Supplementary Notes

<u>Supplementary Note 1: Model fitting</u>

Assume there are two alternatives X & Y and two attributes A & B. When a decision maker is gazing at a certain element, the drift rates are expressed as follows:

$$\text{gaze at } X_A: \quad v_t = v_{t-1} + d[(X_A - \theta Y_A) + (\phi X_B - \gamma Y_B)]$$

$$\text{gaze at } X_B: \quad v_t = v_{t-1} + d[(\phi X_A - \gamma Y_A) + (X_B - \theta Y_B)]$$

$$\text{gaze at } Y_A: \quad v_t = v_{t-1} + d[(\theta X_A - Y_A) + (\gamma X_B - \phi Y_B)]$$

$$\text{gaze at } Y_B: \quad v_t = v_{t-1} + d[(\gamma X_A - \phi Y_A) + (\theta X_B - Y_B)]$$

Over the course of a trial, the dwell time proportion on $X_A, Y_A, X_B,$ & Y_B are assumed to be $p, q, m,$ & n respectively. The drift rate can be re-expressed as follows:

$$v \sim d[\, p(X_A + \phi X_B - \theta Y_A - \gamma Y_B) + q(\theta X_A + \gamma X_B - Y_A - \phi Y_B)$$

$$+ m(\phi X_A + X_B - \gamma Y_A - \theta Y_B) + n(\gamma X_A + \theta X_B - \phi Y_A - Y_B)]$$

$$v \sim d[\,(pX_A - qY_A) + (mX_B - nY_B) + \theta(qX_A - pY_A) + \theta(nX_B - mY_B)$$

$$+ \phi(mX_A - nY_A) + \phi(pX_B - qY_B) + \gamma(nX_A - mY_A) + \gamma(qX_B - pY_B)]$$

In addition, an attribute weight parameter ω is introduced to the model. Let ω be the weight on attribute A and $1 - \omega$ be the weight on attribute B. The re-expressed full model is as follows:

$$v \sim d[\, \omega(pX_A - qY_A) + (1 - \omega)(mX_B - nY_B) + \theta\omega(qX_A - pY_A) + \theta(1 - \omega)(nX_B - mY_B)$$

$$+ \phi\omega(mX_A - nY_A) + \phi(1 - \omega)(pX_B - qY_B) + \gamma\omega(nX_A - mY_A) + \gamma(1 - \omega)(qX_B - pY_B)]$$

Besides the scaling parameter d, attribute weight parameter ω, and attentional discounting parameter $\theta, \phi, \& \gamma$, there were two other free parameters in DDM: 1). Non-decision time t_{er} 2). Within-trial variability σ. Boundary separation was fixed ($a = 1$) as in Krajbich et al. (2010), and between-trial variability was not included. Note that in the initial model fitting process, we allowed boundary separation to vary and fixed within-trial variability due to restrictions of the wiener likelihood sampler in RSTAN. We later transformed the boundary separation parameter to within-trial variability σ using a simple transformation(see below). Uniform priors were used for the attentional parameters and attribute weight parameter. For the other parameters, we followed Wiecki et al. (2013) and put gamma distributions on a and t_{er}. Specifically, $a \sim Gamma\ (3,2), t_{er} \sim Gamma\ (0.8,2)$.

Model convergence was assessed by Gelman–Rubin statistic $\hat{R}$ (Gelman & Rubin, 1992). The largest $\hat{R}$ was 1.001 across five datasets, suggesting successful model convergence. We also checked effective sample sizes to ensure estimation accuracy. Effective sample sizes were larger than 2500 for all the parameters in our datasets, which is enough for inference (Carpenter et al., 2017).

We used the following equations to transform parameters from RSTAN to the aDDM/maaDDM formulation (Ractliff, 1978):

$$v: \quad v = \frac{2*v_{stan}}{a_{stan}*1000} \qquad d: d = \frac{2*d_{stan}}{a_{stan}*1000} \qquad \sigma: \quad \sigma = \frac{2}{a_{stan}} * \sqrt{\frac{1}{1000}}$$

We also used the following equations to transform our attribute-wise attentional discount parameter when fitting the model: $\phi = \frac{1-\lambda}{\lambda}, \quad \gamma = \frac{\gamma^*}{\lambda}$

$$v = d[\ \lambda\omega(pX_A - qY_A) + \lambda(1-\omega)(mX_B - nY_B) + \theta\lambda\omega(qX_A - pY_A)$$

$$+\ \theta\lambda(1-\omega)(nX_B - mY_B) + (1-\lambda)\omega\,(mX_A - nY_A) + (1-\lambda)(1$$

$$-\omega)(pY_B - qY_B) + \gamma^*\omega(\ nX_A - mY_A) + \gamma^*(1-\omega)(qX_B - pY_B)]$$

Supplementary Note 2: Parameter recovery

We generated data with the following parameters: $d \in \{0.0002, 0.0003, 0.00035, 0.0004\}$, $\sigma \in$

$\{0.02, 0.025\}$, $t_{er} \in \{300, 500, 600\}$, $\omega \in \{0.2, 0.5, 0.8\}$, $\theta \in \{0.2, 0.5, 0.8\}$, $\phi \in$

$\{0.2, 0.5, 0.8\}$, $\gamma \in \{0.2, 0.5, 0.8\}$. We sampled 480 combinations of those parameter values. For

each combination, 10 datasets (with 100 trials per dataset) were simulated using the empirical

subjective value ratings and gaze patterns from the food gamble task. The correlations, mean

absolute errors (MAE), and root mean square errors (RMSE) between the true parameters and

the recovered parameters are listed in the Table S4. Figure S5 also visualizes the different levels

of true values of and their corresponding recovered values. Additionally, we also calculated the

correlations between the recovered parameters (Table S5).

We also conducted additional parameter recovery checks using the method from Hawkins

& Heathcote (2021). We started with the maaDDM parameters estimated from the full model in

each dataset. We then randomly sampled a parameter vector from the posterior distribution of

model parameters and used them to generate data with the same properties as the real data (i.e.,

same gaze data, values, and number of the trials). Then we estimated the best-fitting model

parameters. This exercise confirms that parameter recovery works well (Fig S10).

Supplementary Note 3: value rescaling

In the food + nutrition label task, we calculated the healthiness of each food as follows.

On the nutrition label, there are three possible colors, green, yellow, and red. Green labels add to

healthiness, red labels subtract from healthiness, and yellow labels are neutral. We used "1" to represent green, "0" to represent yellow, and "-1" to represent red. For each food, we recorded the number of green components, yellow components, and red components. We then summed them up to represent how healthy the food is. For example, if a product contained 1 green, 3 yellow, and 0 red components, then the health information for this product would be "+1". We then rescaled these values to go from 1 to 10, to align with the other datasets.

To ensure that it was reasonable to rescale all the attribute values from 1 to 10, we did several analyses with the clothing + brand dataset and food + nutrition dataset.

In the clothing + brand dataset, brand labels were scrambled for half of the trials. We used those trials to test if dwell time on options positively influenced choices even when the ratings were non-positive. In other words, we tested whether negatively rated clothing images were truly aversive. We selected trials where both options contained non-positively rated clothing images. We found a positive relationship between the dwell time advantage and choice probability, even in these cases (Figure S2).

