Supplement 1

Model Equations for the Proposed Multilevel Models in Study 1 and 2

In this Supplement, we formally describe both our proposed multilevel models. First, we decompose the individual IAT scores of each person i in task (domain) j additively into a latent within and a latent between component:

$$IAT_{ij} = IAT_j^B + IAT_{ij}^W \tag{A1}$$

where IAT_j^B represents the between IAT component (i.e., latent mean IAT score of cluster j) and IAT_{ij}^W denotes the within IAT component (i.e., IAT score of person i participating in task j centered within clusters). Next, we insert the within and between IAT components as latent predictors for the direct attitude measure at each level, resulting in the following random coefficients (or random-intercept-and-random-slopes) model:

Moderation model:

Level 1: $\qquad Exp_{ij} = b_{0j} + b_{1j} IAT_{ij}^W + \varepsilon_{ij} \tag{A2}$

Level 2: $\qquad b_{0j} = \gamma_{00} + \gamma_{01} IAT_j^B + v_{0j} \tag{A3}$

$\qquad\qquad b_{1j} = \gamma_{10} + \gamma_{11} IAT_j^B + v_{1j} \tag{A4}$

where Exp_{ij} is the direct attitude score of person i participating in task j, b_{0j} are the random intercepts, b_{1j} are the random slopes, and ε_{1j} is the residual term for person i participating in task j. The random intercepts b_{0j} are regressed upon IAT_j^B with γ_{00} being the fixed intercept, γ_{01} being the fixed slope, and v_{0j} being the residual term of the latent regression. Similarly, the random slopes b_{1j} are regressed on IAT_j^B with γ_{10} being the fixed intercept, γ_{11} being the fixed slope, and v_{1j} being the residual term of the latent regression.

Mediated moderation model:

Level 1: $\qquad Exp_{ij} = b_{0j} + b_{1j} IAT_{ij}^W + \varepsilon_{ij} \tag{A5}$

Level 2: $\quad b_{0j} = \gamma_{00} + \gamma_{01}IAT_j^B + \gamma_{02}ln(\sigma_{IAT_j}^2) + \upsilon_{0j}$ (A6)

$\quad b_{1j} = \gamma_{10} + \gamma_{11}IAT_j^B + \gamma_{12}\,ln(\sigma_{IAT_j}^2) + \upsilon_{1j}$ (A7)

where $ln(\sigma_{IAT_j}^2)$ is the log variance of the IAT scores of cluster j and added as an additional

predictor to explain the random intercepts and random slopes with γ_{02} and γ_{12} being the

corresponding fixed slopes, respectively. To test the indirect effect of the mean IAT scores (IAT

difficulties) on the random slopes via the log variances of the IAT scores we added the following

linear regression in addition to Equation 7:

Level 2: $\quad ln(\sigma_{IAT_j}^2) = \gamma_{20} + \gamma_{21}IAT_j^B + \vartheta_j$ (A8)

where $ln(\sigma_{IAT_j}^2)$ is regressed on IAT_j^B with γ_{20} being the fixed intercept and γ_{21} being the fixed

slope of that regression. Substituting Equation 8 into Equation 7 for the random slopes gives the

following Equation:

Level 2: $\quad b_{1j} = \gamma_{10} + \gamma_{11}IAT_j^B + \gamma_{12}\,(\gamma_{20} + \gamma_{21}\,IAT_j^B + \vartheta_j) + \upsilon_{1j}$

$$b_{1j} = \underbrace{\gamma_{10} + \gamma_{12}\gamma_{20}}_{mean\_slope} + \overbrace{\underbrace{\gamma_{11}IAT_j^B}_{direct1} + \underbrace{\gamma_{12}\gamma_{21}\,IAT_j^B}_{indirect1}}^{total\;Effect} + \gamma_{12}\vartheta_j + \upsilon_{1j} \qquad (A9)$$

Following a similar logic, the random intercepts can be replaced by:

Level 2: $\quad b_{0j} = \gamma_{00} + \gamma_{01}IAT_j^B + \gamma_{02}\,(\gamma_{20} + \gamma_{21}\,IAT_j^B + \vartheta_j) + \upsilon_{0j}$

$$b_{0j} = \underbrace{\gamma_{00} + \gamma_{02}\,\gamma_{20}}_{mean\_interc} + \underbrace{\gamma_{01}IAT_j^B}_{direct2} + \underbrace{\gamma_{02}\,\gamma_{21}\,IAT_j^B}_{indirect2} + \gamma_{02}\vartheta_j + \upsilon_{0j} \qquad (A10)$$

Supplement 2

Criteria for Excluding Participants in Study 1 and 2

In this Supplement we describe the criteria for excluding participants following recommendations by Hussey et al. (2018).

Participants were excluded when they did not have complete IAT data. The criteria used to judge the IAT data as incomplete were as follows: a) the number of blocks or the number of trials per block did not match the correct number, and b) no clarity on which domain, attribute category, attribute set, block order or IAT type was processed.

Participants were also excluded when their responses indicated that they did not work on the IAT properly. The criteria used to judge the processing of the IAT as improper were as follows: a) participants responded faster than 300 ms in 35% or more of the trials in any one practice block, faster than 300 ms in 25% or more of the trials in any one test block, or faster than 300 ms in 10% or more of the trials in all of the test blocks combined, b) participants were slower than 10,000 ms in 10% or more of the responses in all of the test blocks combined, and c) participants had an error rate of 50% or more in any one practice block, of 40% or more in all of the practice blocks combined, of 40% or more in any one test block, or of 30 % or more in all of the test blocks combined.

References

Hussey, I., Hughes, S., & Nosek, B. A. (2018). *The implicit and explicit Attitudes, Identities, and Individual Differences (AIID) Dataset.* https://doi.org/10.17605/OSF.IO/PCJWF

Supplement 3

Supplementary Results for Study 1

In this Supplement, we report additional results, including a) descriptive statistics, b) unstandardized parameter estimates of the two main multilevel models, c) unstandardized estimates of the most relevant variables resulting from the two main multilevel models per domain, d) tests of the moderators complementarity and social sensitivity as confounding variables, e) tests of the true-score variance of the attitude domains as a confounding variable, f) meta-analytical results, and g) Monte Carlo simulation studies, which we have omitted from the main text for clarity and ease of presentation of our results.

Descriptive statistics of the observed variables per domain

In what follows we report descriptive statistics of the direct attitude measures and the D scores per domain including the mean, variance, skewness, and kurtosis of the direct attitude measures (see Table 1) and the D scores (see Table 2), as well as the correlations between these variables (see Table 3). Note that we used listwise deletion as a method for handling missing values for estimating the descriptive statistics instead of Bayes which we used for the multilevel models. Therefore, the sample size for the descriptive statistics is slightly lower than for the multilevel models (n = 118,830).

Table 1

Mean, Variance, Skewness, and Kurtosis per Domain for the Direct Attitude Measures (Study 1)

Domain	Mean	Variance	Skew	Kurtosis
David Letterman - Jay Leno	-0.471	13.29	-0.001	-0.434
Tall People - Short People	-0.756	4.784	-0.07	1.25
Evolution - Creationism	-3.056	31.786	0.733	-0.685
Wrinkles - Plastic Surgery	-1.697	13.773	0.36	-0.316
Hot - Cold	-0.943	10.242	0.113	-0.135
Lawyers - Politicians	-1.589	5.825	0.042	0.223
Stable - Flexible	1.119	8.887	-0.319	-0.102
Protein - Carbohydrates	-0.134	8.651	-0.117	0.06

Domain	Mean	Variance	Skew	Kurtosis
Protestants - Catholics	-0.546	6.68	-0.008	0.825
Dogs - Cats	-1.427	18.415	0.173	-0.827
Jazz - Teen Pop	-1.733	17.4	0.242	-0.778
Astrology - Science	3.581	14.9	-0.602	-0.389
Meg Ryan - Julia Roberts	0.473	9.221	-0.248	-0.077
Relaxing - Exercising	-1.042	12.325	0.167	-0.484
Gun Control - Gun Rights	-2.986	25.979	0.721	-0.426
Microsoft - Apple	-0.215	16.164	0.209	-0.575
Kobe - Shaq	1.738	10.031	-0.034	-0.071
West Coast - East Coast	0.483	12.563	-0.147	-0.636
Denzel Washington - Tom Cruise	-2.735	10.705	0.209	-0.506
Organized Labor - Management	-0.156	14.635	-0.009	-0.541
New York - California	0.423	10.937	-0.189	-0.361
Team - Individual	1.68	12.513	-0.201	-0.604
Redsox - Yankees	-0.574	14.888	-0.056	0.075
Prolife - Prochoice	2.558	30.37	-0.747	-0.588
George Bush - John Kerry	2.611	24.23	-0.831	-0.42
Lord of the Rings - Harry Potter	-0.285	15.413	0.103	-0.666
Bill Clinton - Hillary Clinton	-0.843	7.747	0.083	0.097
Coffee - Tea	0.068	17.773	0.197	-0.86
Television - Books	2.446	12.654	-0.371	-0.486
Artists - Musicians	1.028	5.911	-0.178	0.216
Pepsi - Coke	1.468	17.408	-0.339	-0.564
Private - Public	-2.777	7.989	0.328	0.346
Briefs - Boxers	2.829	16.808	-0.616	-0.248
Capital Punishment - Imprisonment	2.433	13.477	-0.419	-0.497
Reason - Emotions	-1.041	10.681	0.054	-0.46
Friends - Family	0.52	8.929	-0.51	0.198
State - Church	-0.676	13.997	0.13	-0.699
Mother Teresa - Princess Diana	-1.35	8.768	0.341	0.227
Innocence - Wisdom	3.015	5.541	-0.233	-0.092
Helpers - Leaders	-0.189	9.479	-0.105	-0.157
Southerners - Northerners	1.164	9.194	-0.239	0.335
Burger King - McDonald's	-0.051	10.663	-0.068	-0.366
Canadian - American	-0.347	8.904	-0.216	0.061
Speed - Accuracy	2.279	7.083	-0.523	-0.065
Urban - Rural	-0.537	15.688	0.233	-0.631
African Americans - European Americans	0.872	4.179	-0.09	3.635
Pants - Skirts	-1.36	12.582	0.065	-0.338
Mountains - Ocean	0.86	10.065	-0.256	-0.368

