

THIS MANUSCRIPT IS CURRENTLY UNDER REVIEW

Supplementary Materials

Formally Testing the Competitive Cue Claim in Person Perception Research:

Does the Use of Social Categories Diminish the Use of Other Cues?

Public repository of data and analysis code: <https://osf.io/sywmh>

Table of Contents

Data Exclusions	3
Exploring the Robustness of the Posterior Parameter Correlation	4
Excluding Correlations with $SD > 0.20$	4
Excluding Correlations with $SD > 0.15$	4
Excluding Correlations with $SD > 0.10$	4
Method: Novel Speeded Classification Experiments	6
Stimuli	6
Procedure	7
MCI Model Fit	10
Analytic Assessment	10
Visual Assessment	12
Parameter Estimates	27
References	41

Data Exclusions**Table S1***Participant Exclusions by Dataset*

Source	Studies	Method	Cues	Initial N	Exclusions	Final N
Axt et al. (2018)	1a–1d	Academic JBT	Individuating, Gender	4,377	385 (8.8%)	3,992
Axt et al. (2018)	2a–3	Academic JBT	Individuating, Group Identity	1,376	80 (5.8%)	1,296
Axt et al. (2018)	4	Dating JBT	Individuating, Race	1,100	120 (10.9%)	980
Current Research	1	Academic JBT	Individuating, Gender	1,892	17 (0.9%)	1,875
Current Research	1a	Speeded Gender Classification	Gender, Emotion	96	1 (1.0%)	95
Current Research	1b	Speeded Emotion Classification	Emotion, Gender	367	17 (4.6%)	350
Current Research	2a	Speeded Race Classification	Race, Emotion	208	6 (2.9%)	202
Current Research	2b	Speeded Emotion Classification	Emotion, Race	208	10 (4.8%)	198
Current Research	3a	Speeded Race Classification	Race, Gender	203	12 (5.9%)	191
Current Research	3b	Speeded Gender Classification	Gender, Race	261	17 (6.5%)	244
Current Research	4	Speeded Trustworthiness Classification	Emotion, Race	209	50 (23.9%)	159
Navon et al. (2021)	2–17	IAT	Individuating, Race	2,152	88 (4.1%)	2,064
Todd et al. (2021)	1, 3, 4	Age Classification	Age, Race	210	0 (0.0%)	210
Todd et al. (2021)	1; 3–6	Race Classification	Race, Age	383	0 (0.0%)	383

Note. The 'Exclusions' column shows the total number of participants excluded, with the corresponding percentage of the excluded initial sample in parentheses.

Exploring the Robustness of the Posterior Parameter Correlation

We re-ran the analysis reported in the main text three additional times, applying progressively stricter criteria for including datasets based on the precision of their parameter correlation estimates. Specifically, we excluded datasets with correlation standard deviations (*SDs*) greater than 0.20, 0.15, and 0.10. As detailed below, the posterior meta-analytic correlation remained non-credible, and substantial heterogeneity remained present, thereby supporting the conclusion from the main text that the relationship between cues is highly variable and not consistently competitive.

Excluding Correlations with *SD* > 0.20

After removing four datasets with correlation *SD* > .20, the analysis included $k = 10$ datasets. The resulting posterior meta-analytic correlation remained non-credibly different from zero ($r = -0.11$, $\text{BCI}_{95\%} [-0.43, 0.21]$). The Bayes Factor, $\text{BF}_{01} = 2.45$, suggests about two and a half times greater evidence in favor of a non-zero correlation. The posterior probability of a negative correlation was $P(r < 0) = .76$. Substantial heterogeneity remained across studies ($\tau = 0.51$, $\text{BCI}_{95\%} [0.33, 0.78]$, $I^2 > 99\%$).

Excluding Correlations with *SD* > 0.15

Applying a stricter cutoff, this analysis included $k = 8$ datasets. The meta-analytic was still not credibly different from zero ($r = -0.25$, $\text{BCI}_{95\%} [-0.57, 0.07]$). The Bayes Factor was inconclusive, offering only anecdotal evidence for an effect over the null, $\text{BF}_{10} = 1.25$. The posterior probability of a negative correlation was $P(r < 0) = .94$. As before, substantial heterogeneity was observed ($\tau = 0.45$, $\text{BCI}_{95\%} [0.27, 0.72]$, $I^2 > 99\%$).

Excluding Correlations with *SD* > 0.10

The most stringent analysis included only $k = 4$ datasets with the most precise correlation estimates. The result was once again a non-credible meta-analytic correlation ($r = -0.17$, $\text{BCI}_{95\%} [-0.67, 0.35]$). The Bayes Factor was inconclusive, $\text{BF}_{01} = 1.54$, and the posterior probability of a negative correlation was $P(r < 0) = .76$. Heterogeneity among this smaller set of studies remained very high ($\tau = 0.55$, $\text{BCI}_{95\%} [0.28, 1.04]$, $I^2 > 99\%$).

Method: Novel Speeded Classification Experiments

Stimuli

Experiments 1–7 relied on a stimulus set of faces selected from the Chicago Face Database (CFD; Ma et al., 2015).

Experiments 1a–1b (Speeded Gender & Emotion Classification, Respectively)

We relied on the same sample of 40 face images used by Klein and Sherman (2024). The stimuli comprised 10 female models and 10 male models from the CFD. Two face images were selected per model: one image in which they display a smile and another in which they display a scowl. We then applied facial morphing techniques, using the morphing tool *WebMorph* (Debruine, 2018), to manipulate ambiguity in the original stimulus set of 40 faces. Smiling and scowling face images, per model, were morphed together, such that each model was associated with two additional face images, besides their original smiling and scowling images: one image with 55% properties of their smiling image and 45% properties of their scowling image remained (i.e., an ambiguous smile), and another in which 45% properties of their smiling image and 55% properties of their scowling image remained (i.e., an ambiguous scowl). Female and male models were then randomly paired and morphed by their expression-matched faces images (e.g., both with unambiguous smiles). Morphs along gender generated ambiguous male (i.e., 45% female, 55% male) and ambiguous female (55% female, 45% male) face images. Because all image stimuli retained at least 55% properties of one level of gender and expression, each stimulus was associated with a “correct” classification. In total, this procedure produced a stimulus set of 160 face images. The same stimulus set size was generated for Experiments 2a–3b.

Experiments 2a–2b (Speeded Race & Emotion Classification, Respectively)

The same stimulus generation procedure was applied for Experiments 2a and 2b, with race instead of gender being the social category crossed with emotion expression: 10 models from each social category (Black, White) were selected, along with smiling and scowling face images per model, and their emotion and social category levels were morphed accordingly.

Experiment 3a–3b (Speeded Race & Gender Classification, Respectively)

The stimulus generation procedure for Experiments 3a and 3b was similar to that of Experiments 1a–2b. Ten face images were selected per race $\times$ gender level: 10 Black female faces, 10 Black male faces, 10 White female faces, and 10 White male faces. Face images were randomly paired between levels of race, producing two additional images: one image with 55% (45%) properties of the Black (White) face and another with 45% (55%) properties the Black (White) face. The same random pairing and morphing procedure was also applied along levels of gender. Finally, ambiguous race- and ambiguous gender-morphed images were paired (e.g., ambiguous Black face paired with ambiguous male face) and morphed together, generating face images that were simultaneously ambiguous along both race and gender.

