Supplemental information for "Domain-general object recognition predicts human ability to tell real from AI-generated faces"

To investigate the effect of careless responses on our hypotheses, we compared models fitted to the data for the first 250 participants regardless of attention check errors (ACEs – First250) to models fitted to the data for the 250 participants who committed no ACE (250NoACE). In study 1, these two samples overlap in that 218 of the participants are the same, but the 250NoACE sample replaced 32 participants who made 1 or more ACE, out of a total of 12 attention checks (22 participants made 1, 4 made 2, 2 made 3, 2 made 4, 1 made 6 and 1 made 9). Study 2 had 11 attention checks and 35 participants were replaced (22 participants made 1 ACE, 9 made 2, 1 made 3, 1 made 4 and 2 made 5).

The data and code we make available at https://osf.io/82krc/ (Study 1) and https://osf.io/cgbf5/ (Study 2) provide all the details about models. Here, we provide summary tables about the models in the main article (the 250NoAce models) compared to the corresponding First250 model in each study. Note that we do not provide a First250 comparison for the M2 model in Study 1, because it was inadmissible, with some eigenvalues smaller than zero and negative variances.

Tables 1 (Study 1) and 2 (Study 2) summarizes the fit indices for each model, as well as the effect of ACEs on standardized loadings, covariances, regressions and amount of variance explained for each variable. Within each category, the variables were sorted starting from those where ACEs had the least influence to those with the most influence.

Table 1

Metric	250NoACE	First250	Comparison
Chi-Square (χ^2) Test	67.517 (df=47, p=0.026)	73.913 (df=47, p=0.007)	Slightly worse fit for First250
CFI (Comparative Fit Index)	0.987	0.971	250NoACE has a better fit
TLI (Tucker-Lewis Index)	0.982	0.962	250NoACE has a better fit
RMSEA (Root Mean Square Err	0.049 (90% CI: 0.019-0.073, p=0.530)	0.051 (90% CI: 0.019-0.078, p=0.469)	Almost identical RMSEA values
SRMR (Standardized RMR)	0.042	0.053	250NoACE does better
AIC (Akaike Information Criter	-4871.83	-3711.99	Lower for 250NoACE
BIC (Bayesian Information Crit	-4757.96	-3607.07	Lower for 250NoACE
SABIC (Sample-size Adjusted B	-4856.26	-3705.31	Lower for 250NoACE
Loadings (standardized)			
AIFaceTest2	.719	.720	
AIFaceTest3	.572	.562	
AIFaceTest4	.705	.716	
AIFaceTest1	.663	.631	
CFMT2	.705	.663	
VFMT2	.753	.706	
VFMT1	.781	.726	
3AFCMatch1	.690	.630	
CFMT1	.676	.613	
3AFCMatch2	.618	.516	higher for 250NoACE
NOMT2	.753	.434	higher for 250NoACE
NOMT1	.781	.423	higher for 250NoACE
Covariances			
CFMT1<--> CFMT2	.711	.722	
NOMT1 <--> NOMT2	.547	.567	
VFMT1 <--> VFMT2	.292	.266	
3AFCMatch1 <-->3AFCMatch2	.480	.411	
O <--> FACE	.807	.676	higher for 250NoACE
R-square			
AIFaceTest2	.517	.519	
AIFaceTest3	.327	.316	
AIFaceTest4	.498	.513	
AIFaceTest1	.439	.398	higher for 250NoACE
CFMT2	.497	.439	higher for 250NoACE
VFM2T	.567	.499	higher for 250NoACE
3AFCMatch1	.476	.397	higher for 250NoACE
CFMT1	.457	.376	higher for 250NoACE
VFM1T	.610	.527	higher for 250NoACE
3AFCMatch2	.382	.266	higher for 250NoACE
NOMT1	.297	.179	higher for 250NoACE
AIFT	.234	.088	higher for 250NoACE
NOMT2	.339	.189	higher for 250NoACE
Regressions			
f-> AIFT	-.013	.155	
o -> AIFT	.494	.169	

Table 2

Metric	250NoACE	First250	Comparison
Chi-Square (œá–≤) Test	49.690 (df=42, p=0.194)	47.106 (df=35, p=0.083)	Slightly better for 250NoACE
CFI (Comparative Fit Index)	0.989	0.985	250NoACE has a better fit
TLI (Tucker-Lewis Index)	0.986	0.976	250NoACE has a better fit
RMSEA (Root Mean Square Error)	0.027 (CI: 0.000-0.053, p=0.922)	0.037 (CI: 0.000-0.062, p=0.775)	250NoACE performs better
SRMR (Standardized RMR)	0.047	0.045	First250 slightly better
AIC (Akaike Information Criterion)	-1155.408	-1064.036	Lower for 250NoACE
BIC (Bayesian Information Criterion)	-1070.893	-954.87	Lower for 250NoACE
SABIC (Sample-size Adjusted BIC)	-1146.975	-1053.143	Lower for 250NoACE
Loadings (standardized)			
PaperFolding <- g	.562	.552	
AIFaceT2	.683	.694	
Badeley <-g	.440	.428	
NOMT1 <- g	.391	.404	
Vocab <-g	.562	.533	
3AFCMatch2 <- O	.573	.540	
AIFaceT4	.678	.712	
AIFaceT1	.699	.663	
NOMT1 <- O	.400	.439	
NOMT2 <- O	.390	.342	
PaperFolding <- O	.476	.528	
NOMT2 <- g	.380	.444	
AIFaceT3	.667	.736	
3AFCMatch2 <- g	.227	.296	
3AFCMatch1 <- O	.525	.433	
3AFCMatch1 <- g	.208	.364	higher for First250
Covariances			
o <--> g	.000	.000	
NOMT1 <--> NOMT2	.610	.596	
3AFCMatch1<--> 3AFCMatch2	.346	.435	
R-square			
3AFCMatch1 <- g	.319	.320	
3AFCMatch2 <- g	.380	.379	
Vocab <-g	.275	.285	
Badeley <-g	.194	.183	
AIFaceT2	.467	.482	
NOMT2 <- g	.296	.314	
AIFT	.157	.136	
PaperFolding <- g	.543	.583	
NOMT1 <- g	.313	.355	
AIFaceT4	.460	.507	
AIFaceT1	.489	.439	
AIFaceT3	.445	.541	
Regressions			
o -> AIFT	.396	.360	
g-> AIFT	-.009	.078	