Domain	Mean	Variance	Skew	Kurtosis
Receiving - Giving	1.432	7.471	-0.111	-0.388
Republicans - Democrats	2.987	22.115	-0.83	-0.409
Muslims - Jews	1.482	6.236	0.292	1.52
Tradition - Progress	1.471	9.431	-0.134	-0.348
Corporations - Nonprofits	2.841	11.067	-0.255	-0.427
Drinking - Abstaining	1.903	17.459	0.035	-0.891
Asians - Whites	0.693	4.422	0.541	2.11
Meat - Vegetables	0.825	16.998	0.165	-0.706
Foreign Places - American Places	-0.519	8.615	0.068	0.079
Traditional Values - Feminism	1.733	19.879	-0.348	-0.661
Numbers - Letters	1.416	11.042	-0.014	-0.246
Tax Reductions - Social Programs	1.617	16.812	-0.519	-0.267
Athletic People - Intelligent People	2.164	6.31	-0.07	-0.052
Japan - United States	1.733	9.32	-0.6	0.071
Hiphop - Classical	1.384	18.465	-0.153	-0.758
Security - Freedom	2.314	6.608	-0.15	-0.016
Gay People - Straight People	1.726	8.571	0.543	0.277
Single - Married	1.801	15.135	-0.522	-0.316
Jocks - Nerds	1.615	9.855	-0.187	0.017
Jews - Christians	-0.028	6.207	-0.008	1.03
Old People - Young People	0.702	6.221	-0.016	0.371
50 Cent - Britney Spears	-0.085	8.418	-0.076	0.232
Rich People - Beautiful People	1.484	5.386	0.02	0.481
Conservatives - Liberals	3.156	20.885	-0.839	-0.208
Fat People - Thin People	2.596	5.686	0.287	0.019
Winter - Summer	2.682	17.097	-0.668	-0.292
Atheism - Religion	0.712	27.404	-0.117	-1.061
Dramas - Comedies	0.897	9.873	-0.253	-0.264
Realism - Idealism	0.076	11.166	-0.099	-0.592
Career - Family	1.699	9.305	-0.399	-0.02
Strong - Sensitive	0.696	9.63	-0.024	-0.252
Solitude - Companionship	1.647	11.532	-0.073	-0.42
Technology - Nature	1.855	7.998	-0.246	0.003
Effort - Talent	0.338	9.043	0.019	-0.378
National Defense - Education	4.25	7.978	-0.606	0.26
Rebellious - Conforming	-0.474	13.556	0.298	-0.405
Determinism - Free will	4.693	9.532	-0.831	0.66
Manufactured - Natural	2.994	7.272	-0.234	-0.011
Night - Morning	-1.302	15.331	0.236	-0.737
Past - Future	2.183	7.866	-0.334	0.024

Domain	Mean	Variance	Skew	Kurtosis
Difficult - Simple	0.518	11.702	0.178	-0.494
Money - Love	2.964	6.711	-0.663	0.626
Chaos - Order	4.101	11.411	-0.578	-0.295
Punishment - Forgiveness	3.512	12.092	-0.56	-0.086
Poor People - Rich People	0.143	7.166	0.041	0.328
Avoiding - Approaching	3.287	10.471	-0.652	0.468
Skeptical - Trusting	1.363	13.768	-0.328	-0.413

Table 2

Mean, Variance, Skewness, and Kurtosis of the D Scores for the IATs per Domain (Study 1)

Domain	Mean	Variance	Skew	Kurtosis
David Letterman - Jay Leno	0.005	0.141	-0.095	-0.369
Tall People - Short People	0.011	0.295	-0.005	-0.746
Evolution - Creationism	0.013	0.298	0.071	-0.682
Wrinkles - Plastic Surgery	0.013	0.241	0.068	-0.396
Hot - Cold	0.014	0.378	0.002	-0.867
Lawyers - Politicians	0.014	0.141	-0.01	-0.158
Stable - Flexible	0.027	0.27	-0.088	-0.592
Protein - Carbohydrates	0.027	0.248	-0.118	-0.681
Protestants - Catholics	0.033	0.22	-0.024	-0.535
Dogs - Cats	0.041	0.253	-0.108	-0.676
Jazz - Teen Pop	0.041	0.232	-0.017	-0.569
Astrology - Science	0.043	0.293	-0.115	-0.56
Meg Ryan - Julia Roberts	0.046	0.172	-0.032	-0.412
Relaxing - Exercising	0.047	0.346	-0.085	-0.776
Gun Control - Gun Rights	0.048	0.195	0.149	-0.387
Microsoft - Apple	0.048	0.228	0.031	-0.535
Kobe - Shaq	0.05	0.169	-0.098	-0.386
West Coast - East Coast	0.062	0.283	-0.143	-0.693
Denzel Washington - Tom Cruise	0.072	0.176	-0.161	-0.332
Organized Labor - Management	0.082	0.227	-0.072	-0.481
New York - California	0.092	0.213	-0.045	-0.515
Team - Individual	0.119	0.275	-0.328	-0.432
Redsox - Yankees	0.122	0.206	-0.237	-0.302
Prolife - Prochoice	0.123	0.223	-0.281	-0.392
George Bush - John Kerry	0.124	0.222	-0.194	-0.401
Lord of the Rings - Harry Potter	0.135	0.247	-0.184	-0.47
Bill Clinton - Hillary Clinton	0.157	0.184	-0.305	-0.266
Coffee - Tea	0.158	0.215	-0.208	-0.43

Domain	Mean	Variance	Skew	Kurtosis
Television - Books	0.164	0.172	-0.152	-0.369
Artists - Musicians	0.164	0.149	-0.148	-0.302
Pepsi - Coke	0.169	0.221	-0.233	-0.499
Private - Public	0.17	0.23	-0.08	-0.488
Briefs - Boxers	0.174	0.23	-0.355	-0.272
Capital Punishment - Imprisonment	0.181	0.165	-0.151	-0.26
Reason - Emotions	0.184	0.224	-0.315	-0.307
Friends - Family	0.189	0.171	-0.188	-0.196
State - Church	0.19	0.235	-0.218	-0.498
Mother Teresa - Princess Diana	0.196	0.146	-0.174	0.062
Innocence - Wisdom	0.199	0.218	-0.114	-0.28
Helpers - Leaders	0.206	0.236	-0.266	-0.388
Southerners - Northerners	0.207	0.248	-0.333	-0.377
Burger King - McDonald's	0.226	0.148	-0.312	-0.047
Canadian - American	0.233	0.223	-0.466	0.003
Speed - Accuracy	0.238	0.187	-0.239	-0.232
Urban - Rural	0.239	0.262	-0.361	-0.403
African Americans - European Americans	0.245	0.19	-0.292	-0.16
Pants - Skirts	0.245	0.211	-0.298	-0.345
Mountains - Ocean	0.248	0.191	-0.319	-0.051
Receiving - Giving	0.259	0.15	-0.213	-0.022
Republicans - Democrats	0.26	0.308	-0.466	-0.57
Muslims - Jews	0.267	0.137	-0.188	-0.036
Tradition - Progress	0.268	0.232	-0.312	-0.255
Corporations - Nonprofits	0.27	0.218	-0.345	-0.193
Drinking - Abstaining	0.281	0.281	-0.387	-0.363
Asians - Whites	0.284	0.163	-0.352	-0.127
Meat - Vegetables	0.298	0.251	-0.433	-0.183
Foreign Places - American Places	0.298	0.209	-0.338	-0.149
Traditional Values - Feminism	0.302	0.263	-0.472	-0.403
Numbers - Letters	0.302	0.222	-0.286	-0.351
Tax Reductions - Social Programs	0.307	0.187	-0.404	-0.154
Athletic People - Intelligent People	0.32	0.136	-0.26	-0.307
Japan - United States	0.326	0.163	-0.305	-0.168
Hiphop - Classical	0.329	0.206	-0.392	-0.19
Security - Freedom	0.333	0.161	-0.207	-0.438
Gay People - Straight People	0.339	0.194	-0.316	-0.186
Single - Married	0.342	0.23	-0.426	-0.247
Jocks - Nerds	0.352	0.181	-0.381	-0.056
Jews - Christians	0.372	0.183	-0.505	0.087

Domain	Mean	Variance	Skew	Kurtosis
Old People - Young People	0.38	0.152	-0.343	0.024
50 Cent - Britney Spears	0.386	0.151	-0.355	-0.074
Rich People - Beautiful People	0.388	0.128	-0.393	-0.138
Conservatives - Liberals	0.39	0.319	-0.8	-0.005
Fat People - Thin People	0.416	0.18	-0.366	-0.028
Winter - Summer	0.432	0.266	-0.579	-0.145
Atheism - Religion	0.44	0.241	-0.501	-0.086
Dramas - Comedies	0.464	0.194	-0.459	-0.085
Realism - Idealism	0.517	0.228	-0.563	-0.082
Career - Family	0.518	0.153	-0.401	-0.022
Strong - Sensitive	0.52	0.196	-0.38	-0.216
Solitude - Companionship	0.526	0.186	-0.428	-0.191
Technology - Nature	0.531	0.171	-0.498	0.079
Effort - Talent	0.539	0.156	-0.38	-0.172
National Defense - Education	0.621	0.121	-0.487	0.428
Rebellious - Conforming	0.633	0.212	-0.58	-0.125
Determinism - Free will	0.649	0.14	-0.41	0.076
Manufactured - Natural	0.654	0.13	-0.437	-0.046
Night - Morning	0.657	0.162	-0.59	0.224
Past - Future	0.673	0.155	-0.553	0.239
Difficult - Simple	0.688	0.212	-0.79	0.47
Money - Love	0.756	0.101	-0.703	0.871
Chaos - Order	0.76	0.182	-0.888	0.903
Punishment - Forgiveness	0.812	0.121	-0.747	0.782
Poor People - Rich People	0.829	0.134	-0.746	0.544
Avoiding - Approaching	0.933	0.107	-0.757	0.691
Skeptical - Trusting	0.938	0.131	-1.142	1.819

Table 3

Correlations Between the Observed Variables Across Domains (Study 1)

Variable	1	2	3	4	5	6	7	8
1. Mean D score								
2. Variance D score	-.51**							
	[-.65, -.35]							
3. Skew D score	-.89**	.29**						
	[-.93, -.84]	[.10, .47]						
4. Kurtosis D score	.85**	-.63**	-.83**					
	[.78, .90]	[-.73, -.49]	[-.88, -.75]					
5. Mean direct att[a]	.25*	-.12	-.17	.23*				
	[.05, .43]	[-.31, .09]	[-.36, .03]	[.03, .42]				
6. Variance direct att	-.18	.46**	.04	-.21*	.18			
	[-.37, .02]	[.28, .60]	[-.16, .24]	[-.39, -.01]	[-.03, .36]			
7. Skew direct att	-.29**	.09	.39**	-.27**	-.44**	-.06		
	[-.46, -.09]	[-.12, .28]	[.21, .55]	[-.44, -.07]	[-.59, -.26]	[-.25, .15]		
8. Kurtosis direct att	.10	-.32**	-.02	.16	-.02	-.59**	.02	
	[-.11, .29]	[-.49, -.12]	[-.22, .18]	[-.04, .35]	[-.22, .18]	[-.71, -.45]	[-.19, .22]	
9. I-E cor	-.52**	.56**	.28**	-.47**	-.09	.70**	.02	-.39**
	[-.65, -.35]	[.40, .68]	[.08, .45]	[-.62, -.30]	[-.29, .11]	[.58, .79]	[-.18, .22]	[-.55, -.20]

Note. att = attitude; I-E cor = implicit explicit correlation. Values in square brackets indicate the 95% confidence interval for each correlation.