Experiment 4 (Speeded Trustworthiness Classification)

Experiment 4 relied on the same 40 neutral, unmorphed face images as those selected for Experiments 2a–2b. Each upright image was then vertically inverted. The final stimulus set, therefore, comprised 80 face images, with two images (upright, inverted) per model.

Procedure

As previously mentioned, target faces varied along two sets of cues (e.g., gender and emotion cues in Experiments 1a and 1b). In each experiment, participants were instructed to “quickly and accurately” classify a series of target faces along a one of the two sets of cues (e.g., gender classification in Experiment 1a). Therefore, the two sets of cues varied in relevance, depending

on the classification study to which participants were assigned. In Experiment 1a (i.e., speeded gender classification), gender cues were relevant and emotion cues were irrelevant. In Experiment 1b (i.e., speeded emotion classification), the opposite was true.

On each trial, stimuli appeared in the following sequence: a fixation cross for 500 ms, a target face for 200 ms, and, finally, a pattern mask until a response was recorded. After a block of practice trials, ranging between 10 and 16 trials across the experiments, participants in all experiments completed 160 classification trials. In Experiments 1a–3b, there were three test blocks of 60, 60, and 40 trials, with each face image being presented once. In Experiment 4, there were four test blocks of 40 images, with each face image being presented twice. In all experiments, self-paced breaks were provided between blocks. Participants then reported demographic information before being debriefed. These experiments were approved by the ##### Institutional Review Board (protocol number: #####).

Table S2*Participant Demographics Per Dataset*

Source	Studies	Method	Age		Gender (%)			Race/Ethnicity (%)				
			$\bar{X}$	SD	M	F	NB	W	B	A	L	MR
Current Research	1	Academic JBT	32.1	14.7	31.4	63.8	4.9	55.6	12.6	10.5	12.3	9.0
Axt et al. (2018)	4	Dating JBT	21.5	3.3	35.2	64.2	0.6	—	—	—	—	—
Axt et al. (2018)	1a–1d	Academic JBT	35.2	14.2	36.7	62.4	0.9	69.9	6.1	10.0	0.1	13.9
Axt et al. (2018)	2a–3	Academic JBT	18.8	1.8	32.2	67.5	0.3	64.3	9.0	16.7	3.9	6.1
Current Research	1a	Speeded Gender Classification	20.1	2.9	19.1	79.8	1.1	17.0	2.1	48.9	20.2	11.7
Current Research	1b	Speeded Emotion Classification	20.0	2.3	17.7	80.0	2.3	17.1	1.4	52.6	12.9	16.0
Current Research	2a	Speeded Race Classification	19.8	3.3	28.2	65.3	6.4	18.8	2.0	49.5	19.3	10.4
Current Research	2b	Speeded Emotion Classification	19.6	1.9	20.7	75.3	4.0	15.7	1.0	55.6	16.2	11.6
Current Research	3a	Speeded Race Classification	20.2	2.3	26.7	71.7	1.6	17.3	3.7	52.4	18.3	8.4
Current Research	3b	Speeded Gender Classification	20.0	1.9	21.3	72.5	6.1	15.2	3.7	50.8	18.0	12.3
Current Research	4	Speeded Trustworthiness Classification	19.7	2.9	13.8	77.9	8.1	7.5	1.3	47.8	22.0	19.4
Navon et al. (2021)	2–17	IAT	34.1	14.3	35.8	62.4	1.8	67.8	7.0	9.9	—	15.2
Todd et al. (2021)	1, 3, 4	Age Classification	—	—	26.0	73.0	1.0	58.7	2.6	28.1	5.6	5.1
Todd et al. (2021)	1; 3–6	Race Classification	—	—	37.2	62.0	0.8	70.4	2.5	17.2	6.2	3.7

Note. M = Male, F = Female, NB = Nonbinary, W = White or European American, B = Black or African American, A = Asian American or Pacific Islander, L = Latinx or Hispanic, and MR = reported other or more than one race/ethnicity. Some participants did not report their gender or race/ethnicity.

MCI Model Fit

Analytic Assessment

Model fit, the ability for the model to predict observed response data, was partly assessed via posterior predictive checks (PPCs). That is, we assessed the divergence between the observed response data and those predicted by the model's posterior distribution. Analytically, we conducted PPCs via two fit statistics. First, T_1 indicates the model's ability to account for discrepancies between observed and expected responses across all participants, providing an assessment of mean-level fit. Second, T_2 indicates the model's ability to account for discrepancies between observed and expected covariances, providing an assessment of individual-level fit. As stated in the main text, we conducted posterior predictive checks using 1,000 samples from each of the fitted models' posterior distribution, and the means and covariances of the posterior were then compared to those in the observed data to produce T_1 (mean) and T_2 (covariance) statistics. Both T_1 and T_2 p -values reflect the probability that discrepancies between the expected and posterior-predicted data are greater than those between the expected and observed data (Heck et al., 2018, p. 273). The T_1 statistic is defined by a chi-square distribution and, therefore, must be considered vis-à-vis the size of our dataset. The larger the dataset, the more sensitive the statistic will be at identifying even minor deviations of the model from the data. With $N_{\text{obs}} > 10,000$ observations per dataset, we expect a high level of sensitivity. Therefore, using the w -statistic, we describe the degree of misfit between the model's expectations and the observed aggregate responses. We rely on $w < .10$ as a descriptive benchmark for small degrees of misfit.

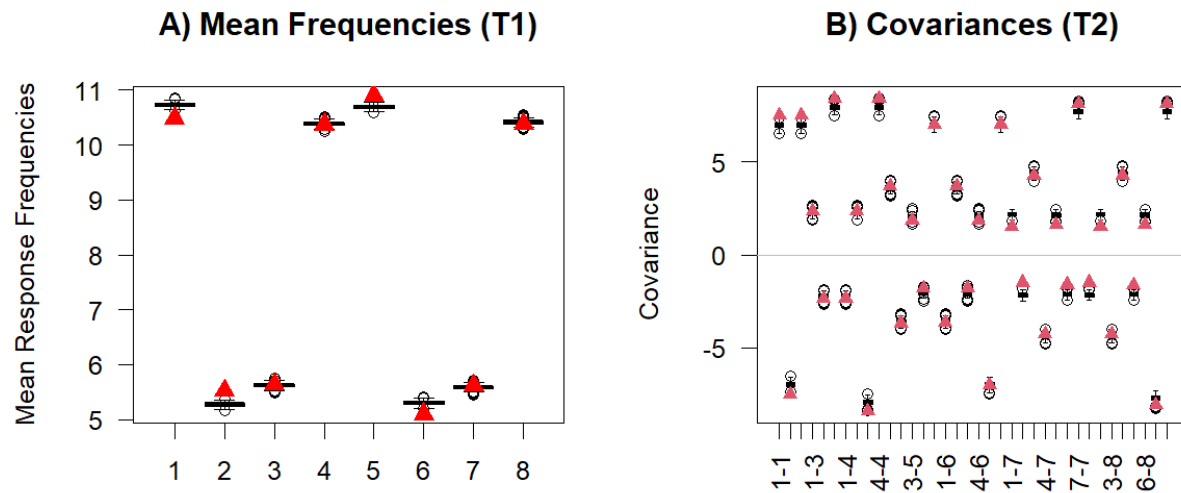
Table S2*Model Fit Statistics by Dataset*