Additionally, we looked at the probability of choosing the first seen item as a function of the first fixation duration. The duration of the first fixation generally had a positive effect on choice (except when the absolute value difference was zero, Figure S2).

We did a similar check in the food & nutrition dataset. We selected trials where both options contained non-positively rated food items. We also found a positive relationship between dwell time and choice, even when the food items' ratings were non-positive (Figure S2). Also, the duration of the first fixation was generally correlated with choice (except when the absolute value difference was four, Figure S2).

We further ran two mixed-effects logistic regressions on these same sets of trials.

1). *choose left* on *value difference (clothing + brand: left clothing rating – right clothing rating; food & nutrition: left food rating – right food rating), and dwell fraction difference (left -right).* The coefficients on dwell fraction were as follows: brand & clothing: $\beta = 0.42$, p $= 10^{-17}$; food & nutrition label: $\beta = 0.15$, p $= 10^{-17}$.

2). *choose first seen option* on *value difference (clothing + brand: left clothing rating – right clothing rating; food & nutrition: left food rating – right food rating), and first fixation duration.* The coefficients on first fixation duration were as follows: *clothing + brand*: $\beta = 2.04$, p $= 10^{-12}$; food & nutrition label: $\beta = 0.28$, $p = 0.0023$. Therefore, we concluded that it was reasonable to include these data in the main-text analyses.

To test whether the unhealthy labels in the food + nutrition dataset, as well as the brand labels in the clothing + brand task were aversive or not, we looked at the effect of label dwell time on choice in mixed-effects models. Note that every trial comprises one healthy food and one unhealthy food in the food & nutrition datasets. Similarly, every choice consists of one high-ranked brand and one low-ranked brand in the brand & clothing dataset. We looked at

1). *Choose left option* on *attribute difference (left minus right) and dwell time advantage on left brand/nutrition label.* The coefficients on dwell time advantage were as follows: *clothing + brand*: $\beta = 0.31$, p $= 10^{-17}$; food & nutrition label: $\beta = 0.82$ p $= 10^{-17}$).

2). *Choose higher ranked brand/ healthier food* on *absolute attribute difference* and *dwell time advantage on higher ranked brand/ healthier food.* The coefficients on dwell time advantage were as follows: *clothing + brand*: $\beta = 0.31$, p $= 10^{-17}$; food & nutrition label: $\beta = 0.81$, p $= 10^{-17}$).

Here, we saw a positive relationship between label dwell time and choices (Figure S3 & S4). Therefore, we concluded that it was reasonable to assign positive values to these labels.

Supplementary Note 4: model comparison

The model comparison results reported in the main text were based on aggregated WAICs. Here, we reported individual level WAIC to ensure the robustness of the model comparison results.

According to individual level WAIC, 35 out of 42 subjects in the food gamble task and 23 out of 33 subjects in the money gamble task were better described by the model without ω; 20 out of 28 subjects in the clothing + brand task, 11/20 out of 20 subjects in the compassion/envy condition of the dictator game, and 39/40 out of 48 subjects in the food & TL/GDA nutrition label task were better described by the model with ω (see Table S6 for more details).

Supplementary Note 5: Detailed behavioral results in Table 1

First column: test whether ω^* was significantly different from 0.5 using a two-tailed t test: food gamble: $t(41) = 1.62, p = 0.11, 95\%$ CI (top food): $[0.5, 0.54]$); money gamble: $t(32) = 1.31, p = 0.20, 95\%$ CI (top payoff): $[0.50, 0.51]$); compassion: $t(19) = 10.78, p = 10^{-10},, 95\%$ CI (self-payoff): $[0.69, 0.79]$); envy: $t(19) = 31.07, p = 10^{-17}, 95\%$ CI (self-payoff): $[0.87, 0.93]$); food + GDA label: $t(16) = 10.78, p = 10^{-17},, 95\%$ CI (taste): $[0.79, 0.88]$); food + TL label: $t(47) = 11.81, p = 10^{-15}, 95\%$ CI (taste): $[0.73, 0.83]$); clothing + brand: $t(27) = 13.02, p = 10^{-14},, 95\%$ CI (top food): $[0.76, 0.85]$).

Second & third columns: coefficients of the mixed-effects logistic regression of log(RT) (in milliseconds) on *absolute Attribute A value difference (|left – right|) and absolute Attribute B value difference (|left – right|):* The coefficients were as follows: food

gamble ; β_{top} = -0.039, p =10^{-8}; β_{bottom} = -0.023, p = 0.0011; money gamble: β_{top} = 0.018, p = 0.08; β_{bottom} = −0.0039, p = 0.69; compassion: β_{self} = -0.025, p = 0.34; β_{other} = - 0.033, p = 0.060; envy: β_{self} = -0.064, p = 0.0013; β_{other} = -0.030, p = 0.028; food + GDA label: β_{taste} = -0.12, p = 10^{-12}; $\beta_{nutrition}$ = -0.042, p = 10^{-6}; food + TL label: β_{taste} = -0.089, p = 10^{-9}; $\beta_{nutrition}$ = -0.045, p = 10^{-8} ; clothing + brand: $\beta_{clothing}$ = -0.053, p = 10^{-12}; β_{brand} = 0.0022, p = 0.71.

Fourth column: a mixed-effects logistic regression of *choose better option on Attribute A* (1 or 0) on *absolute attribute A value difference*, *attribute B value difference* (relative to attribute A, i.e., positive when it is in line with attribute A), and *gaze proportion on Attribute A* (within each task). See Supplementary Table 1.

Fifth column: a mixed-effects logistic regression of *choose left option* (1 or 0) on *Attribute A value difference* (left − right), *Attribute B value difference* (left − right), and *gaze proportion difference* (left − right). See Supplementary Table 2.

Supplementary Table 1. The influence of attribute attention on choice

	Main effect on choosing the option that is better on Attribute A			
	Intercept.	$\|VD_{Attribute\ A}\|$	$VD_{Attribute\ B\ to\ A^*}$	$Gaze_{Attribute\ A}$
Food gamble	0.23***	1.06***	0.49***	0.10**
Money gamble	0.029	0.20***	0.23***	0.096*
Compassion	-0.47*	0.34***	0.40***	0.17*
Envy	3.49***	1.82 ***	0.45**	0.14
Food + GDA	1.29***	2.01***	-0.12	0.15*
Food + TL	1.14***	2.12***	0.22†	0.16*
Clothing +Brand	1.087***	1.61***	0.44***	0.079*

Note. This table reported the group level results from the mixed effect logistic regression of choosing better option on *absolute attribute A value difference, attribute B value difference* (relative to attribute A, i.e., positive when it is in line with attribute A*),* and *gaze proportion on Attribute A* (within each task). Attribute A represents the top row in the food gamble and the money gamble task; "self" in the dictator game; taste in the food + nutrition task; clothing appearance in the clothing + brand task. Attribute B represents the bottom row in the food gamble and the money gamble task; "other" in the dictator game; nutrition in the food + nutrition task; brand in the clothing + brand task.
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; † $p < 0.1$

