., et al. (2003). Jags: A program for analysis of bayesian graphical models using gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, pp. 1–10).

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347–356. doi: 10.1111/1467-9280.00067

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

Smith, S. M., & Krajbich, I. (2019). Gaze amplifies value in decision making. *Psychological Science*, *30*(1), 116–128. doi: 10.1177/0956797618810521

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van der Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, 485–493.

Steyvers, M. (2011). *Matjags 1.3: A matlab interface for jags.*

Wabersich, D., & Vandekerckhove, J. (2014). Extending jags: A tutorial on adding custom distributions to jags (with a diffusion model example). *Behavior Research Methods*, *46*(1), 15–28.