Source	Studies	Method	Cues	T1 _{p(ind)}	T1 _p	T2 _p	w
Axt et al. (2018)	1a–1d	Academic JBT	Individuating, Gender	0.507	< .001	< .001	0.063
Axt et al. (2018)	2a–3	Academic JBT	Individuating, Group Identity	0.517	0.097	0.138	0.006
Axt et al. (2018)	4	Dating JBT	Individuating, Race	0.467	< .001	< .001	0.035
Current Research	1	Academic JBT	Individuating, Gender	0.499	0.249	< .001	0.002
Current Research	1a	Speeded Gender Classification	Gender, Emotion	0.526	0.042	0.071	0.009
Current Research	1b	Speeded Emotion Classification	Emotion, Gender	0.528	0.024	0.054	0.005
Current Research	2a	Speeded Race Classification	Race, Emotion	0.464	< .001	0.072	0.053
Current Research	2b	Speeded Emotion Classification	Emotion, Race	0.473	< .001	0.043	0.054
Current Research	3a	Speeded Race Classification	Race, Gender	0.559	< .001	0.032	0.030
Current Research	3b	Speeded Gender Classification	Gender, Race	0.564	0.376	0.001	0.001
Current Research	4	Speeded Trustworthiness Classification	Emotion, Race	0.397	< .001	< .001	0.068
Navon et al. (2021)	2–17	IAT	Individuating, Race	0.389	0.004	0.007	0.010
Todd et al. (2021)	1, 3, 4	Age Classification	Age, Race	0.364	0.002	< .001	0.031
Todd et al. (2021)	1; 3–6	Race Classification	Race, Age	0.346	< .001	< .001	0.192

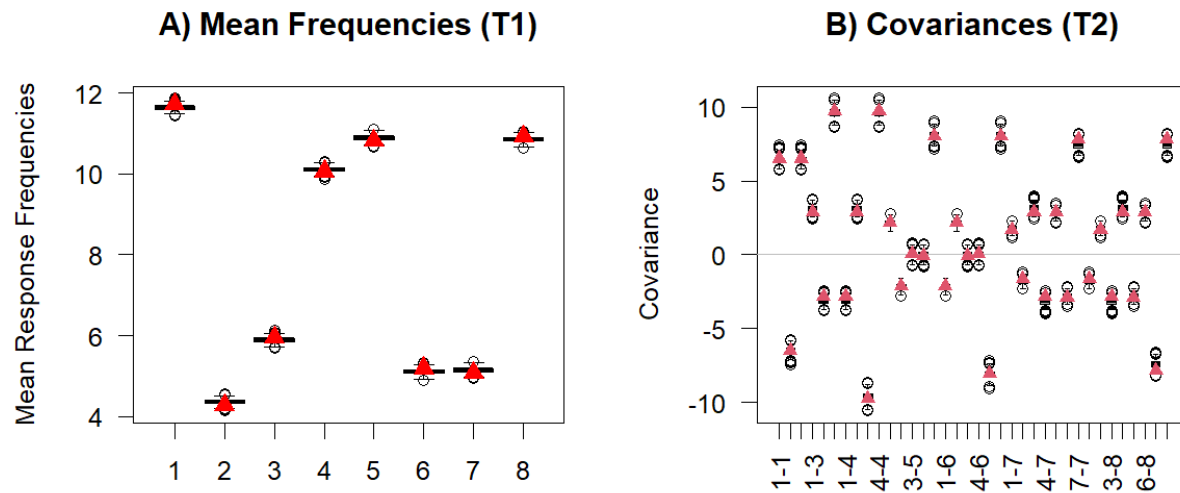
Note. T1_{p(ind)} is the median of the individual-level posterior predictive *p*-values, reflecting how well the model accounts for the response frequencies of a typical participant. The T1_p is the *p*-value for the aggregate T1 statistic, which assesses misfit between the model's predicted and the observed aggregated response frequencies (i.e., mean-level fit). The T2_p is the *p*-value for the T2 statistic, which assesses misfit for the response covariances, providing an index of fit to individual-level response patterns. The *w*-statistic is an effect size measure of misfit for the aggregate T1 statistic, with values < .10 indicating a small degree of misfit. Model fit statistics for speeded classification data come from models accounting for within-participant factors.

Visual Assessment

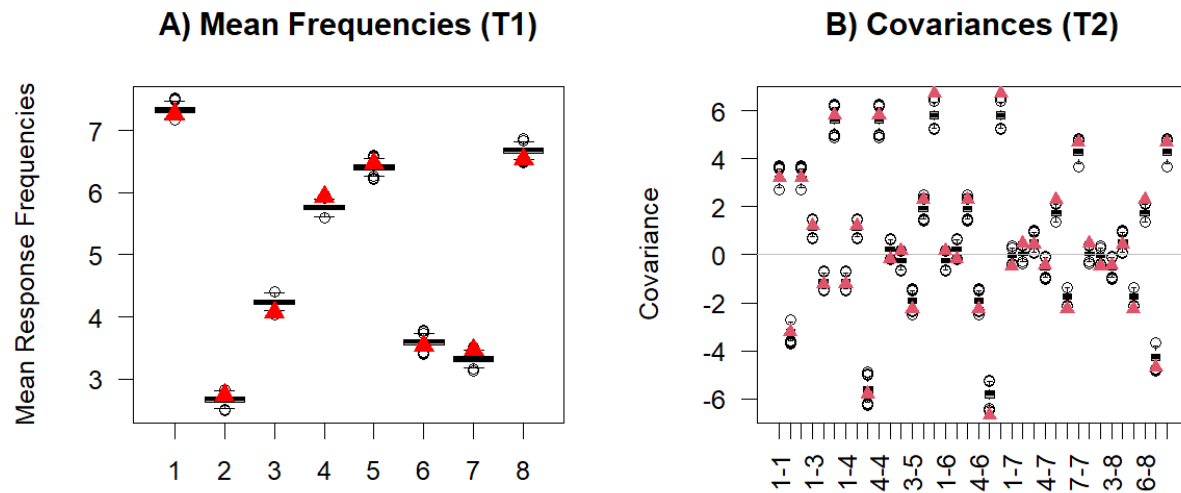
As discussed above, analytic assessment of model fit can be limited under instances of large data, which certainly reflects the current research. Therefore, we also examined the model's fit by plotting the observed and predicted response patterns and visually inspecting the divergence. Specifically, we plotted predicted versus observed mean frequencies (T_1) and covariances (T_2). The observed means and covariances generally fall within the range of box-plotted model expectations, indicating good fit. Visual inspection is both inherently Bayesian and a common tool for MPT modeling using the current estimation strategy (e.g., Calanchini et al., 2021; Smith et al., 2020). Visualizations of model fit for speeded classification data come from models accounting for within-participant factors.