[a] We took the absolute value of the means of the direct attitude measures so that the values ranged from 0 to 9.
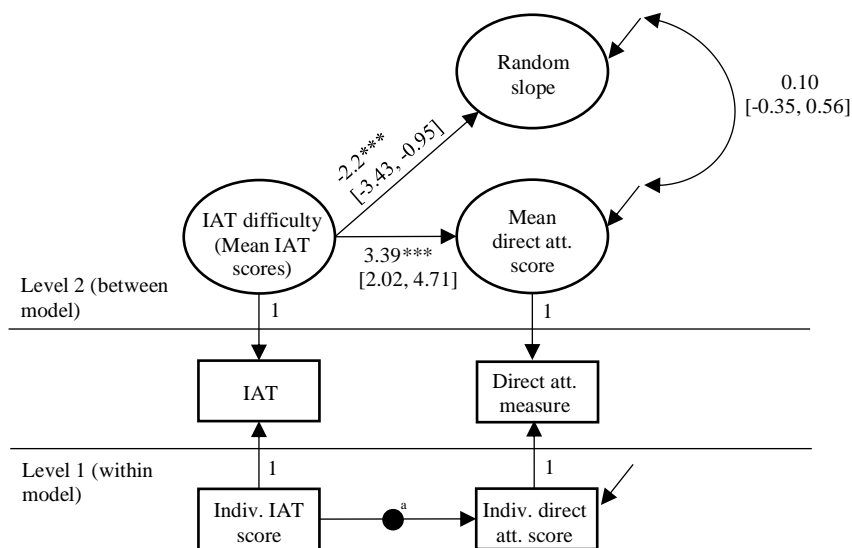
* indicates $p < .05$. ** indicates $p < .01$.

Unstandardized parameter estimates of the two main multilevel models

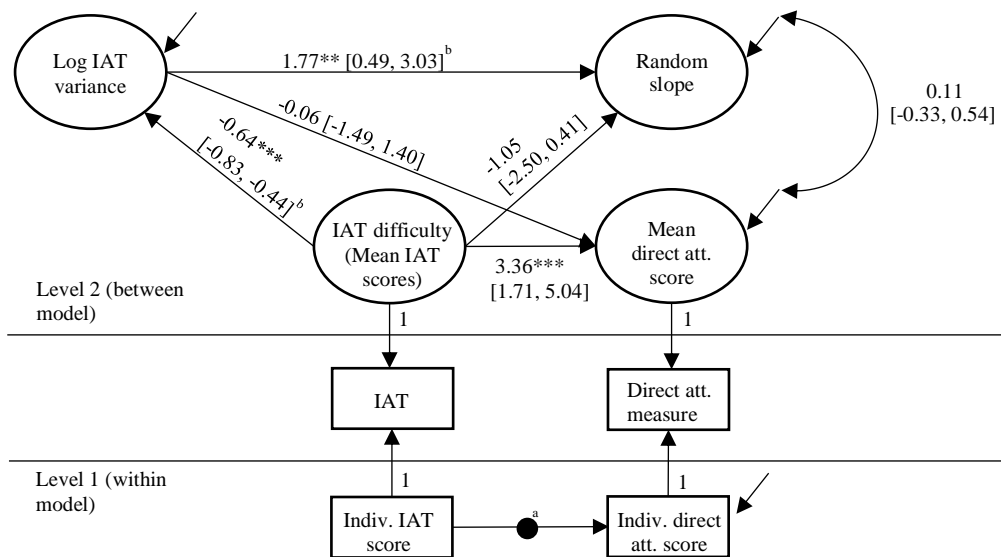Results of the unstandardized parameter estimates of the two main multilevel models are presented in Figure 1.

Figure 1

Unstandardized Parameter Estimates for the Multilevel Moderation Model (Panel A) and the Multilevel Mediated Moderation Model (Panel B) in Study 1

A

B

Note. Circles represent latent and rectangles observed variables. Numbers without brackets ascribed to single-headed arrows are path coefficients, numbers without brackets ascribed to double-headed arrows are covariances, and numbers in square brackets are 95% credible intervals. Darkened circles represent random slopes which are also depicted on Level 2. IAT

= Implicit Association Test; Log IAT variance = Log IAT true-score variance; att = attitude;

Indiv = individual.

[a] Mplus does not provide the unstandardized within effect across all cluster (Muthén & Muthén, 1998-2017; Schuurman et al., 2016).

[b] The indirect effect of IAT difficulty on the random slope via log IAT true-score variance was significant, $B = -1.10$, p$SD = 0.46$, 95% *C.I.* [-2.10, -0.30].

*$p < .05$, **$p < .01$, ***$p < .001$

Unstandardized estimates of the most relevant variables resulting from the two main multilevel models per domain

For a better overview of the test difficulty of the attitude IATs, the true-score variance of the attitude IATs and their relationship with the direct attitude measures we present below the corresponding unstandardized estimates resulting from our two main multilevel models. Table 4 displays the test difficulty of the IATs and their relationship with the direct attitude measures based on the multilevel moderation model. Table 5 displays the test difficulty of the IATs, the log true-score variance of the IATs and their relationship with the direct attitude measures based on the multilevel mediated moderation model.

Table 4

Domains and Their Corresponding IAT Test Difficulty and Slope Estimated From the Multilevel Moderation Model (Study 1)

Domain	IAT difficulty[a]	Slope[b]
Lawyers - Politicians	0.010	2.139
Protein - Carbohydrates	0.014	2.282
Protestants - Catholics	0.019	2.773
Tall People - Short People	0.019	1.008
David Letterman - Jay Leno	0.020	4.409
Wrinkles - Plastic Surgery	0.020	3.261
Hot - Cold	0.022	1.495
Evolution - Creationism	0.023	6.473
Jazz - Teen Pop	0.027	4.553
Gun Control - Gun Rights	0.035	5.917

Domain	IAT difficulty[a]	Slope[b]
Microsoft - Apple	0.035	4.617
Kobe - Shaq	0.038	3.441
Relaxing - Exercising	0.041	2.357
Astrology - Science	0.043	3.712
Dogs - Cats	0.052	5.126
Stable - Flexible	0.057	2.064
Meg Ryan - Julia Roberts	0.064	3.453
West Coast - East Coast	0.070	3.764
Denzel Washington - Tom Cruise	0.085	3.475
New York - California	0.085	2.999
Organized Labor - Management	0.087	4.053
Redsox - Yankees	0.109	4.854
Team - Individual	0.112	2.826
George Bush - John Kerry	0.123	6.723
Lord of the Rings - Harry Potter	0.123	4.956
Prolife - Prochoice	0.128	7.320
Artists - Musicians	0.142	2.032
Bill Clinton - Hillary Clinton	0.145	2.677
Briefs - Boxers	0.151	4.137
Private - Public	0.151	1.904
Television - Books	0.155	3.987
Coffee - Tea	0.157	3.873
Pepsi - Coke	0.163	4.777
Reason - Emotions	0.175	2.847
State - Church	0.185	3.780
Innocence - Wisdom	0.191	1.164
Friends - Family	0.194	2.820
Capital Punishment - Imprisonment	0.195	2.693
Helpers - Leaders	0.200	2.443
Mother Teresa - Princess Diana	0.205	3.028
Southerners - Northerners	0.210	2.327
Canadian - American	0.224	3.215
Urban - Rural	0.236	4.416
Mountains - Ocean	0.237	3.107
Pants - Skirts	0.239	2.037
African Americans - European Americans	0.240	1.655
Speed - Accuracy	0.242	1.788
Burger King - McDonald's	0.243	3.718
Republicans - Democrats	0.255	5.797
Corporations - Nonprofits	0.262	3.286
Muslims - Jews	0.262	2.847

Domain	IAT difficulty[a]	Slope[b]
Receiving - Giving	0.262	1.610
Asians - Whites	0.272	1.958
Foreign Places - American Places	0.275	2.488
Meat - Vegetables	0.277	4.828
Traditional Values - Feminism	0.277	5.647
Drinking - Abstaining	0.279	3.601
Tradition - Progress	0.288	3.176
Numbers - Letters	0.301	2.669
Athletic People - Intelligent People	0.310	2.021
Tax Reductions - Social Programs	0.315	4.275
Hiphop - Classical	0.319	4.949
Japan - United States	0.325	3.418
Gay People - Straight People	0.336	2.985
Single - Married	0.343	3.087
Jocks - Nerds	0.351	2.602
Security - Freedom	0.351	2.192
Old People - Young People	0.369	1.191
Conservatives - Liberals	0.373	5.224
50 Cent - Britney Spears	0.381	2.193
Rich People - Beautiful People	0.387	1.169
Jews - Christians	0.393	2.593
Atheism - Religion	0.414	5.781
Winter - Summer	0.425	3.485
Fat People - Thin People	0.429	1.481
Dramas - Comedies	0.469	2.442
Strong - Sensitive	0.511	1.662
Solitude - Companionship	0.519	2.509
Realism - Idealism	0.521	3.289
Technology - Nature	0.527	2.781
Career - Family	0.529	2.246
Effort - Talent	0.542	2.670
National Defense - Education	0.604	3.342
Manufactured - Natural	0.637	1.975
Rebellious - Conforming	0.638	3.223
Determinism - Free will	0.647	1.923
Night - Morning	0.663	2.881
Past - Future	0.667	1.171
Difficult - Simple	0.711	2.810
Money - Love	0.741	2.122
Chaos - Order	0.747	3.143
Poor People - Rich People	0.809	1.924

Domain	IAT difficulty[a]	Slope[b]
Punishment - Forgiveness	0.817	1.614
Skeptical - Trusting	0.913	1.988
Avoiding - Approaching	0.921	0.738

Note. IAT = Implicit Association Test.

[a] IAT difficulty corresponds to the latent mean IAT scores per domain.

[b] The slope corresponds to the within effect of attitude IAT scores on direct attitude scores per domain.