Supplementary Table 2: The influence of option attention on choice

	Main effect on choosing the left option			
	Intercept.	$VD_{Attribute\ A}$	$VD_{Attribute\ B}$	$Gaze_{left\ to\ right}$
Food gamble	0.026	0.86***	0.80***	1.01***
Money gamble	0.016	1.54***	1.62***	1.13***
Compassion	0.079	1.50***	2.09***	1.19***
Envy	-0.16	4.79***	0.75	0.79***
Food + GDA	0.10	2.06***	0.31*	1.25***
Food + TL	0.017	1.97***	0.61***	1.21***
Clothing +Brand	0.076	1.53***	0.76***	1.43***

Note. This table reported the group level results from the mixed effect logistic regression of *choose left option* (1 or 0) on *Attribute A value difference* (left – right), *Attribute B value difference* (left – right), and *gaze proportion difference* (left – right) within each task. Attribute A represents the top row in the food gamble and the money gamble task; "self" in the dictator game; taste in the food + nutrition task; clothing appearance in the clothing + brand task. Attribute B represents the bottom row in the food gamble and the money gamble task; "other" in the dictator game; nutrition in the food + nutrition task; brand in the clothing + brand task.
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; † $p < 0.1$

Supplementary Table 3: correlation between attentional bias parameters

	θ, ϕ	θ, γ	ϕ, γ	$\theta*\phi, \gamma$
Food gamble	-0.089	0.66***	0.049	0.67***
Money gamble	0.31	0.71***	0.093	0..47**
Compassion	-0.69**	0.35	-0.17	0.47**
Envy	0.28	0.56*	0.065	0.34
Food + GDA	0.1	0.39**	0.019	0.37*
Food + TL	-0.07	0.097	0.1	0.22
Clothing +Brand	-0.28	0.5**	-0.2	0.5***

Note. This table reported the correlations between attentional parameters estimated from the best fitted maaDDM models in each dataset. θ represents the option-level attentional bias; ϕ represents the attribute-level attentional bias; γ represents the attentional bias on the non-fixated attribute of the non-fixated option.
 * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$;

Supplementary Table 4: parameter recovery results

	d	σ	t_{er}	ω	θ	ϕ	γ
Cor.	0.79	0.88	0.90	0.97	0.84	0.75	0.81
MAE	4.47e-05	0.0009	41.78	0.048	0.12	0.13	0.12
RMSE	6.02e-05	0.0011	54.23	0.061	0.15	0.17	0.16

Note. This table reported the summary statistics of the parameter recovery. Cor. Indicates the correlation between the true parameters we simulated and the recovered parameters we estimated from the model fitting with simulated datasets from different parameter combinations; MAE = Mean Absolute Error; RMSE = Root Mean Square Error.
d represents the value scaling parameter; σ represents the within-trial variability parameter; t_{er} represents non-decision time; ω represents the attribute weighting parameter; θ, ϕ, γ represent attentional bias parameters on different locations;

Supplementary Table 5: Correlations between recovered parameters

	d	σ	t_{er}	ω	θ	ϕ	γ
d		-0.480	0.532	0.029	-0.213	-0.005	-0.208
σ			-0.468	-0.029	0.061	0.084	0.042
t_{er}				-0.009	-0.047	-0.058	-0.046
ω					0.028	-0.204	-0.163
θ						0.084	0.256
ϕ							-0.227
γ							

Note. The values shown in this table represent the correlations between the recovered parameters. We reported the correlations between the attentional bias parameters in Supplementary Table 3. These values indicate correlations between parameters that arise due to model-fitting, rather than due to true correlations in the data.

Supplementary Table 6: Individual-level model comparison results based on WAIC

Dataset	$\gamma - free$	$\gamma = \theta\phi$	$\gamma = \theta$	$\gamma = \phi$
Food gamble	10	9	**13**	10
Money gamble	3	6	**12**	**12**
Compassion	1	**9**	4	6
Envy	5	**9**	4	2
Food + GDA label	9	**20**	5	14
Food + TL label	9	**22**	4	13
Clothing + brand	10	**15**	3	0

Note. The values shown in this table represents the number of participants whose data were best described by different variants of maaDDM models after accounting for attribute weighting. $\gamma - free$ represents the model with γ as a free parameter; $\gamma = \theta\phi, \gamma = \theta$, $\gamma = \phi$ represent three restricted versions of the models.

Supplementary Table 7: Full aggregated WAIC results

Dataset	$\gamma - free$	$\gamma = \theta\phi$	$\gamma = \theta$	$\gamma = \phi$
Food gamble	20167.6	**20160.7**	20171.9	20236.9
Money gamble	20167.2	20230.3	**20159.0**	20263.4
Compassion	3580.6	**3568.6**	3590.8	3604.4
Envy	6442.2	**6441.8**	6467.3	6484.8
Food + GDA label	15423.8	**15403.4**	15452.4	15454.5
Food + TL label	15563.7	**15557.9**	15623.2	15583.8
Clothing + brand	17191.9	**17190.2**	17217.5	17444.4