Figure S1*Posterior predictive check for Studies 1a–1d of Axt et al. (2018)*

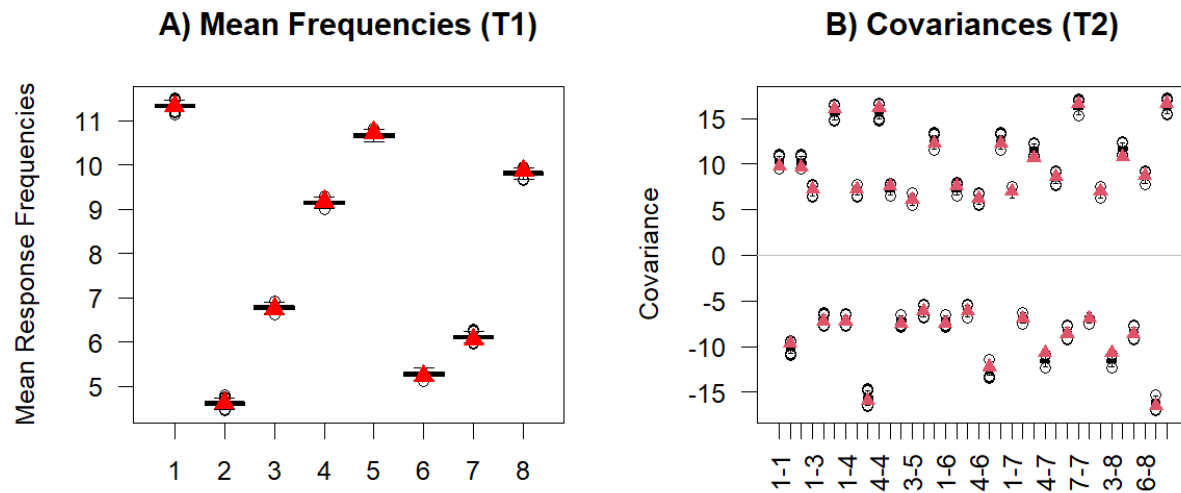
Note. Observed data are shown as red triangles; model-predicted distributions are shown as boxplots. Panel A shows the fit to mean response frequencies (T1). Panel B shows the fit to response covariances (T2).

Figure S2*Posterior predictive check for Studies 2a–3 of Axt et al. (2018)*

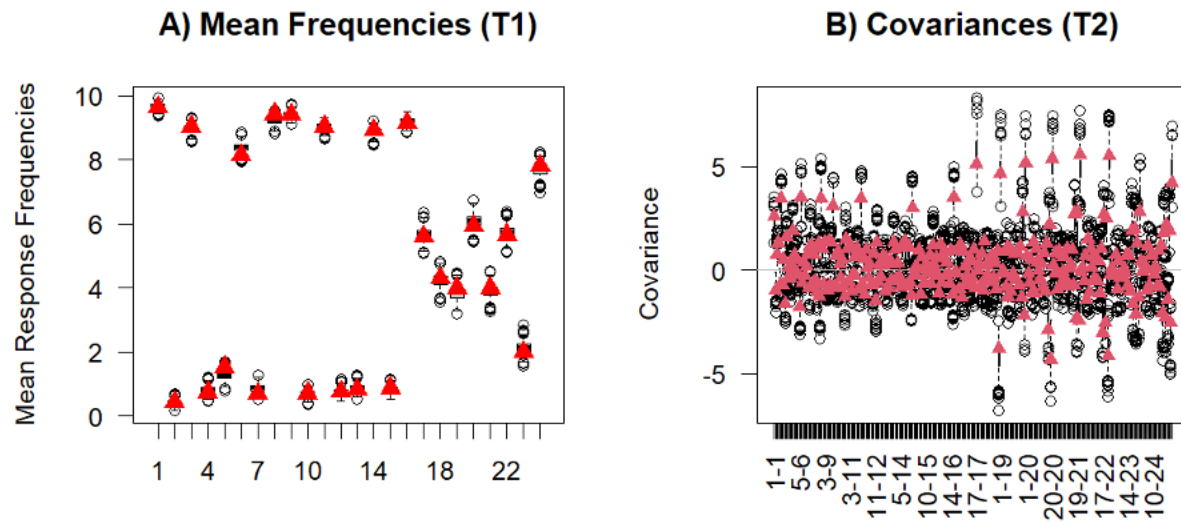
Note. Observed data are shown as red triangles; model-predicted distributions are shown as boxplots. Panel A shows the fit to mean response frequencies (T1). Panel B shows the fit to response covariances (T2).

Figure S3*Posterior predictive check for Study 4 of Axt et al. (2018)*

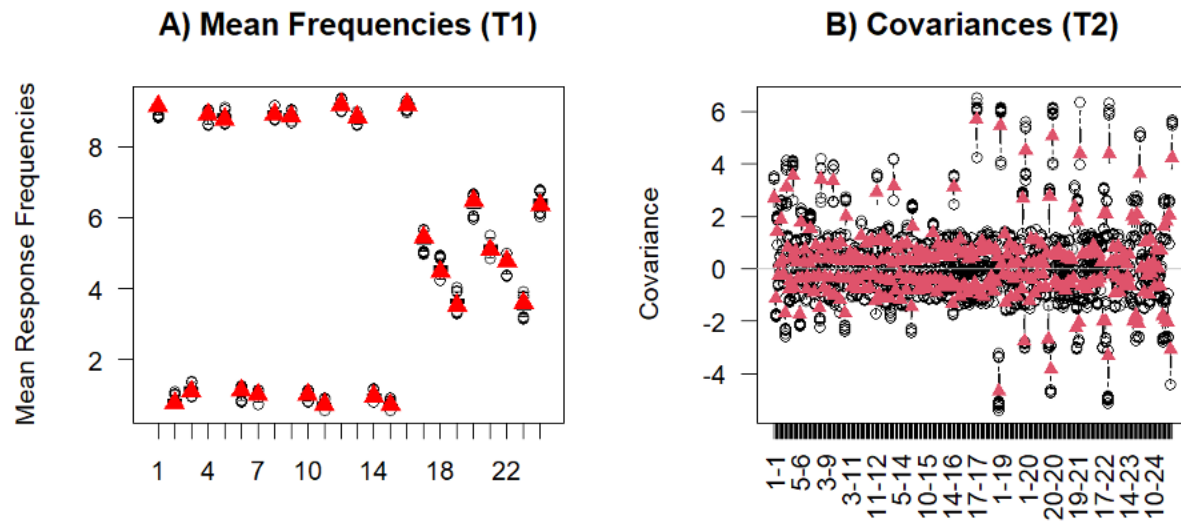
Note. Observed data are shown as red triangles; model-predicted distributions are shown as boxplots. Panel A shows the fit to mean response frequencies (T1). Panel B shows the fit to response covariances (T2).

Figure S4*Posterior predictive check for Judgment Bias Task Data of Current Research*

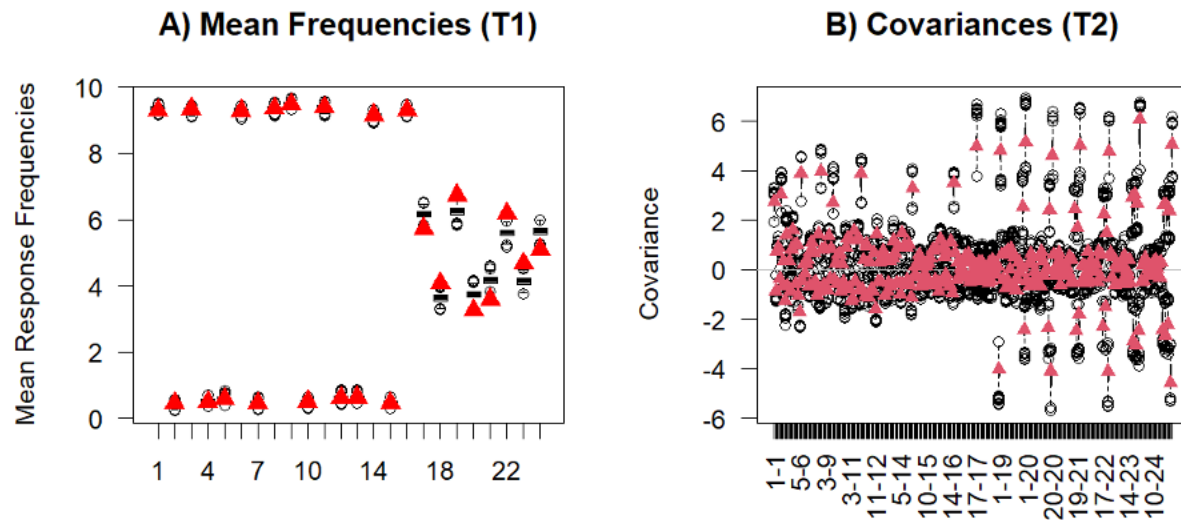
Note. Observed data are shown as red triangles; model-predicted distributions are shown as boxplots. Panel A shows the fit to mean response frequencies (T1). Panel B shows the fit to response covariances (T2).