Table 5

Domains and Their Corresponding IAT Test Difficulty, IAT True-Score Variance and Slope Estimated From the Multilevel Mediated Moderation Model (Study 1)

Domain	IAT difficulty[a]	Log IAT variance[b]	Slope[c]
Wrinkles - Plastic Surgery	0.007	-1.437	3.418
David Letterman - Jay Leno	0.009	-1.982	4.543
Protein - Carbonhydrates	0.013	-1.404	2.387
Hot - Cold	0.016	-1.033	1.519
Lawyers - Politicians	0.018	-1.972	2.265
Evolution - Creationism	0.029	-1.248	6.705
Relaxing - Exercising	0.033	-1.059	2.617
Protestants - Catholics	0.038	-1.523	3.119
Kobe - Shaq	0.040	-1.768	3.102
Astrology - Science	0.041	-1.244	3.930
Tall People - Short People	0.042	-1.19	0.927
Stable - Flexible	0.043	-1.349	2.281
Dogs - Cats	0.044	-1.374	4.915
Meg Ryan - Julia Roberts	0.065	-1.828	3.356
Jazz - Teen Pop	0.068	-1.499	4.817
New York - California	0.070	-1.539	3.150
Microsoft - Apple	0.071	-1.535	4.472
Gun Control - Gun Rights	0.072	-1.663	6.612
Denzel Washington - Tom Cruise	0.091	-1.687	3.212
West Coast - East Coast	0.099	-1.248	3.704
Organized Labor - Management	0.108	-1.483	4.107
Team - Individual	0.117	-1.322	2.330
George Bush - John Kerry	0.122	-1.488	6.745
Redsox - Yankees	0.125	-1.58	4.869

Domain	IAT difficulty[a]	Log IAT variance[b]	Slope[c]
Prolife - Prochoice	0.127	-1.505	7.774
Coffee - Tea	0.132	-1.501	3.742
Lord of the Rings - Harry Potter	0.134	-1.404	4.773
Artists - Musicians	0.156	-1.841	2.201
Briefs - Boxers	0.157	-1.495	3.902
Pepsi - Coke	0.157	-1.551	5.054
Bill Clinton - Hillary Clinton	0.164	-1.690	2.733
Television - Books	0.164	-1.760	4.343
Capital Punishment - Imprisonment	0.168	-1.769	2.854
Innocence - Wisdom	0.172	-1.539	1.263
Reason - Emotions	0.177	-1.426	3.335
Private - Public	0.180	-1.504	1.517
Friends - Family	0.190	-1.671	2.930
Helpers - Leaders	0.209	-1.511	2.710
Canadian - American	0.212	-1.507	3.035
Mother Teresa - Princess Diana	0.213	-1.946	2.566
State - Church	0.213	-1.487	3.847
Southerners - Northerners	0.219	-1.353	2.408
Burger King - McDonald's	0.229	-1.919	3.841
Speed - Accuracy	0.238	-1.590	1.231
African Americans - European Americans	0.245	-1.593	2.049
Pants - Skirts	0.245	-1.571	1.935
Mountains - Ocean	0.246	-1.639	3.542
Republicans - Democrats	0.253	-1.247	5.874
Urban - Rural	0.253	-1.328	4.234
Tradition - Progress	0.258	-1.508	3.067
Receiving - Giving	0.264	-1.805	1.11
Muslims - Jews	0.275	-1.971	2.681
Corporations - Nonprofits	0.276	-1.595	3.574
Meat - Vegetables	0.285	-1.394	4.631
Asians - Whites	0.289	-1.859	1.745
Tax Reductions - Social Programs	0.298	-1.631	3.853
Foreign Places - American Places	0.301	-1.587	2.962
Athletic People - Intelligent People	0.308	-1.971	2.356
Traditional Values - Feminism	0.309	-1.307	5.828
Drinking - Abstaining	0.310	-1.272	3.604
Numbers - Letters	0.311	-1.521	2.271
Gay People - Straight People	0.317	-1.697	2.492
Japan - United States	0.329	-1.883	3.456
Jocks - Nerds	0.330	-1.745	2.692

Domain	IAT difficulty[a]	Log IAT variance[b]	Slope[c]
Security - Freedom	0.349	-1.855	2.196
Hiphop - Classical	0.360	-1.632	5.047
Single - Married	0.367	-1.468	3.548
Jews - Christians	0.377	-1.753	2.447
Rich People - Beautiful People	0.377	-2.068	1.175
50 Cent - Britney Spears	0.386	-1.891	1.914
Old People - Young People	0.386	-1.855	1.260
Conservatives - Liberals	0.398	-1.157	5.447
Fat People - Thin People	0.421	-1.660	1.273
Winter - Summer	0.429	-1.354	3.260
Atheism - Religion	0.442	-1.429	5.957
Dramas - Comedies	0.485	-1.617	2.481
Strong - Sensitive	0.504	-1.644	1.419
Solitude - Companionship	0.507	-1.692	2.623
Realism - Idealism	0.518	-1.519	2.797
Technology - Nature	0.522	-1.749	2.875
Career - Family	0.530	-1.919	2.485
Effort - Talent	0.541	-1.852	2.301
National Defense - Education	0.609	-2.042	2.508
Rebellious - Conforming	0.626	-1.545	3.108
Manufactured - Natural	0.632	-2.043	2.457
Determinism - Free will	0.642	-1.954	1.594
Past - Future	0.680	-1.790	1.336
Difficult - Simple	0.695	-1.585	2.442
Night - Morning	0.700	-1.811	2.622
Money - Love	0.741	-2.170	1.766
Chaos - Order	0.760	-1.710	3.121
Poor People - Rich People	0.820	-1.998	1.480
Punishment - Forgiveness	0.822	-2.100	1.821
Avoiding - Approaching	0.917	-2.283	1.047
Skeptical - Trusting	0.942	-1.973	1.888

Note. IAT = Implicit Association Test.

[a] IAT difficulty corresponds to the latent mean IAT scores per domain.

[b] Log IAT variance corresponds to the log true-score variance of the IAT scores per domain;

the closer log IAT true-score variance is to zero the higher the true-score variance.

[c] The slope corresponds to the within effect of attitude IAT scores on direct attitude scores per

domain.

Testing the additional moderators complementarity and social sensitivity as confounding variables

      Definitions and explanations of the moderators for the coders. Both moderators were defined and explained according to Greenwald et al. (2009). That is, with regard to social sensitivity coders received the following information: "Please judge the extent to which self-reporting the construct assessed by the measure might activate concerns about the impression that the response would make on others. For example, self-reporting attitudes toward Black Americans is something that you should judge to be considerably more socially sensitive than self-reporting attitudes toward brands of yogurt. Please make each judgment on a scale of 1–7 (1 = *not at all likely to be affected by social desirability concerns*, 7 = *extremely likely to be affected by social desirability concerns*). To repeat for clarity, the social sensitivity measure for IAT measures should be judged to be the sensitivity associated with self-reporting the same attitude, belief, or self-related measure (note: Black/White IAT is a 7)." In addition to Greenwald's explanations coders were also told to consider whether, when comparing the target categories, the more positive or more negative evaluation of one of the two targets is condemned by the general public (public opinion in the USA) and how strong this condemnation is in each case. If the strength of condemnation differs between the comparisons, coders should choose the more socially sensitive comparison (e.g., in a Men/Women IAT the most socially sensitive comparison is evaluating women more negatively than men and not evaluating men more positively than women or women more positively than men or men more negatively than women).

      With regard to complementarity coders received the following information: "for some preferences, liking one alternative implies disliking a complementary alternative. For example, having a positive attitude toward a candidate of one political party might imply having a negative attitude toward a political competitor from another party, but it might not imply having a negative attitude toward another candidate from the same party. In contrast,

having a positive attitude toward one brand of yogurt might not imply having a negative

attitude toward other brands of yogurt. To rate complementarity, please estimate the extent to

which liking one of the two IAT target categories in a measure implied disliking the other.

Please make each judgment on a scale of 1–9 (1 = *extremely noncomplementary*, 9 =

extremely complementary). Complementarity is high when liking one target category implies

disliking its contrasted category." In addition to Greenwald's explanations coders were also

told that their judgments should not be based on their own understanding of the target

categories but on whether the target categories are complementary from an empirical

perspective that is from the perspective of the general public in the USA (e.g. in an African

American/European American IAT one may not see the target categories as complementary at

all, but in American society they are complementary to some degree, because there are people

for whom the positive evaluation of European Americans is inherently accompanied by a

more negative evaluation of African Americans).

Coding procedure. Each domain was coded with respect to the moderators

complementarity and social sensitivity by two coders who were otherwise not involved in the

project. Accordingly, both coders were blind to the central results of the study. Prior to

coding, they underwent training. First, they received the definitions and additional

explanations about the moderators and had the opportunity to discuss them and ask questions.

Second, they coded the first ten domains and discussed any disagreements with a supervisor

until agreement was reached. This deepened their understanding of the moderators and gave

them a baseline against which to compare their assessment of the following domains. After

completing the training, they then coded the remaining 85 domains independently.

Interrater reliability. Interrater reliability between the two coders was established

based on the 85 domains that the two coders rated independently. Because the scales of the

moderators consisted of multiple response categories, we chose the intraclass correlation

(ICC) for metric variables as the interrater reliability index. Because we assume that the

domains were randomly drawn but the coders were not (we selected the coders based on their

skills, their availability, etc.) we applied a two-way mixed effects model once for the single

raters (ICC3) and once for the mean of the raters (ICC3k; cf. Shrout & Fleiss, 1979). All ICCs

were calculated using the package psych (Revelle, 2021). Results can be found in Table 6. For

both moderators, interrater reliability was good with ICCs ≥ .67 (cf. Cicchetti, 1994).

Accordingly, we averaged the coders' ratings for each of the moderators and used the average

ratings as the final measures for the moderators.

Table 6

Interrater Reliability between coders for the two moderators social sensitivity and

complementarity

Moderator	IRR	Index
Social sensitivity	.67*** [.53, .77]	ICC3
Social sensitivity	.80*** [.69, .87]	ICC3k
Complementarity	.70*** [.57, .79]	ICC3
Complementarity	.82*** [.72, .88]	ICC3k

Note. IRR = interrater reliability; ICC intraclass correlation. Values in square brackets

indicate the 95% confidence interval.

*p < .05, **p < .01, ***p < .001.

Multilevel multivariate moderation models. We ran two multilevel models to test

the effects of the moderators social sensitivity and complementarity both of which were

implemented in a similar manner as the multilevel moderation model in the main text of

Study 1. The first model equaled the multilevel moderation model with the difference that in

addition to the moderator IAT test difficulty the moderators social sensitivity and

complementarity were included on Level 2 (see Figure 2 Panel A). The second model equaled

the first model with the difference that only the moderators social sensitivity and

complementarity and not the moderator IAT test difficulty were included on Level 2 (see

Figure 2 Panel B). This procedure allowed us to estimate the variation in the slopes IAT test

difficulty explained over and above the other two moderators.