Note. The values shown in this table represents the aggregated WAICs across participants in each dataset. $\gamma - free$ represents the model with γ as a free parameter; $\gamma = \theta\phi, \gamma = \theta, \gamma = \phi$ represent three restricted versions of the models.
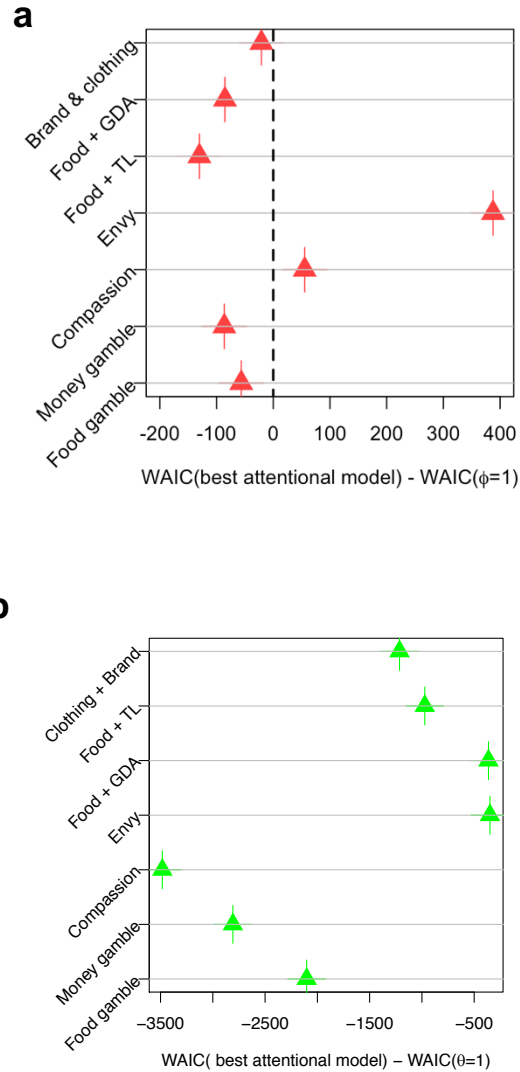
Supplementary Figure 1. Model comparisons between the best attentional model and the model with a). $\phi = 1$ (no attribute attentional bias) and b). $\theta = 1$ (no option attentional bias). The x-axis represents the WAIC difference between the two competing models. The best attentional model is the model with parameters $d, \sigma, t_{er}, \theta$, and ϕ in the food gamble and money gamble tasks; The best attentional model is the model with parameters $d, \sigma, t_{er}, \omega, \theta$, and ϕ in the other tasks.
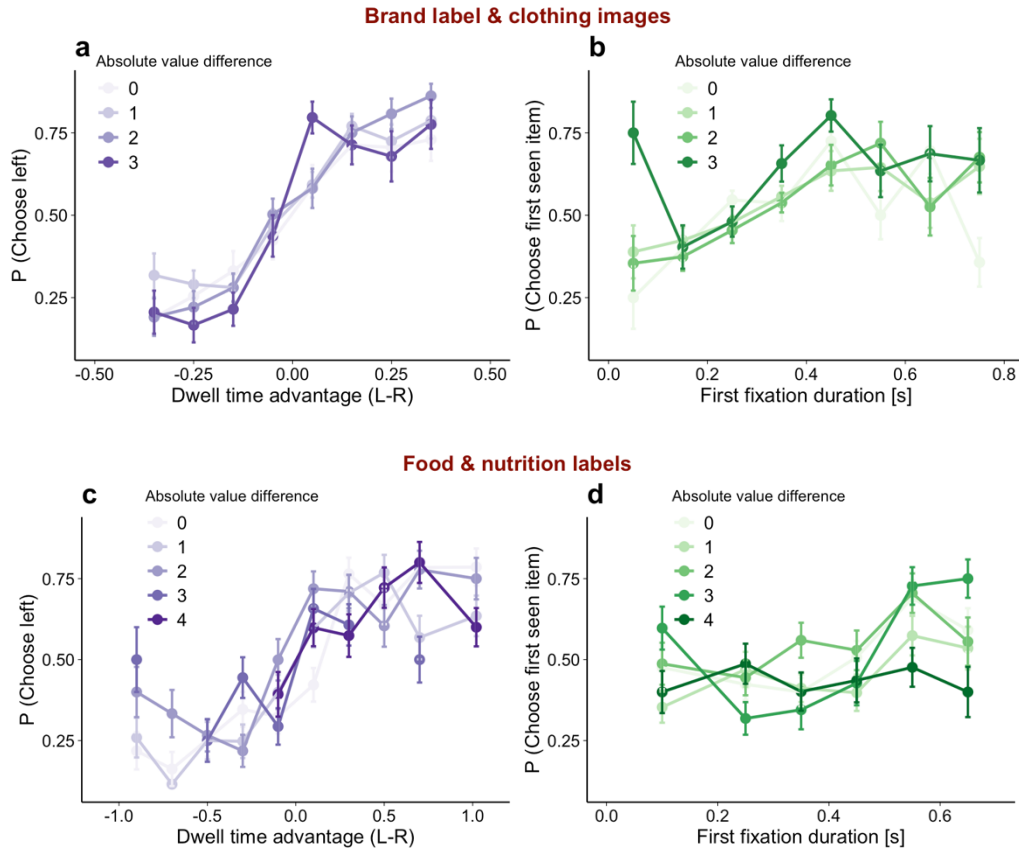
Supplementary Figure 2. Effect of dwell time on choice with non-positively rated items in (a-b) brand + clothing task and (c-d) food + nutrition label task. (a, c) represent the probability of choosing left as a function of dwell time advantage and (b, d) represent the probability of choosing first seem item as a function of the duration of the first fixation. Only trials containing non-positively rated clothing items in the scrambled brand + clothing dataset were used for this analysis and only trials containing non-positively rated food items in the food + nutrition label task were used for the analysis. Note that the analysis for the food + nutrition label task pooled the TL and GDA label trials. The absolute value difference was calculated as the absolute value difference between left and right clothing image ratings in the scrambled brand + clothing task, and the difference between the food ratings in the food + nutrition label task.
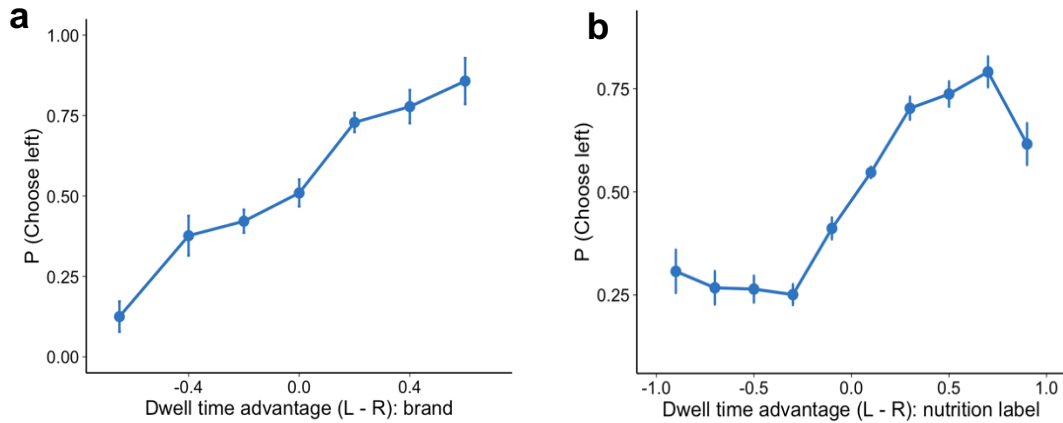