Figure S5*Posterior predictive check for Experiment 1a of Current Research*

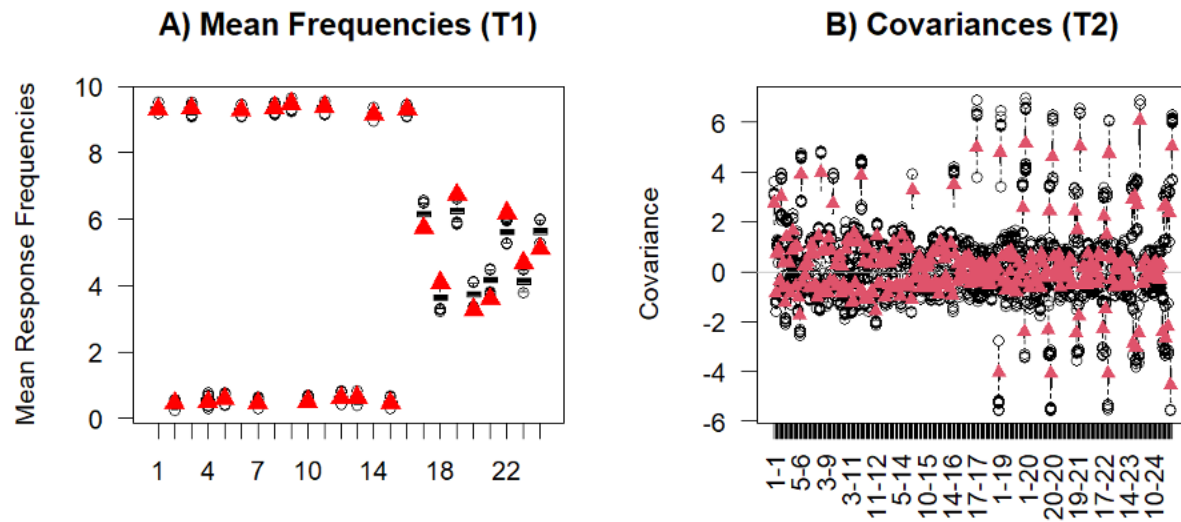
Note. Observed data are shown as red triangles; model-predicted distributions are shown as boxplots. Panel A shows the fit to mean response frequencies (T1). Panel B shows the fit to response covariances (T2).

Figure S6*Posterior predictive check for Experiment 1b of Current Research*

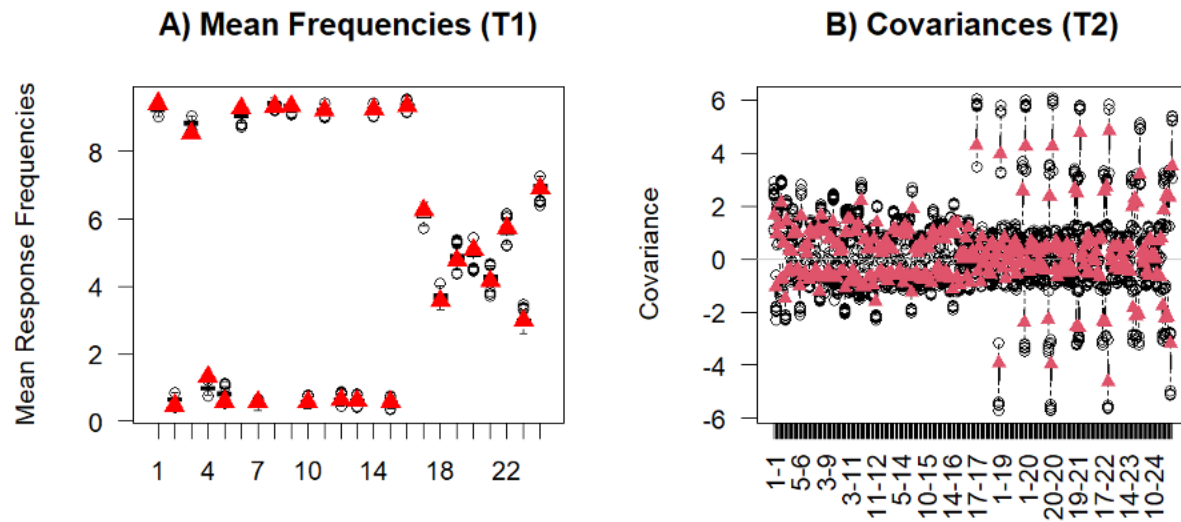
Note. Observed data are shown as red triangles; model-predicted distributions are shown as boxplots. Panel A shows the fit to mean response frequencies (T1). Panel B shows the fit to response covariances (T2).

Figure S7*Posterior predictive check for Experiment 2a of Current Research*

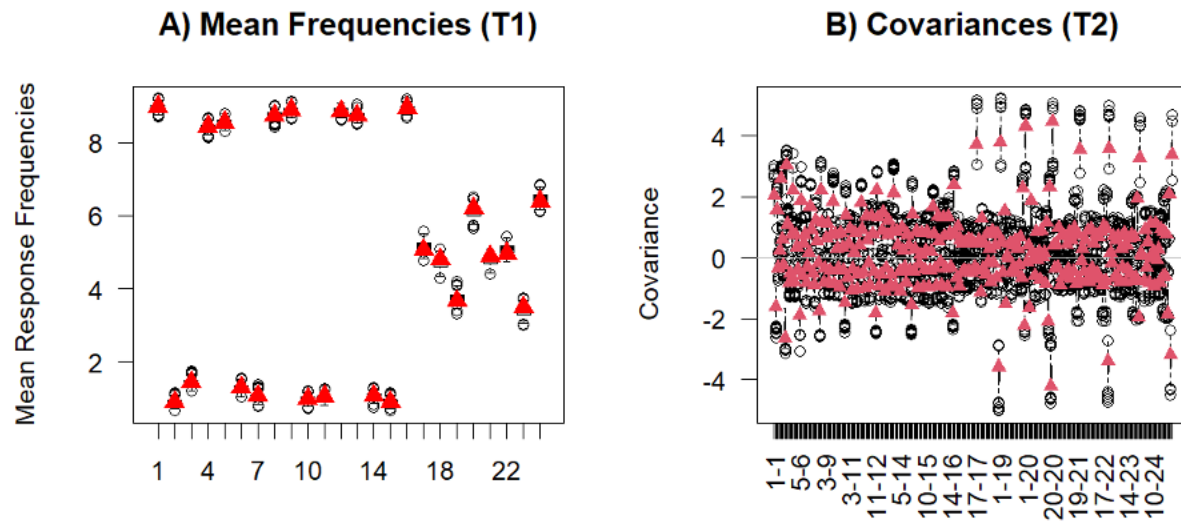
Note. Observed data are shown as red triangles; model-predicted distributions are shown as boxplots. Panel A shows the fit to mean response frequencies (T1). Panel B shows the fit to response covariances (T2).