Figure 2

 Standardized Parameter Estimates for the Multilevel Multivariate Moderation Model

Including IAT Test Difficulty, Social Sensitivity and Complementarity as Moderators (Panel

A) and for the Multilevel Multivariate Moderation Model Including Social Sensitivity and

Complementarity as Moderators (Panel B) in Study 1
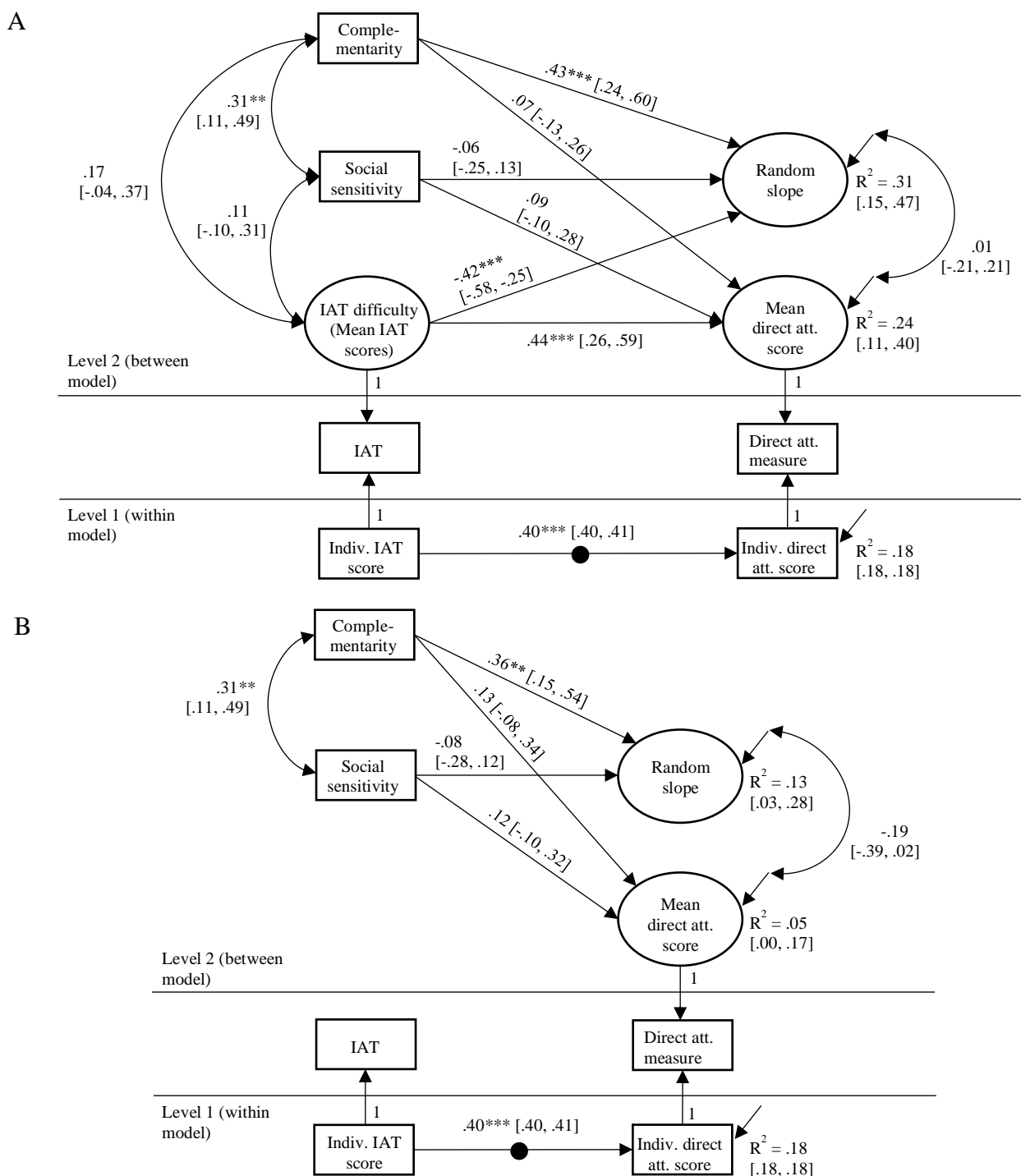
Note. Circles represent latent and rectangles observed variables. Numbers without brackets ascribed to single-headed arrows are path coefficients, numbers without brackets ascribed to double-headed arrows are correlation coefficients, and numbers in square brackets are 95% credible intervals. Darkened circles represent random slopes which are also depicted on Level 2. R^2 = the coefficient of determination; IAT = implicit association test; att = attitude; Indiv = individual.

*p < .05, **p < .01, ***p < .001.

Tests of the true-score variance of the to-be-measured constructs as a confounding variable

We ran two multilevel models to test the true-score variance of the to-be-measured constructs as a confounding variable using the true-score variance of the direct attitude measures as a proxy for the true-score variance of the to-be-measured constructs. The first model is a correlational model where we only model the correlations between the Level 2 variables IAT test difficulty, the log true-score variance of the direct attitude measure and the mean of the direct attitude measure (see Figure 3 Panel A). The second model was implemented in a similar manner as the multilevel moderation model in the main text of Study 1 with a few noteworthy differences. On Level 1 we exchanged the independent and the dependent variable and regressed the individual IAT scores on the individual direct attitude scores. This was necessary to estimate the log true-score variance of the direct attitude scores on Level 2. Furthermore we then used the log true-score variance of the direct attitude scores as an additional moderator to the IAT test difficulty to explain the random slope (see Figure 3 Panel B).

Figure 3

Standardized Parameter Estimates for the Multilevel Correlational Model (Panel A) and the

Multilevel Moderation Model (Panel B) Testing the True-Score Variance of the Direct

Attitude Measure as a Confounding Variable in Study 1

A
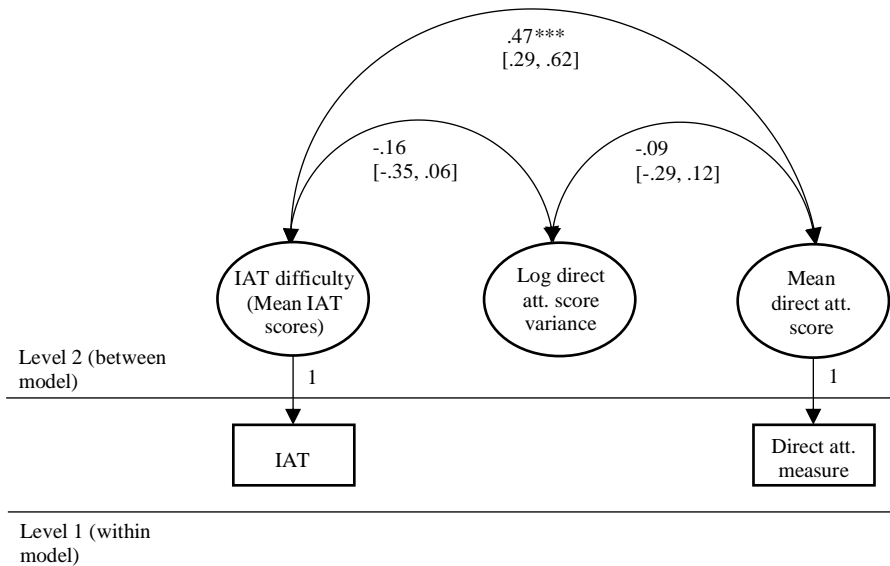
B
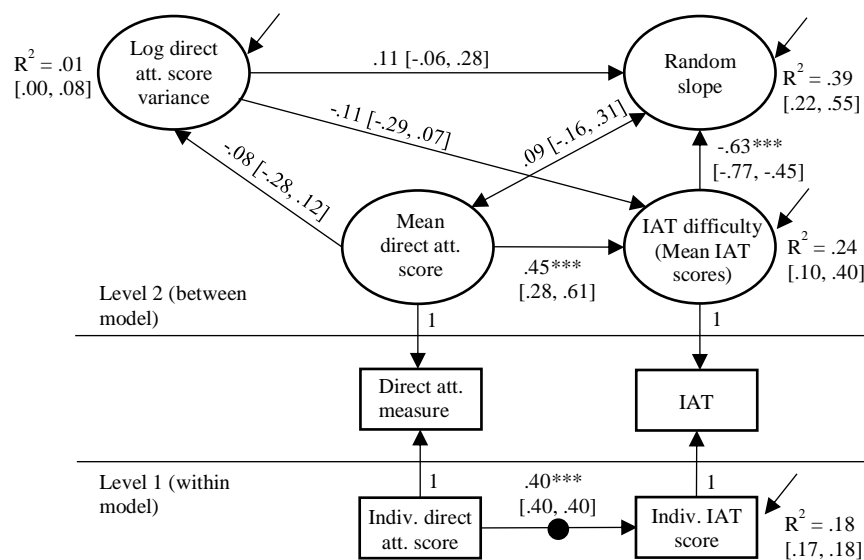
Note. Circles represent latent and rectangles observed variables. Numbers without brackets

ascribed to single-headed arrows are path coefficients, numbers without brackets ascribed to

double-headed arrows are correlation coefficients, and numbers in square brackets are 95%

credible intervals. Darkened circles represent random slopes which are also depicted on Level

2. R^2 = the coefficient of determination; IAT = implicit association test; att = attitude; Indiv =

individual.

*p < .05, **p < .01, ***p < .001.

Meta-Analytical Results

We conducted additional meta-analyses to allow for better comparisons with previous studies testing moderators of IAT relations.

Moderation hypothesis (H1): In order to conduct a meta-analysis we treated the 95 domains as 95 studies resulting in 95 IAT-direct attitude correlation coefficients following Cheung and Jak's (2016) split/analyze/meta-analyze (SAM) approach. We used meta-regression to assess the moderating effect of IAT test difficulty. Accordingly, we assumed a mixed effect model (Harrer et al., 2022). Before effect size pooling we Fisher's z-transformed the correlations. As estimator of between study heterogeneity we chose the restricted maximum likelihood procedure (Veroniki et al., 2016; Viechtbauer, 2005). The final results were as follows: The pooled association between attitude IAT scores and direct attitude scores was $r = .41$, $p < .001$, 95% CI [.38, .44]. The between-study heterogeneity variance was estimated at $\tau^2 = 0.03$, 95% CI [0.02, 0.036], with an I^2 value of 96.9%, 95% CI [96.6%, 97.2%]. Accordingly, moderator analysis was warranted. IAT test difficulty explained 24.06% of the heterogeneity in effect sizes. Unsurprisingly, the test of moderators was significant, $QM(1) = 29.61$, $p < .001$, as well as the unstandardized regression weight of IAT test difficulty, $B = -0.35$, $p < .001$, 95% CI [-0.48, -0.23]. The analysis was done in R using the meta package (Schwarzer, 2007).