Supplementary Figure 3. Probability of choosing left as a function of dwell time advantage for a) brand label and b) nutrition label. Trials containing brand labels were selected from the brand + clothing dataset for the analysis. Note that the analysis for the food + nutrition label task combined the TL and GDA label trials.
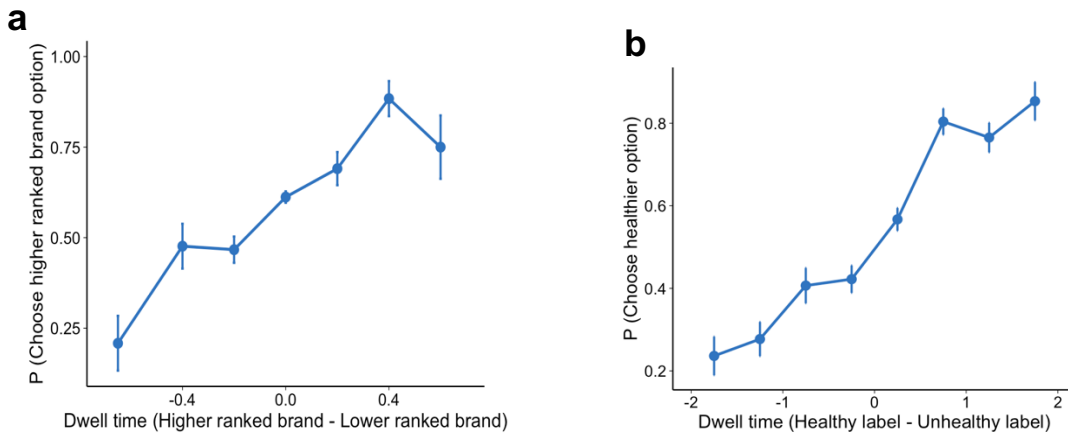
Supplementary Figure 4. Probability of choosing a) the higher ranked brand option as a function of the dwell advantage on the higher ranked brand, and b) the healthier food option as a function of the dwell advantage on the healthy food nutrition label. Trials containing brand labels were selected from the brand + clothing dataset for the analysis. Note that the analysis for the food + nutrition label task combined the TL and GDA label trials.
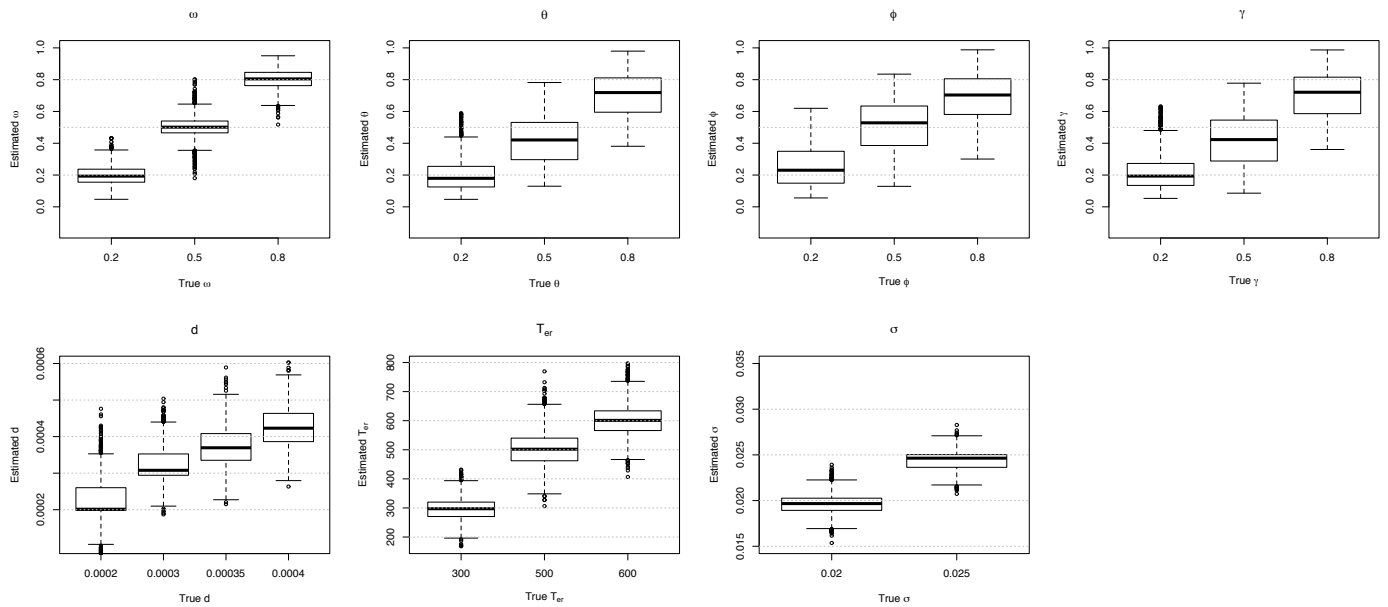
Supplementary Figure 5. Parameter recovery results. The comparisons between the recovered parameters from the model fitting and their true values. The black horizontal lines indicate the median estimated values and rectangles indicate the interquartile ranges. The bars represent the lowest/highest value excluding outliers, and the dots represent outliers that are outside 1.5 times the interquartile range above the upper/lower quartiles. ω represents the attribute weighting parameter; θ, ϕ, γ represent attentional bias parameters on different locations; d represents the value scaling parameter; t_{er} represents non-decision time; σ represents the within-trial variability parameter.
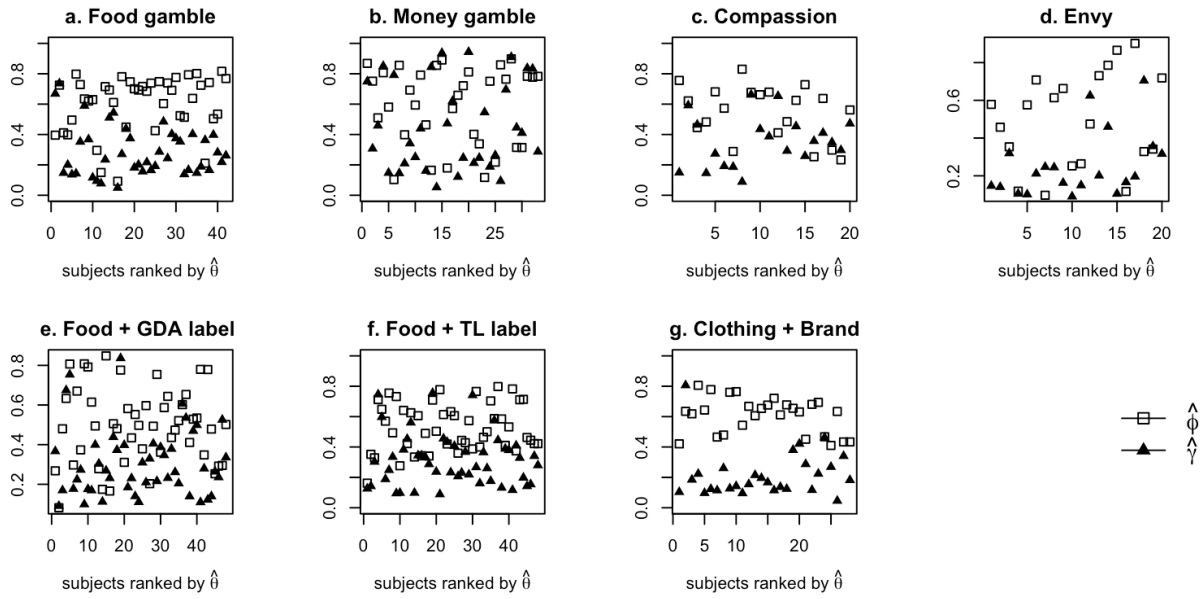
Supplementary Figure 6. Individual differences in attentional biases. The x-axis represents subjects ranked by their option-level attentional parameter θ. The symbols represent each subject's attribute-level attentional parameter ϕ and the diagonal element attentional parameter γ.
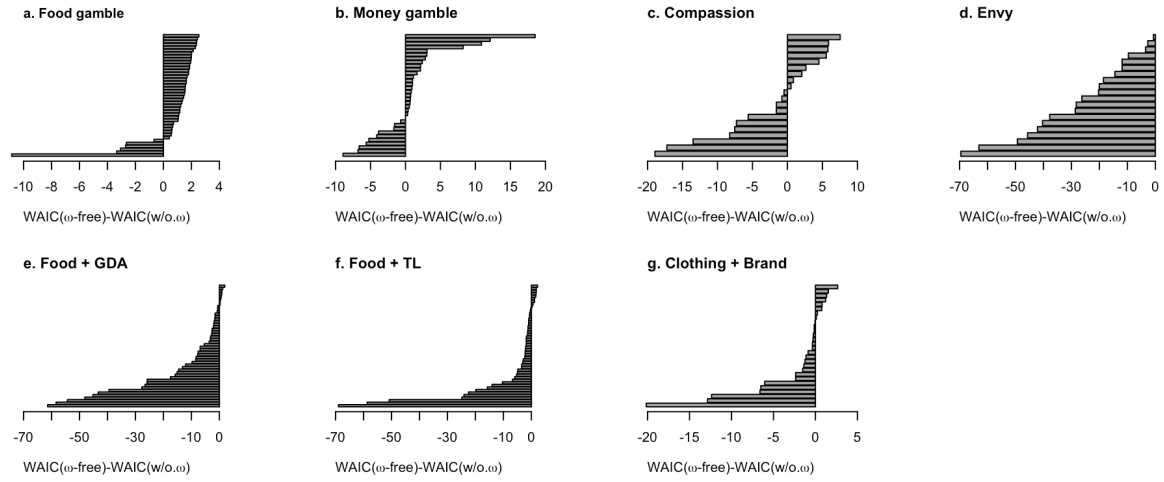
Supplementary Figure 7. Individual WAIC differences between the full model and the restricted model without the attribute weighting parameter ω. Negative differences indicate a better fit with the full model.