Figure S8*Posterior predictive check for Experiment 2b of Current Research*

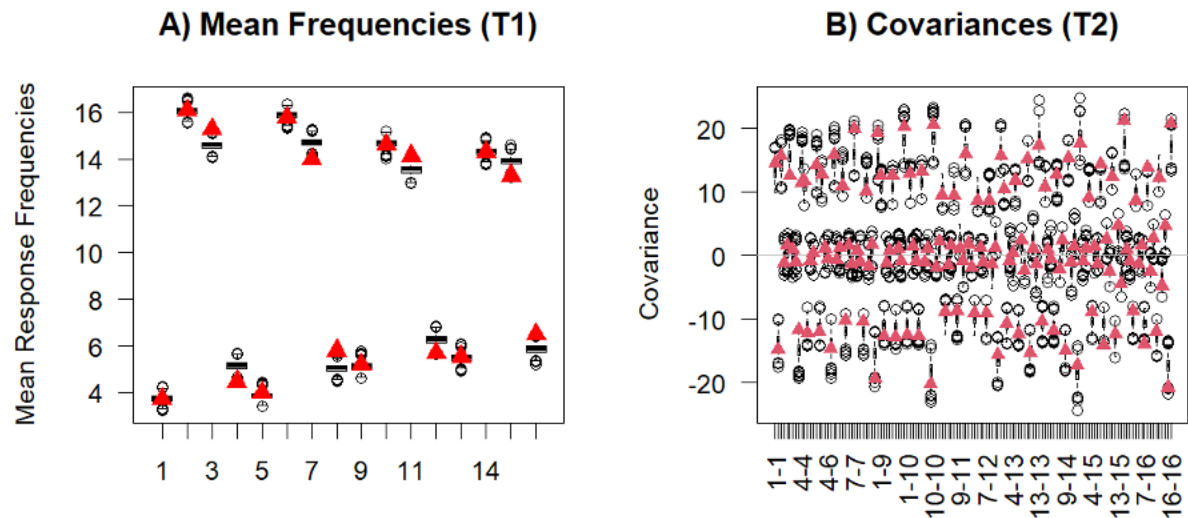
Note. Observed data are shown as red triangles; model-predicted distributions are shown as boxplots. Panel A shows the fit to mean response frequencies (T1). Panel B shows the fit to response covariances (T2).

Figure S9*Posterior predictive check for Experiment 3a of Current Research*

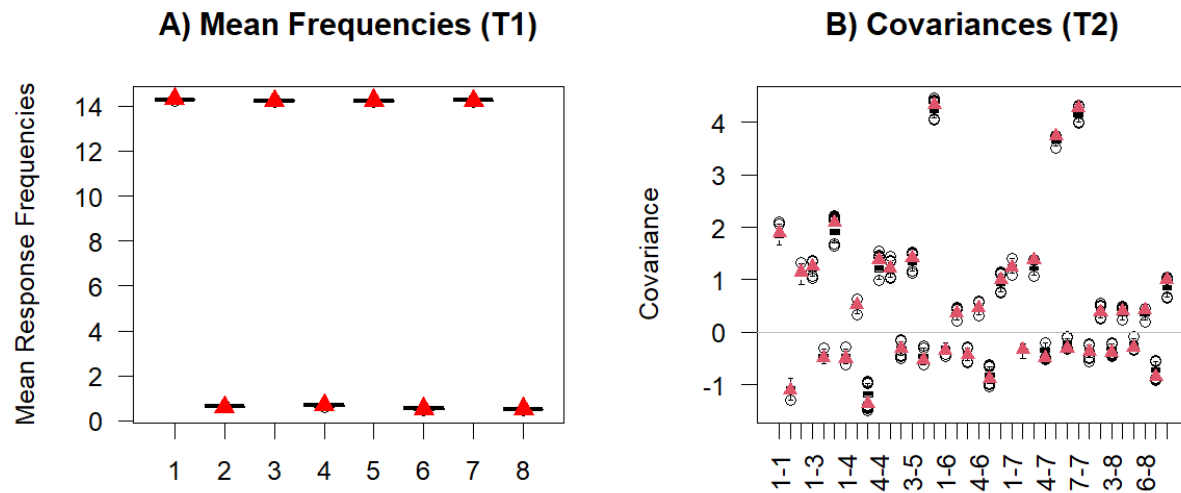
Note. Observed data are shown as red triangles; model-predicted distributions are shown as boxplots. Panel A shows the fit to mean response frequencies (T1). Panel B shows the fit to response covariances (T2).

Figure S10*Posterior predictive check for Experiment 3b of Current Research*

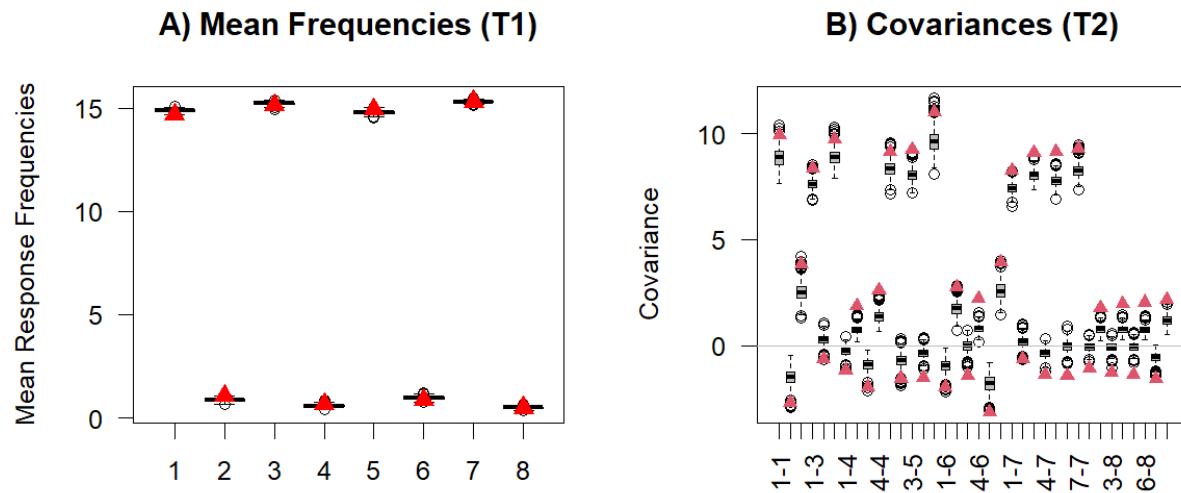
Note. Observed data are shown as red triangles; model-predicted distributions are shown as boxplots. Panel A shows the fit to mean response frequencies (T1). Panel B shows the fit to response covariances (T2).