Mediated moderation hypothesis (H2): We proceeded as outlined above only this time we included IAT test difficulty as well as IAT total variance as moderators into the meta-regression. Accordingly, the results corresponded to the previous results except for the moderation analysis. IAT test difficulty and IAT total variance together explained 36.9% of the heterogeneity in effect sizes. Unsurprisingly, the test of moderators was significant once again, $QM(2) = 54.35$, $p < .001$. The unstandardized regression weight of IAT total variance was significant, $B = 1.27$, $p < .001$, 95% CI [0.70, 1.85], as well as the unstandardized

regression weight of IAT test difficulty, $B = -0.2$, $p < .01$, 95% CI [-0.33, -0.06]. Interestingly, in contrast to the multilevel approach this result does not suggest a full but only a partial mediation. However, the analysis did not allow for estimating indirect effects so that conclusions concerning the mediation hypothesis should be treated with caution. Again the analysis was done in R using the meta package (Schwarzer, 2007).

Monte Carlo Simulations

We run two Monte Carlo simulation studies to evaluate the statistical performance of the multilevel models used in the application in Study 1. In the first simulation study, we simulated data based on the *multilevel moderation model* used to test the moderation hypothesis (H1) in the main text (see Figure 2A). In the second simulation study, we simulated data based on the *multilevel mediated moderation model* used to test the mediated moderation hypothesis (H2) in the main text (see Figure 2B). All data was generated based on the parameter values found in the applications. For each simulation study, we generated 500 data sets (or replications). The sample size matched the sample size in our study (i.e., 95 clusters and 1.300 observations per cluster, resulting in 123500 observations in total). We ran the simulations in Mplus using Bayes estimation with a minimum number of 2500 MCMC iterations and a minimum convergence criterium of .005. This means that at least 2500 MCMC iterations were used to estimate each model. If the minimum convergence criterium of .005 (default in Mplus is .05) was not met after 2500 iterations, more iterations were used to reach the minimum convergence criterium. Preliminary analyses were performed to ensure that the simulated models converged properly while maintaining an acceptable computation time. In line with our application, we used the default priors implemented in Mplus. The statistical performance of the models was evaluated using parameter estimation bias (PEB), standard error bias (SEB), mean square error (MSE), 95% coverage rate (CO). Additionally, we computed the statistical power of all model parameters. Table 7 shows the results of the moderation model. Table 8 shows the results of the mediated moderation model.

Table 7

Monte Carlo Simulation Study: Moderation Model

	Pop_Par	AVE_Par	PEB	STD_Pop	AVE_SE	SEB	MSE	CO	POWER
$\sigma^2_{IAT_W}$	0.20	0.20	0.00	0.00	0.00	0.00	0.00	0.95	1.00
σ^2_ε	9.28	9.28	0.00	0.04	0.04	0.02	0.00	0.95	1.00
$\boldsymbol{\gamma_{11}}$	**-2.20**	**-2.23**	**0.01**	**0.61**	**0.63**	**0.02**	**0.38**	**0.95**	**0.94**
$\boldsymbol{\gamma_{01}}$	**3.39**	**3.40**	**0.00**	**0.64**	**0.70**	**0.09**	**0.41**	**0.96**	**0.99**
$\sigma_{v_{ov_1}}$	0.10	0.11	0.04	0.23	0.24	0.02	0.05	0.93	0.07
$\mathbb{E}(\boldsymbol{IAT^B_j})$	**0.29**	**0.29**	**0.00**	**0.02**	**0.02**	**-0.02**	**0.00**	**0.95**	**1.00**
γ_{00}	-0.14	-0.16	0.08	0.26	0.26	0.02	0.07	0.94	0.10
$\boldsymbol{\gamma_{10}}$	**3.75**	**3.76**	**0.00**	**0.23**	**0.23**	**0.01**	**0.05**	**0.95**	**1.00**
$\sigma^2_{IAT_B}$	0.05	0.06	0.02	0.01	0.01	0.11	0.00	0.96	1.00
$\sigma^2_{v_o}$	2.31	2.43	0.05	0.38	0.38	-0.01	0.16	0.93	1.00
$\sigma^2_{v_1}$	1.82	1.89	0.04	0.28	0.30	0.07	0.08	0.95	1.00

Note. Focal parameters of the analysis are bold. Pop_Par: population parameter; AVE_Par: average parameter estimate across all simulated data sets; PEB: relative parameter estimation bias; STD_Pop: standard deviation of the parameter estimates across all simulated data sets; AVE_SE: average standard error (here: posterior standard deviation) across all simulated data sets; SEB: standard error estimate; MSE: mean square error; CO: 95% coverage rate; POWER: statistical power from a frequentist standpoint. $\sigma^2_{IAT_W}$: within variance of the IAT measure; σ^2_ε: unexplained within variance of the direct measures; $\gamma_{11}, \gamma_{01}, \gamma_{00}, \gamma_{10}$: fixed effects (see Equations A3-A4 in Supplement 1); $\sigma_{v_{ov_1}}$: covariance of the random intercepts and random slopes; $\mathbb{E}(IAT^B_j)$: overall mean (expected value) of the latent IAT cluster means; $\sigma^2_{IAT_B}$: between variance of the IAT measure; $\sigma^2_{v_o}$: unexplained random intercept variance; $\sigma^2_{v_1}$: unexplained random slope variance.

Table 8

Monte Carlo Simulation Study: Mediated Moderation Model

	Pop_Par	AVE_Par	PEB	STD_Pop	AVE_SE	SEB	MSE	CO	POWER
σ^2_ε	9.28	9.28	0.00	0.04	0.04	0.03	0.00	0.95	1.00
$\boldsymbol{\gamma_{12}}$	**1.77**	**1.77**	**0.00**	**0.64**	**0.66**	**0.02**	**0.41**	**0.95**	**0.78**
$\boldsymbol{\gamma_{11}}$	**-1.05**	**-1.07**	**0.02**	**0.73**	**0.74**	**0.01**	**0.54**	**0.96**	**0.31**

	Pop_Par	AVE_Par	PEB	STD_Pop	AVE_SE	SEB	MSE	CO	POWER
γ_{21}	**-0.64**	**-0.63**	**-0.01**	**0.10**	**0.10**	**0.05**	**0.01**	**0.95**	**1.00**
γ_{02}	**-0.06**	**-0.04**	**-0.33**	**0.73**	**0.76**	**0.04**	**0.53**	**0.95**	**0.05**
γ_{01}	**3.36**	**3.35**	**0.00**	**0.79**	**0.86**	**0.08**	**0.63**	**0.97**	**0.97**
$\sigma_{v_{ov_1}}$	0.11	0.12	0.05	0.22	0.23	0.06	0.05	0.96	0.09
$\mathbb{E}(IAT_j^B)$	0.29	0.29	-0.01	0.02	0.02	0.06	0.00	0.96	1.00
$mean\_interc$	-0.22	-0.19	-0.14	1.09	1.13	0.03	1.20	0.96	0.06
$mean\_slope$	6.31	6.32	0.00	0.97	0.97	0.00	0.94	0.94	1.00
$\mathbb{E}(ln(\sigma_{IAT_j}^2))$	-1.44	-1.44	0.00	0.04	0.04	0.02	0.00	0.95	1.00
$\sigma_{IAT_B}^2$	0.05	0.06	0.02	0.01	0.01	0.05	0.00	0.94	1.00
$\sigma_{v_o}^2$	2.32	2.44	0.05	0.38	0.38	0.01	0.16	0.92	1.00
$\sigma_{v_1}^2$	1.70	1.79	0.05	0.28	0.29	0.03	0.09	0.94	1.00
σ_{ϑ}^2	0.05	0.05	0.03	0.01	0.01	0.03	0.00	0.94	1.00
$\gamma_{02}\,\gamma_{21}$	**0.04**	**0.03**	**-0.24**	**0.45**	**0.49**	**0.08**	**0.20**	**0.95**	**0.05**
$\gamma_{12}\gamma_{21}$	**-1.10**	**-1.10**	**0.00**	**0.44**	**0.46**	**0.06**	**0.19**	**0.96**	**0.77**

Note. Focal parameters of the analysis are bold. Pop_Par: population parameter; AVE_Par: average parameter estimate across all simulated data sets; PEB: relative parameter estimation bias; STD_Pop: standard deviation of the parameter estimates across all simulated data sets; AVE_SE: average standard error (here: posterior standard deviation) across all simulated data sets; SEB: standard error estimate; MSE: mean square error; CO: 95% coverage rate; POWER: statistical power from a frequentist standpoint. σ_ε^2: unexplained within variance of the direct measures; $\gamma_{12}, \gamma_{11}, \gamma_{21}, \gamma_{02}, \gamma_{01}$: fixed effects (see Equations A9-A10 in Supplement 1); $\gamma_{02}\,\gamma_{21}$, $\gamma_{12}\gamma_{21}$: indirect effects; $\sigma_{IAT_B}^2$: between variance of the IAT measure; $\sigma_{v_o}^2$: unexplained random intercept variance; $\sigma_{v_1}^2$: unexplained random slope variance; σ_{ϑ}^2: unexplained variability in the log-variance of the IAT measure; $\sigma_{v_{ov_1}}$: covariance of the random intercepts and random slopes; $\mathbb{E}(IAT_j^B)$: overall mean (expected value) of latent IAT cluster means; $\mathbb{E}(ln(\sigma_{IAT_j}^2))$: overall mean (expected value) of latent log-variance of the IAT measure; $mean\_interc$: mean of the random intercepts (see Equations A9-A10 in Supplement 1); $mean\_slope$: mean of the random slopes (see Equations A9-A10 in Supplement 1).

References

Cheung, M. W.-L., & Jak, S. (2016). Analyzing Big Data in Psychology: A Split/Analyze/Meta-Analyze Approach. *Frontiers in Psychology*, *7:738*, 738. https://doi.org/10.3389/fpsyg.2016.00738

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: Iii. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*(1), 17–41. https://doi.org/10.1037/a0015575

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2022). *Doing Meta-Analysis with R: A Hands-On Guide* (First edition). CRC Press.

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide* (8th ed.).