Food gamble
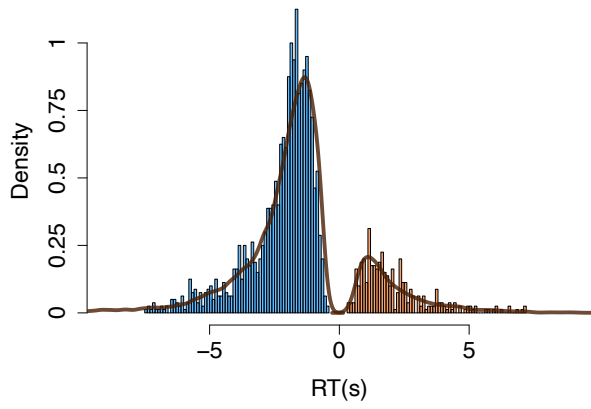
Food gamble
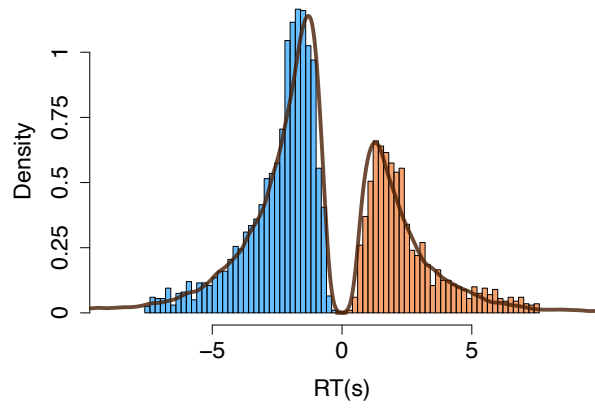
Money gamble
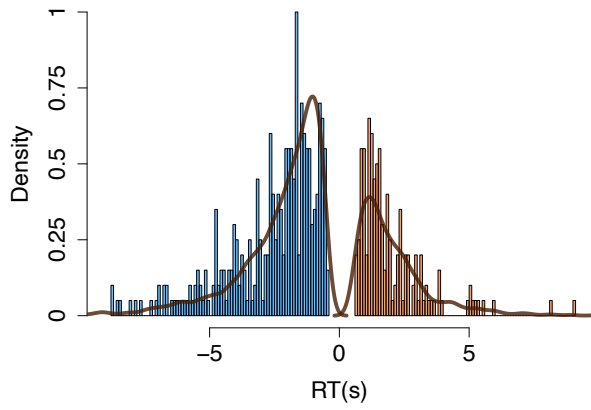
Money gamble
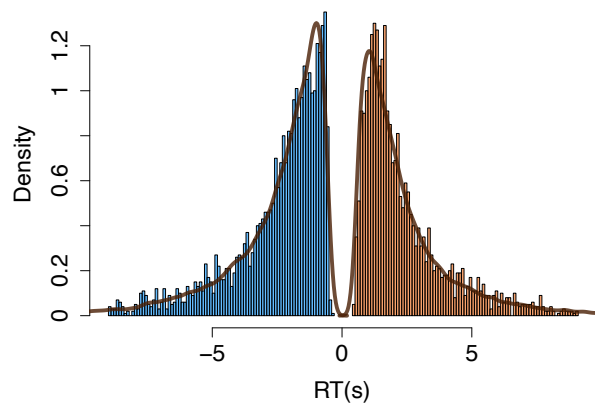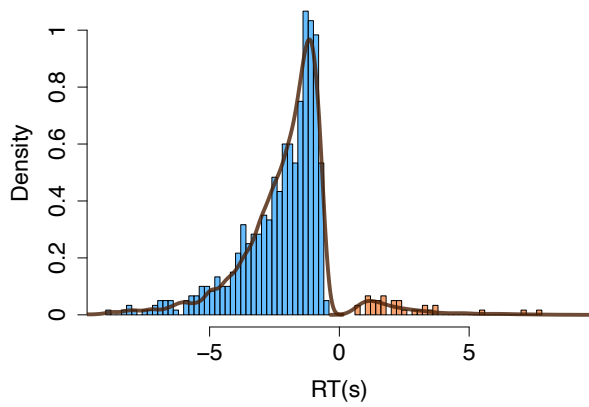
Compassion
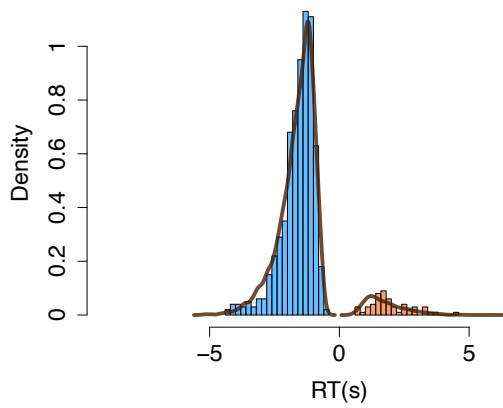
Compassion