Figure S11*Posterior predictive check for Experiment 4 of Current Research*

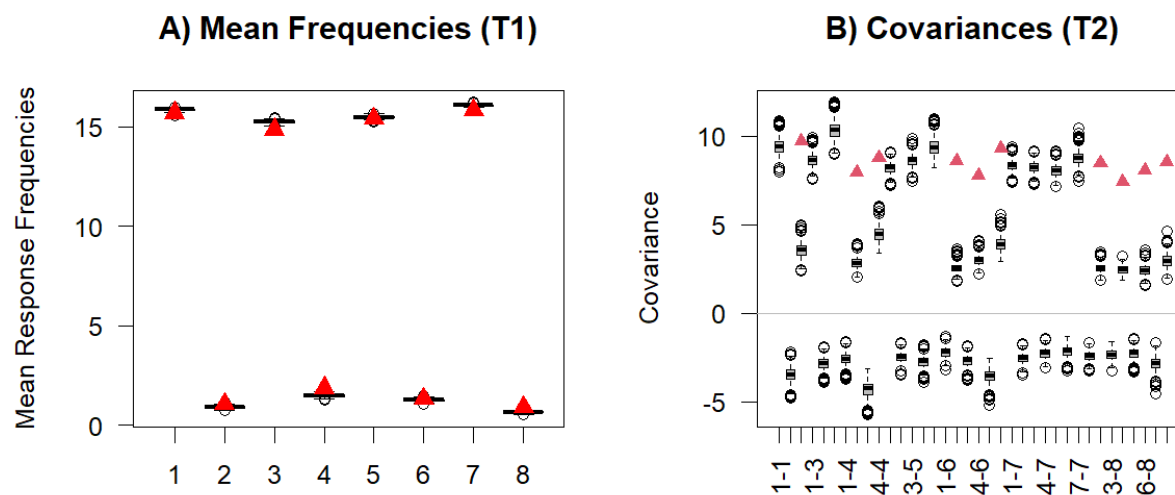
Note. Observed data are shown as red triangles; model-predicted distributions are shown as boxplots. Panel A shows the fit to mean response frequencies (T1). Panel B shows the fit to response covariances (T2).

Figure S12*Posterior predictive check for Studies 2–17 of Navon et al. (2021)*

Note. Observed data are shown as red triangles; model-predicted distributions are shown as boxplots. Panel A shows the fit to mean response frequencies (T1). Panel B shows the fit to response covariances (T2).

Figure S13*Posterior predictive check for Studies 1, 3, 4 of Todd et al. (2021)*

Note. Observed data are shown as red triangles; model-predicted distributions are shown as boxplots. Panel A shows the fit to mean response frequencies (T1). Panel B shows the fit to response covariances (T2).

Figure S14*Posterior predictive check for Studies 1, 3–6 of Todd et al. (2021)*

Note. Observed data are shown as red triangles; model-predicted distributions are shown as boxplots. Panel A shows the fit to mean response frequencies (T1). Panel B shows the fit to response covariances (T2).

Parameter Estimates**Table S3***Parameter Estimates for Axt et al. (2018), Studies 1a–1d*

Parameter	μ	95% BCI Lower	95% BCI Upper
Use of Individuating (C_1)	0.328	0.320	0.335
Use of Gender (C_2)	<0.000	<0.000	<0.000
Response Bias (g)	0.513	0.506	0.519

Note. C_1 reflects the use of individuating cues and C_2 reflects the use of gender cues. The g parameter reflects a response bias. Values represent the posterior mean (μ) and the 95% Bayesian Credibility Interval (BCI).

Table S4*Parameter Estimates for Axt et al. (2018), Studies 2a–3*

Parameter	μ	95% BCI Lower	95% BCI Upper
Use of Individuating (C_1)	0.363	0.351	0.377
Use of Group Identity (C_2)	0.002	<0.000	0.006
Response Bias (g)	0.545	0.533	0.557

Note. C_1 reflects the use of individuating cues and C_2 reflects the use of group identity cues. The g parameter reflects a response bias. Values represent the posterior mean (μ) and the 95% Bayesian Credibility Interval (BCI).

Table S5*Parameter Estimates for Axt et al. (2018), Study 4*

Parameter	μ	95% BCI Lower	95% BCI Upper
Use of Individuating (C_1)	0.284	0.269	0.300
Use of Race (C_2)	0.001	<0.000	0.002
Response Bias (g)	0.546	0.532	0.560

Note. C_1 reflects the use of individuating cues and C_2 reflects the use of race cues. The g parameter reflects a response bias. Values represent the posterior mean (μ) and the 95% Bayesian Credibility Interval (BCI).

Table S6*Parameter Estimates for JBT Data From Current Research*

Parameter	μ	95% BCI Lower	95% BCI Upper
Use of Individuating (C_1)	0.258	0.249	0.268
Use of Gender (C_2)	0.004	0.001	0.008
Response Bias (g)	0.576	0.560	0.594

Note. C_1 reflects the use of individuating cues and C_2 reflects the use of gender cues. The g parameter reflects a response bias. Values represent the posterior mean (μ) and the 95% Bayesian Credibility Interval (BCI).

Table S7*Parameter Estimates for Experiment 1a (Speeded Emotion Classification)*

Parameter	Condition	μ	95% BCI Lower	95% BCI Upper
Use of Emotion (C₁)	Baseline (Unambiguous)	0.820	0.802	0.838
	Gender Ambiguity	0.863	0.845	0.881
	Emotion Ambiguity	0.168	0.151	0.186
Use of Gender (C₂)	Baseline (Unambiguous)	0.100	0.038	0.152
	Gender Ambiguity	0.017	<0.000	0.056
	Emotion Ambiguity	0.011	<0.000	0.027
Response Bias (g)	Baseline (Unambiguous)	0.455	0.407	0.504
	Gender Ambiguity	0.579	0.532	0.626
	Emotion Ambiguity	0.569	0.547	0.592

Note. C₁ reflects the use of emotion cues and C₂ reflects the use of gender cues. The g parameter reflects a response bias. The 'Baseline' condition used unambiguous cues. The 'Gender Ambiguity' and 'Emotion Ambiguity' conditions used morphed faces to introduce ambiguity in the respective cue. Values represent the posterior mean (μ) and the 95% Bayesian Credibility Interval (BCI).

Table S8*Parameter Estimates for Experiment 1b (Speeded Gender Classification)*

Parameter	Condition	μ	95% BCI Lower	95% BCI Upper
Use of Gender (C₁)	Baseline (Unambiguous)	0.853	0.822	0.882
	Emotion Ambiguity	0.871	0.840	0.899
	Gender Ambiguity	0.214	0.173	0.253
Use of Emotion (C₂)	Baseline (Unambiguous)	0.245	0.064	0.417
	Emotion Ambiguity	0.038	0.001	0.119
	Gender Ambiguity	0.164	0.134	0.194
Response Bias (g)	Baseline (Unambiguous)	0.303	0.168	0.450
	Emotion Ambiguity	0.437	0.310	0.567
	Gender Ambiguity	0.691	0.631	0.752

Note. C₁ reflects the use of gender cues and C₂ reflects the use of emotion cues. The g parameter reflects a response bias. The 'Baseline' condition used unambiguous cues. The 'Emotion Ambiguity' and 'Gender Ambiguity' conditions used morphed faces to introduce ambiguity in the respective cue. Values represent the posterior mean (μ) and the 95% Bayesian Credibility Interval (BCI).