Revelle, W. (2021). *psych: Procedures for Psychological, Psychometric, and Personality Research*. https://CRAN.R-project.org/package=psych

Schuurman, N. K., Ferrer, E., Boer-Sonnenschein, M. de, & Hamaker, E. L. (2016). How to Compare Cross-Lagged Associations in a Multilevel Autoregressive Model. *Psychological Methods*, *21*(2), 206–221. https://doi.org/10.1037/met0000062

Schwarzer, G. (2007). meta: An R package for meta-analysis. *R News*, *7*(3), 40–45. https://cran.rstudio.org/doc/rnews/rnews_2007-3.pdf#page=40

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Veroniki, A. A., Jackson, D., Viechtbauer, W., Bender, R., Bowden, J., Knapp, G., Kuss, O.,

      Higgins, J. P. T., Langan, D., & Salanti, G. (2016). Methods to estimate the between-

      study variance and its uncertainty in meta-analysis. *Research Synthesis Methods*, *7*(1),

      55–79. https://doi.org/10.1002/jrsm.1164

Viechtbauer, W. (2005). Bias and Efficiency of Meta-Analytic Variance Estimators in the

      Random-Effects Model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261–

      293. https://doi.org/10.3102/10769986030003261

Supplement 4

Supplementary Results for Study 2

In this Supplement, we report additional results, including a) descriptive statistics, b) statistics demonstrating the need for multilevel models, c) unstandardized parameter estimates of the two main multilevel models, d) unstandardized estimates of relevant variables resulting from the two main multilevel models per domain, e) meta-analytical results, and f) Monte Carlo simulation studies, which we have omitted from the main text for clarity and ease of presentation of our results.

Descriptive statistics of the observed variables per domain

In what follows we report descriptive statistics of the direct attitude measures and the D scores per domain including the mean, variance, skewness, and kurtosis of the direct attitude measures (see Table 1) and the D scores (see Table 2), as well as the correlations between these variables (see Table 3). Note that we used listwise deletion as a method for handling missing values for estimating the descriptive statistics instead of Bayes which we used for the multilevel models. Therefore, the sample size for the descriptive statistics is slightly lower than for the multilevel models (n = 40,845).

Table 1

Mean, Variance, Skewness, and Kurtosis per Domain for the Direct Attitude Measures (Study 2)

Domain	Mean	Variance	Skew	Kurtosis
Receiving - Giving	1.492	7.294	0.046	-0.411
Wisdom - Innocence	-2.926	6.35	0.33	-0.135
Evolution - Creationism	-3.024	32.454	0.683	-0.822
David Letterman - Jay Leno	-0.779	14.242	0.032	-0.492
Microsoft - Apple	-0.332	18.052	0.185	-0.716
Teen Pop - Jazz	1.851	16.812	-0.349	-0.801
Julia Roberts - Meg Ryan	-0.284	10.649	0.134	-0.174
Hot - Cold	-0.553	10.455	-0.034	-0.101
Numbers - Letters	1.365	9.335	0.209	-0.383
Coffee - Tea	0.068	18.561	0.324	-0.792

Domain	Mean	Variance	Skew	Kurtosis
Gun Rights - Gun Control	3.103	24.717	-0.718	-0.439
New York - California	0.697	10.837	-0.119	-0.504
Redsox - Yankees	-0.272	14.276	-0.11	0.294
Shaq - Kobe	-1.755	9.609	-0.201	-0.21
Atheism - Religion	1.038	28.853	-0.198	-1.118
Boxers - Briefs	-2.859	17.411	0.612	-0.306
Politicians - Lawyers	1.419	5.106	0.111	0.46
Organized Labor - Management	0.283	12.732	-0.199	-0.163
Protestants - Catholics	-0.573	6.604	-0.298	0.86
Reason - Emotions	-1.044	10.728	0.026	-0.709
Speed - Accuracy	2.212	7.397	-0.205	-0.637
Tradition - Progress	1.738	9.759	-0.171	-0.24
Astrology - Science	3.58	14.584	-0.604	-0.293
Carbohydrates - Protein	0.094	9.722	0.023	-0.091
Urban - Rural	-1.001	15.988	0.259	-0.665
Plastic Surgery - Wrinkles	1.575	12.927	-0.535	-0.113
Dogs - Cats	-1.607	18.195	0.29	-0.804
Bill Clinton - Hillary Clinton	-0.899	8.357	0.322	0.074
Exercising - Relaxing	1.035	12.803	-0.04	-0.4
Meat - Vegetables	0.584	17.638	0.143	-0.651
Capital Punishment - Imprisonment	2.45	12.88	-0.461	-0.458
George Bush - John Kerry	3.137	24.382	-0.989	-0.058
Denzel Washington - Tom Cruise	-2.419	9.82	0.001	-0.565
West Coast - East Coast	0.097	12.553	-0.082	-0.53
Artists - Musicians	1.093	5.573	0.039	-0.072
Drinking - Abstaining	1.731	18.141	-0.008	-0.666
Lord of the Rings - Harry Potter	-0.411	15.828	0.199	-0.45
Prolife - Prochoice	2.867	29.504	-0.787	-0.44
Mother Teresa - Princess Diana	-1.295	9.283	0.137	-0.314
Past - Future	2.4	7.981	-0.417	0.152
Realism - Idealism	0.046	10.685	-0.081	-0.696
Hiphop - Classical	1.088	18.747	-0.056	-0.876
Pepsi - Coke	1.349	17.683	-0.349	-0.624
Flexible - Stable	-1.087	8.196	0.285	-0.239
Pants - Skirts	-1.358	11.438	0.017	-0.344
Rebellious - Conforming	-0.312	14.283	0.264	-0.576
Southerners - Northerners	1.412	9.509	-0.306	0.303
Television - Books	1.916	12.743	-0.38	-0.413
Helpers - Leaders	0.49	10.107	-0.264	-0.392
State - Church	-0.947	13.275	0.077	-0.77

Domain	Mean	Variance	Skew	Kurtosis
Burger King - McDonald's	0.081	11.996	0.051	-0.336
Poor People - Rich People	0.279	7.115	0.131	0.493
Strong - Sensitive	0.745	11.01	-0.131	-0.365
Tall People - Short People	-0.709	4.619	0.071	2.192
50 Cent - Britney Spears	-0.052	8.187	0.031	0.395
Corporations - Nonprofits	2.841	12.026	-0.347	-0.467
Muslims - Jews	1.455	6.287	0.076	1.225
Mountains - Ocean	0.648	9.667	-0.216	-0.337
Tax Reductions - Social Programs	1.553	17.885	-0.389	-0.66
Winter - Summer	2.453	16.909	-0.545	-0.464
Rich People - Beautiful People	1.589	5.106	0.198	0.318
African Americans - European Americans	0.803	4.551	-0.03	3.138
Security - Freedom	2.331	7.556	-0.303	-0.078
Traditional Values - Feminism	1.587	19.784	-0.22	-0.864
Asians - Whites	0.551	4.374	0.538	1.929
Jocks - Nerds	1.581	8.973	-0.165	0.035
Athletic People - Intelligent People	2.005	6.685	-0.377	0.162
Gay People - Straight People	1.568	7.648	0.609	0.256
Jews - Christians	0.168	5.351	0.085	1.608
Republicans - Democrats	3.035	19.143	-0.871	-0.06
Conservatives - Liberals	3.347	20.985	-0.873	-0.167
Effort - Talent	0.194	8.969	0.009	-0.316
Dramas - Comedies	0.858	10.056	-0.119	-0.346
Punishment - Forgiveness	3.33	11.986	-0.416	-0.528
Japan - United States	1.799	8.686	-0.573	-0.131
Technology - Nature	1.708	8.532	-0.235	-0.273
Friends - Family	0.637	7.54	-0.401	-0.121
National Defense - Education	4.49	8.885	-0.729	0.138
Fat People - Thin People	2.292	5.756	0.152	0.843
Career - Family	1.773	10.462	-0.456	-0.054
Money - Love	2.993	6.604	-0.721	0.924
Chaos - Order	3.751	13.235	-0.628	-0.265
Night - Morning	-1.342	15.395	0.249	-0.8
Canadian - American	-0.057	7.291	-0.116	0.632
Old People - Young People	0.957	5.921	-0.152	0.569
Avoiding - Approaching	2.834	11.36	-0.355	-0.227
Determinism - Free will	4.694	9.974	-0.827	0.248
Difficult - Simple	0.165	12.672	-0.128	-0.298
Married - Single	-1.878	14.483	0.301	-0.458
Skeptical - Trusting	1.234	14.383	-0.254	-0.42

Domain	Mean	Variance	Skew	Kurtosis
Manufactured - Natural	2.871	7.191	-0.192	0.21
Foreign Places - American Places	-0.422	9.812	-0.036	-0.098
Public - Private	3.181	7.856	-0.387	0.193
Companionship - Solitude	-1.608	10.779	-0.083	-0.456
Team - Individual	2.026	11.133	-0.315	-0.285

Table 2

Mean, Variance, Skewness, and Kurtosis of the D Scores for the IATs per Domain (Study 2)

Domain	Mean	Variance	Skew	Kurtosis
Receiving - Giving	0.001	0.148	-0.003	0.041
Wisdom - Innocence	0.006	0.211	0.18	-0.189
Evolution - Creationism	0.008	0.262	0.194	-0.638
David Letterman - Jay Leno	0.021	0.127	0.201	0.087
Microsoft - Apple	0.025	0.233	-0.003	-0.579
Teen Pop - Jazz	0.027	0.223	-0.12	-0.41
Julia Roberts - Meg Ryan	0.03	0.156	0.072	-0.352
Hot - Cold	0.031	0.236	-0.196	-0.475
Numbers - Letters	0.032	0.211	-0.177	-0.419
Coffee - Tea	0.039	0.237	-0.041	-0.163
Gun Rights - Gun Control	0.042	0.165	-0.28	-0.302
New York - California	0.044	0.216	0.054	-0.512
Redsox - Yankees	0.047	0.219	0.029	-0.333
Shaq - Kobe	0.048	0.115	-0.094	-0.434
Atheism - Religion	0.054	0.28	-0.111	-0.712
Boxers - Briefs	0.056	0.253	-0.056	-0.297
Politicians - Lawyers	0.058	0.125	-0.002	0.022
Organized Labor - Management	0.059	0.207	0.013	-0.511
Protestants - Catholics	0.06	0.246	0.099	-0.576
Reason - Emotions	0.063	0.193	-0.264	0.084
Speed - Accuracy	0.065	0.161	-0.015	-0.228
Tradition - Progress	0.065	0.195	0.144	-0.278
Astrology - Science	0.073	0.222	-0.058	-0.595
Carbohydrates - Protein	0.079	0.229	-0.056	-0.439
Urban - Rural	0.082	0.258	-0.094	-0.626
Plastic Surgery - Wrinkles	0.083	0.153	-0.299	-0.423
Dogs - Cats	0.083	0.231	-0.139	-0.426
Bill Clinton - Hillary Clinton	0.086	0.155	0.044	-0.21
Exercising - Relaxing	0.087	0.235	-0.091	-0.653
Meat - Vegetables	0.088	0.238	-0.351	-0.308