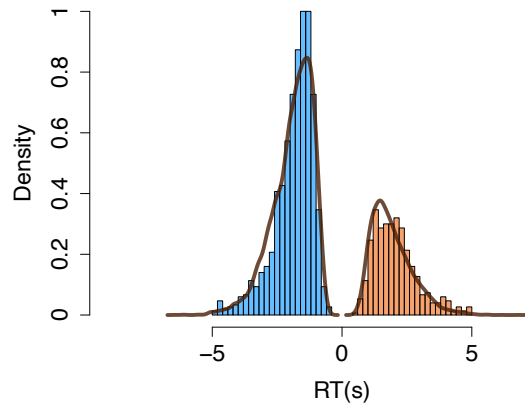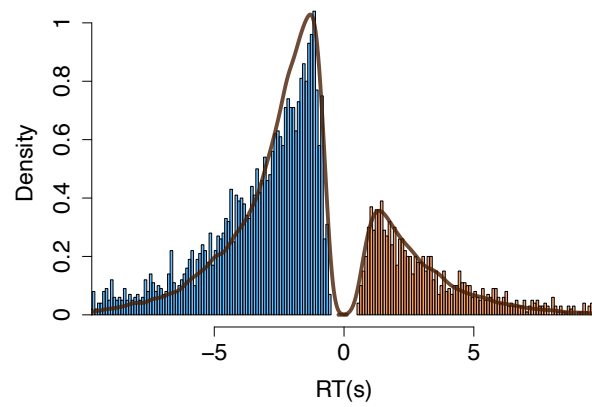
Envy
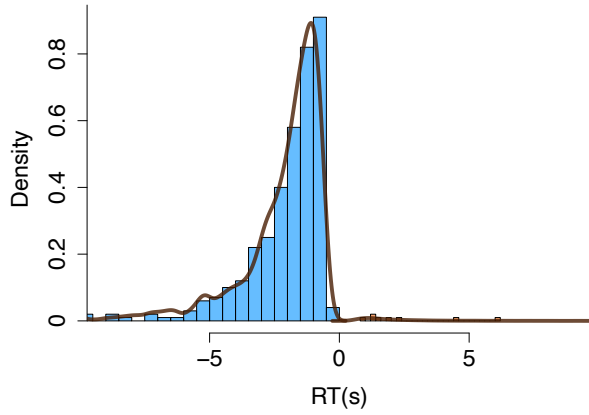
Envy
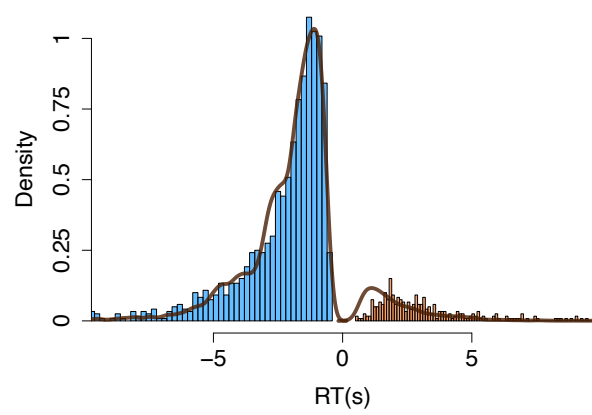
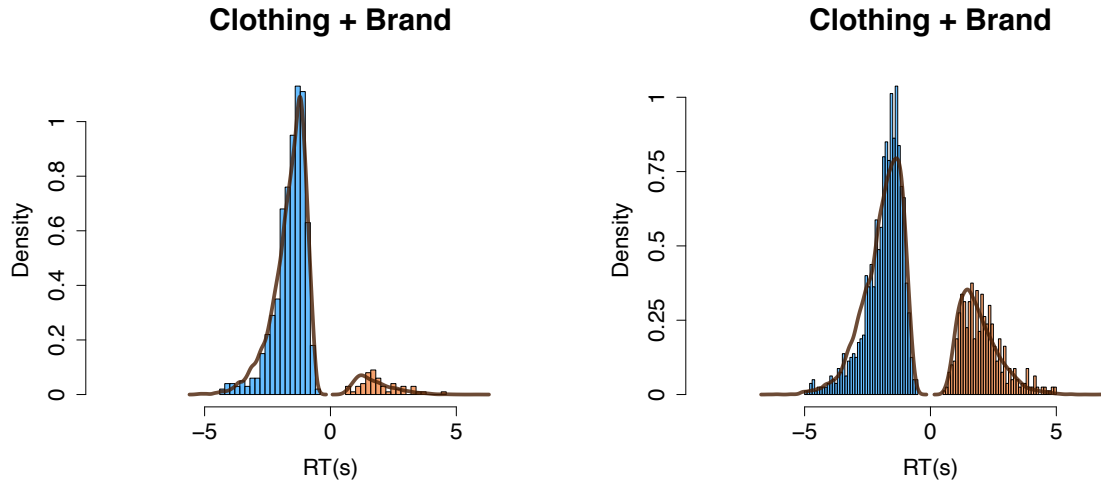**Food + GDA label**
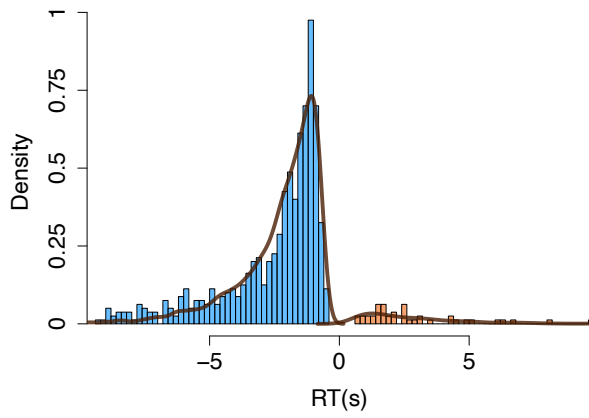
Food + GDA label
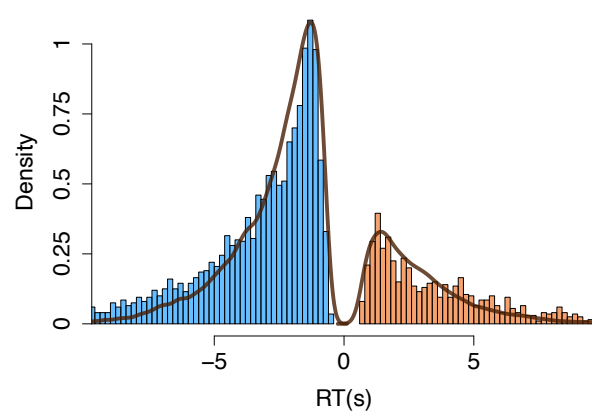
**Food + TL label**