Table S9*Parameter Estimates for Experiment 2a (Speeded Race Classification)*

Parameter	Condition	μ	95% BCI Lower	95% BCI Upper
Use of Race (C₁)	Baseline (Unambiguous)	0.936	0.919	0.951
	Emotion Ambiguity	0.921	0.904	0.937
	Race Ambiguity	0.200	0.177	0.223
Use of Emotion (C₂)	Baseline (Unambiguous)	0.007	<0.000	0.038
	Emotion Ambiguity	0.116	0.025	0.211
	Race Ambiguity	<0.000	0.000	0.002
Response Bias (g)	Baseline (Unambiguous)	0.495	0.415	0.576
	Emotion Ambiguity	0.478	0.390	0.568
	Race Ambiguity	0.474	0.436	0.510

Note. C₁ reflects the use of race cues and C₂ reflects the use of emotion cues. The g parameter reflects a response bias. The 'Baseline' condition used unambiguous cues. The 'Emotion Ambiguity' and 'Race Ambiguity' conditions used morphed faces to introduce ambiguity in the respective cue. Values represent the posterior mean (μ) and the 95% Bayesian Credibility Interval (BCI).

Table S10*Parameter Estimates for Experiment 2b (Speeded Emotion Classification)*

Parameter	Condition	μ	95% BCI Lower	95% BCI Upper
Use of Emotion (C₁)	Baseline (Unambiguous)	0.936	0.919	0.952
	Race Ambiguity	0.930	0.911	0.948
	Emotion Ambiguity	0.200	0.176	0.223
Use of Race (C₂)	Baseline (Unambiguous)	0.001	<0.000	0.004
	Race Ambiguity	0.007	0.001	0.016
	Emotion Ambiguity	<0.000	<0.000	0.002
Response Bias (g)	Baseline (Unambiguous)	0.493	0.412	0.576
	Race Ambiguity	0.485	0.399	0.574
	Emotion Ambiguity	0.475	0.438	0.512

Note. C₁ reflects the use of emotion cues and C₂ reflects the use of race cues. The g parameter reflects a response bias. The 'Baseline' condition used unambiguous cues. The 'Race Ambiguity' and 'Emotion Ambiguity' conditions used morphed faces to introduce ambiguity in the respective cue. Values represent the posterior mean (μ) and the 95% Bayesian Credibility Interval (BCI).

Table S11*Parameter Estimates for Experiment 3a (Speeded Gender Classification)*

Parameter	Condition	μ	95% BCI Lower	95% BCI Upper
Use of Gender (C₁)	Baseline (Unambiguous)	0.784	0.757	0.807
	Race Ambiguity	0.826	0.802	0.848
	Gender Ambiguity	0.139	0.119	0.158
Use of Race (C₂)	Baseline (Unambiguous)	0.093	0.021	0.179
	Race Ambiguity	0.039	0.002	0.095
	Gender Ambiguity	0.009	0.001	0.024
Response Bias (g)	Baseline (Unambiguous)	0.436	0.378	0.489
	Race Ambiguity	0.533	0.479	0.587
	Gender Ambiguity	0.582	0.561	0.604

Note. C₁ reflects the use of gender cues and C₂ reflects the use of race cues. The g parameter reflects a response bias. The 'Baseline' condition used unambiguous cues. The 'Race Ambiguity' and 'Gender Ambiguity' conditions used morphed faces to introduce ambiguity in the respective cue. Values represent the posterior mean (μ) and the 95% Bayesian Credibility Interval (BCI).

Table S12*Parameter Estimates for Experiment 3b (Speeded Race Classification)*

Parameter	Condition	μ	95% BCI Lower	95% BCI Upper
Use of Race (C₁)	Baseline (Unambiguous)	0.881	0.860	0.901
	Gender Ambiguity	0.924	0.906	0.941
	Race Ambiguity	0.192	0.169	0.214
Use of Gender (C₂)	Baseline (Unambiguous)	0.324	0.160	0.458
	Gender Ambiguity	0.052	0.002	0.134
	Race Ambiguity	0.158	0.129	0.185
Response Bias (g)	Baseline (Unambiguous)	0.626	0.526	0.723
	Gender Ambiguity	0.503	0.413	0.589
	Race Ambiguity	0.557	0.511	0.597

Note. C₁ reflects the use of race cues and C₂ reflects the use of gender cues. The g parameter reflects a response bias. The 'Baseline' condition used unambiguous cues. The 'Gender Ambiguity' and 'Race Ambiguity' conditions used morphed faces to introduce ambiguity in the respective cue. Values represent the posterior mean (μ) and the 95% Bayesian Credibility Interval (BCI).

Table S13*Parameter Estimates for Experiment 4 (Speeded Trustworthiness Classification)*

Parameter	Condition	μ	95% BCI Lower	95% BCI Upper
Use of Emotion (C₁)	Upright	0.563	0.517	0.605
	Inverted	0.406	0.356	0.452
Use of Race (C₂)	Upright	0.005	<0.000	0.018
	Inverted	0.005	<0.000	0.018
Response Bias (g)	Upright	0.401	0.343	0.459
	Inverted	0.441	0.382	0.499

Note. C₁ reflects the use of emotion cues and C₂ reflects the use of race cues. The g parameter reflects a response bias. The 'Upright' condition refers to standard face presentation, whereas 'Inverted' refers to faces being flipped upside down. Values represent the posterior mean (μ) and the 95% Bayesian Credibility Interval (BCI).

Table S14*Parameter Estimates for Navon et al. (2021), Studies 2–17*

Parameter	μ	95% BCI Lower	95% BCI Upper
Use of Individuating (C_1)	0.957	0.953	0.961
Use of Race (C_2)	0.001	<0.000	0.002
Response Bias (g)	0.519	0.491	0.547

Note. C_1 reflects the use of individuating cues and C_2 reflects the use of race cues. The g parameter reflects a response bias. Values represent the posterior mean (μ) and the 95% Bayesian Credibility Interval (BCI).

Table S15*Parameter Estimates for Todd et al. (2021), Studies 1, 3, 4*

Parameter	μ	95% BCI Lower	95% BCI Upper
Use of Age (C_1)	0.899	0.553	0.974
Use of Race (C_2)	0.311	0.002	0.929
Response Bias (g)	0.641	0.339	0.839

Note. C_1 reflects the use of age cues and C_2 reflects the use of race cues. The g parameter reflects a response bias. Values represent the posterior mean (μ) and the 95% Bayesian Credibility Interval (BCI).

Table S16*Parameter Estimates for Todd et al. (2021), Studies 1; 3–6*

Parameter	μ	95% BCI Lower	95% BCI Upper
Use of Race (C_1)	0.926	0.911	0.940
Use of Age (C_2)	0.433	0.299	0.548
Response Bias (g)	0.662	0.608	0.716

Note. C_1 reflects the use of race cues and C_2 reflects the use of age cues. The g parameter reflects a response bias. Values represent the posterior mean (μ) and the 95% Bayesian Credibility Interval (BCI).

References

- Calanchini, J., Lai, C. K., & Klauer, K. C. (2021). Reducing implicit racial preferences: III. A process-level examination of changes in implicit preferences. *Journal of Personality and Social Psychology*, *121*, 796–818.
- DeBruine, L. (2018). Debruine/WebMorph: Beta Release 2. Zenodo.
- Smith, C. T., Calanchini, J., Hughes, S., Van Dessel, P., & De Houwer, J. (2020). The impact of instruction-and experience-based evaluative learning on IAT performance: A Quad model perspective. *Cognition and Emotion*, *34*, 21–41.