Food + TL label

Supplementary Figure 8. The figures show the aggregated RT distributions across all the subjects in each task, split by high value/utility difference condition and low value/utility difference condition. Median absolute value differences are used to separate trials into high/low conditions. The densities are from the simulations of the best-fitting models and the histograms are from the aggregated empirical data. The RT distribution for the higher valued option is on the left (i.e., with a negative axis; blue), whereas the lower valued option is on the right (i.e., with a positive axis, orange).
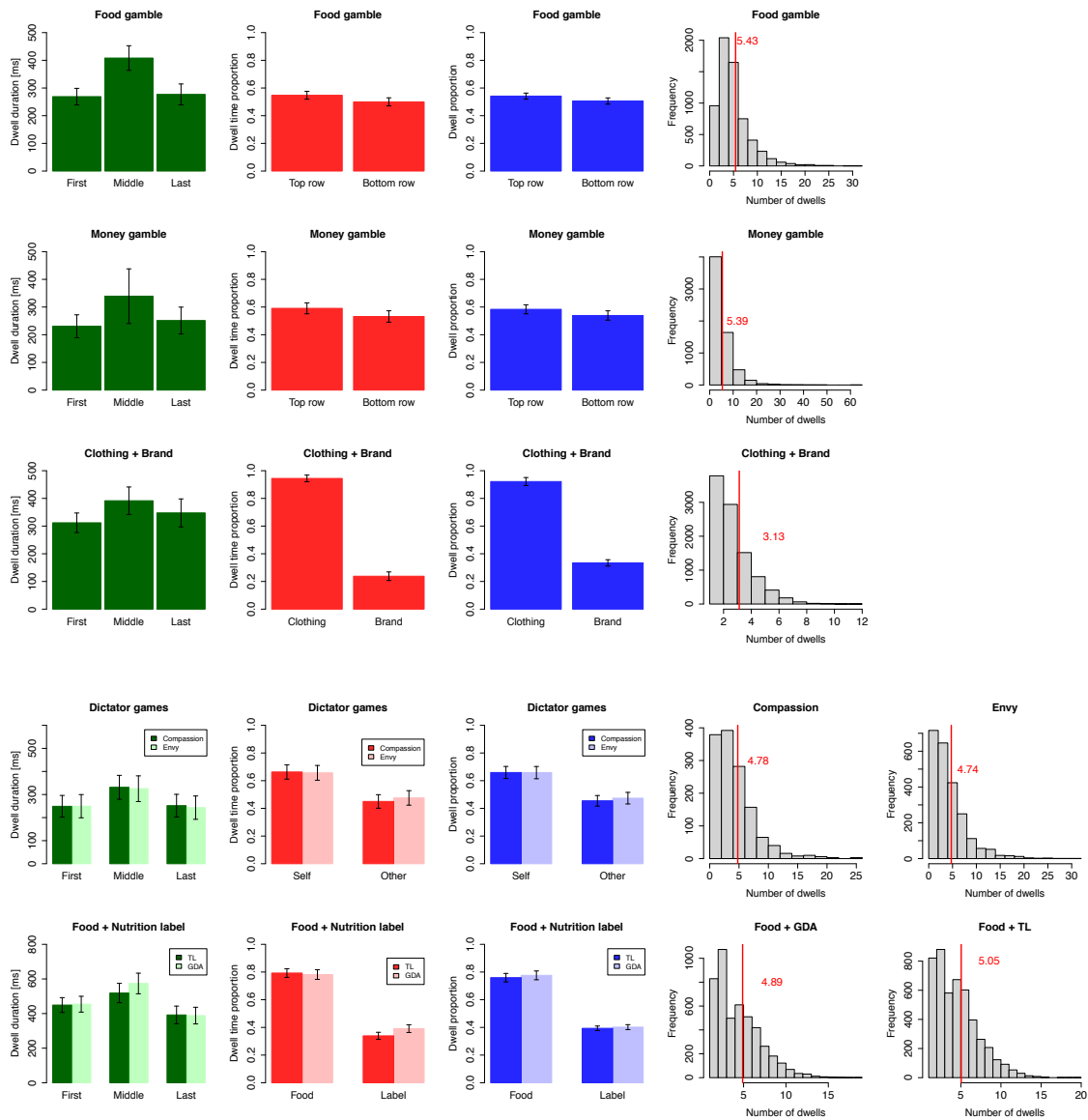
Supplementary Figure 9. Descriptive gaze statistics for each task. The first column displays mean dwell times conditional on being the first, last, or other (middle) dwell of the trial. The second column displays the fraction of total dwell time allocated to each of the two attributes. The third column displays the fraction of dwells allocated to each of the two attributes. The fourth (and fifth) column display histograms of the number of dwells in each trial. Error bars are standard errors clustered by participant. In the histograms, the red line/number indicates the mean of the distribution.
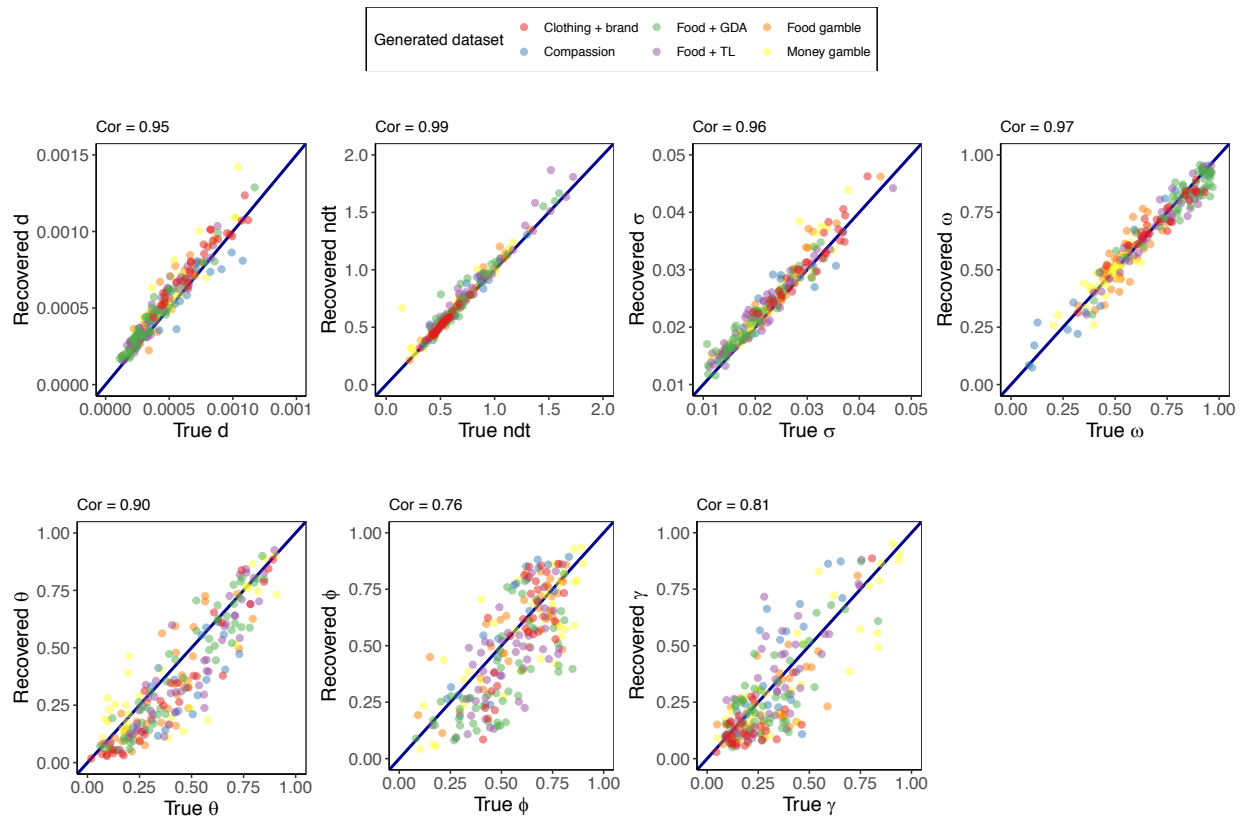
Supplementary Figure 10. Additional parameter recovery results. Each plot displays the correlation between the "true" simulated parameter value (x-axis) and the recovered parameter value (y-axis). Each dot represents a simulated subject and colors indicate different tasks.