Domain	Mean	Variance	Skew	Kurtosis
Capital Punishment - Imprisonment	0.09	0.121	-0.027	-0.192
George Bush - John Kerry	0.103	0.198	-0.085	-0.442
Denzel Washington - Tom Cruise	0.115	0.158	-0.006	0.038
West Coast - East Coast	0.117	0.329	-0.213	-0.718
Artists - Musicians	0.138	0.17	-0.058	-0.044
Drinking - Abstaining	0.14	0.215	-0.138	-0.38
Lord of the Rings - Harry Potter	0.14	0.233	-0.22	-0.696
Prolife - Prochoice	0.141	0.176	-0.457	-0.165
Mother Teresa - Princess Diana	0.147	0.15	-0.264	-0.192
Past - Future	0.151	0.171	-0.193	-0.187
Realism - Idealism	0.152	0.256	-0.189	-0.309
Hiphop - Classical	0.153	0.235	-0.16	-0.496
Pepsi - Coke	0.163	0.214	-0.241	-0.258
Flexible - Stable	0.164	0.194	-0.006	-0.566
Pants - Skirts	0.175	0.238	-0.362	-0.225
Rebellious - Conforming	0.18	0.239	-0.052	-0.407
Southerners - Northerners	0.191	0.274	-0.232	-0.405
Television - Books	0.195	0.228	-0.243	-0.433
Helpers - Leaders	0.196	0.224	-0.144	-0.45
State - Church	0.2	0.215	-0.182	-0.447
Burger King - McDonald's	0.205	0.131	-0.189	-0.503
Poor People - Rich People	0.212	0.183	-0.227	-0.229
Strong - Sensitive	0.215	0.232	-0.323	-0.333
Tall People - Short People	0.216	0.265	-0.154	-0.405
50 Cent - Britney Spears	0.216	0.124	-0.174	-0.256
Corporations - Nonprofits	0.226	0.214	-0.257	-0.398
Muslims - Jews	0.227	0.137	-0.143	-0.036
Mountains - Ocean	0.235	0.181	-0.237	-0.336
Tax Reductions - Social Programs	0.236	0.197	-0.272	-0.371
Winter - Summer	0.245	0.231	-0.27	-0.38
Rich People - Beautiful People	0.247	0.118	-0.418	0.504
African Americans - European Americans	0.253	0.192	-0.409	-0.258
Security - Freedom	0.266	0.15	-0.352	0.345
Traditional Values - Feminism	0.27	0.233	-0.417	-0.209
Asians - Whites	0.272	0.209	-0.303	-0.018
Jocks - Nerds	0.286	0.195	-0.542	0.372
Athletic People - Intelligent People	0.287	0.172	-0.387	-0.176
Gay People - Straight People	0.312	0.224	-0.477	-0.05
Jews - Christians	0.317	0.177	-0.705	1.106
Republicans - Democrats	0.318	0.268	-0.555	-0.176

Domain	Mean	Variance	Skew	Kurtosis
Conservatives - Liberals	0.335	0.309	-0.674	-0.164
Effort - Talent	0.346	0.168	-0.288	0.003
Dramas - Comedies	0.352	0.175	-0.384	0.033
Punishment - Forgiveness	0.358	0.11	-0.306	0.147
Japan - United States	0.361	0.169	-0.377	0.229
Technology - Nature	0.385	0.159	-0.307	-0.262
Friends - Family	0.39	0.197	-0.556	0.118
National Defense - Education	0.393	0.173	-0.521	0.335
Fat People - Thin People	0.407	0.166	-0.386	0.008
Career - Family	0.413	0.206	-0.558	0.004
Money - Love	0.421	0.159	-0.538	0.188
Chaos - Order	0.426	0.19	-0.491	-0.095
Night - Morning	0.44	0.194	-0.254	-0.277
Canadian - American	0.44	0.217	-0.842	0.56
Old People - Young People	0.453	0.148	-0.397	0.118
Avoiding - Approaching	0.472	0.161	-0.635	0.69
Determinism - Free will	0.476	0.121	-0.309	-0.093
Difficult - Simple	0.48	0.182	-0.574	0.339
Married - Single	0.482	0.254	-0.556	-0.319
Skeptical - Trusting	0.493	0.181	-0.457	-0.056
Manufactured - Natural	0.511	0.122	-0.349	0.232
Foreign Places - American Places	0.539	0.186	-0.603	-0.094
Public - Private	0.606	0.201	-0.756	0.736
Companionship - Solitude	0.641	0.241	-0.847	0.491
Team - Individual	0.97	0.111	-1.056	2.431

Table 3

Correlations Between the Observed Variables Across Domains (Study 2)

Variable	1	2	3	4	5	6	7	8
1. Mean D score								
2. Variance D score	-.23*							
	[-.41, -.03]							
3. Skew D score	-.85**	.05						
	[-.90, -.78]	[-.15, .25]						
4. Kurtosis D score	.70**	-.46**	-.68**					
	[.59, .79]	[-.60, -.28]	[-.78, -.56]					
5. Mean direct att[a]	.22*	-.18	-.18	.17				
	[.01, .40]	[-.37, .03]	[-.36, .03]	[-.03, .36]				
6. Variance direct att	-.26*	.41**	.13	-.34**	.20*			
	[-.44, -.06]	[.22, .56]	[-.07, .32]	[-.51, -.15]	[.00, .39]			
7. Skew direct att	-.25*	.13	.33**	-.13	-.52**	-.16		
	[-.43, -.05]	[-.07, .32]	[.14, .50]	[-.32, .07]	[-.66, -.36]	[-.35, .04]		
8. Kurtosis direct att	.20*	-.17	-.20	.24*	-.07	-.58**	.03	
	[.00, .39]	[-.36, .03]	[-.39, .00]	[.04, .42]	[-.27, .13]	[-.70, -.43]	[-.17, .23]	
9. I-E cor	-.40**	.54**	.26*	-.50**	-.05	.71**	-.10	-.32**
	[-.55, -.21]	[.39, .67]	[.06, .43]	[-.64, -.33]	[-.25, .15]	[.59, .80]	[-.30, .10]	[-.49, -.13]

Note. att = attitude; I-E cor = implicit explicit correlation. Values in square brackets indicate the 95% confidence interval for each correlation.

[a] We took the absolute value of the means of the direct attitude measures so that the values ranged from 0 to 9.

* indicates $p < .05$. ** indicates $p < .01$.

Assessing the need for multilevel models

The ICC of both, identity IAT scores and direct attitude scores, was substantial, $ICC_{identity\ IAT} = 0.13$ and $ICC_{direct\ attitude} = 0.19$, respectively, suggesting a considerable dependency in the data. Furthermore, the variance of the random intercepts and the random slopes in the unconditional random intercept and random slope model was substantial as well, $\hat{\sigma}_I^2 = 2.90$, with a 95% credible interval of *C.I.* [2.18, 3.91], and $\hat{\sigma}_S^2 = 2.35$, with a 95% credible interval of *C.I.* [1.75, 3.26], respectively. This result also indicates strong variation in the data across clusters. Both results further suggest that consistent with our proposed models, Level 2 predictors should be considered.
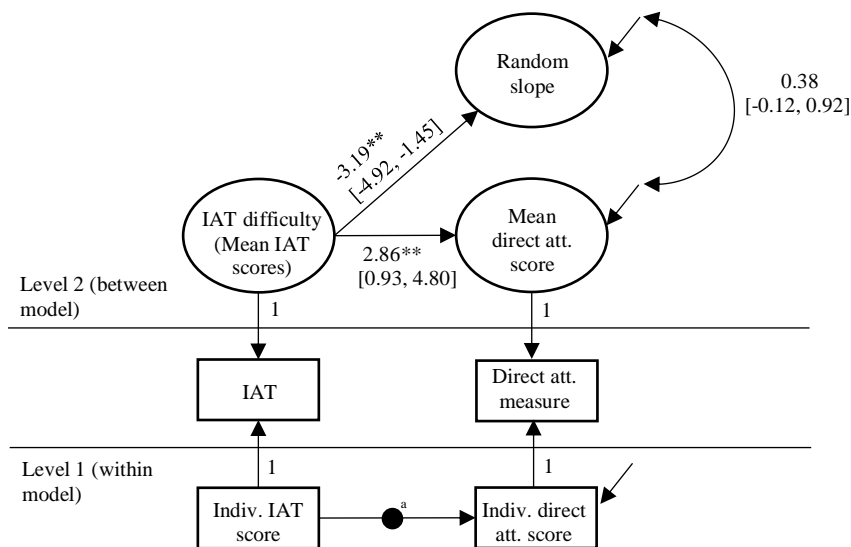
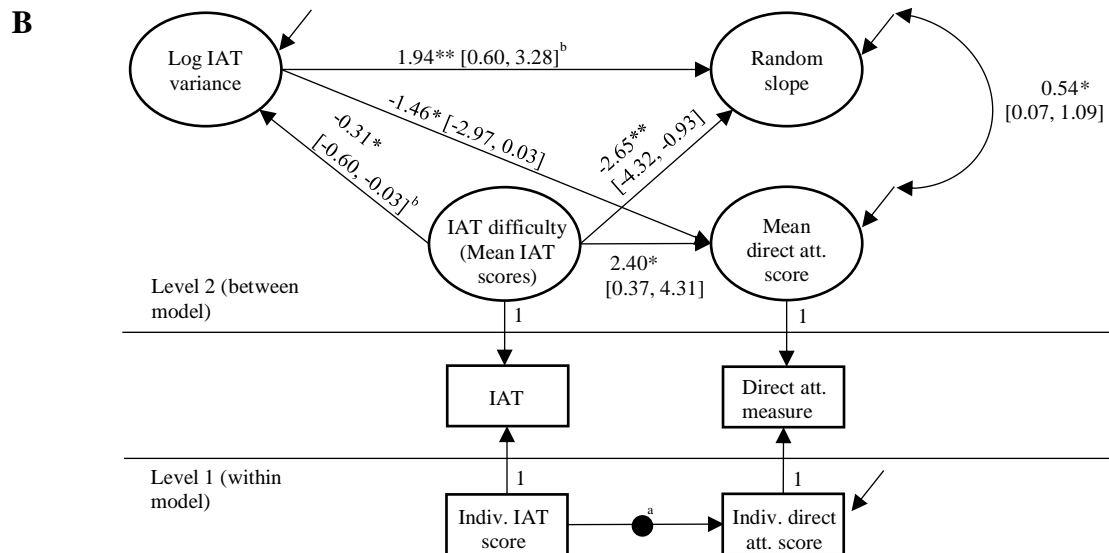**Unstandardized parameter estimates of the two main multilevel models**

Results of the unstandardized parameter estimates of the multilevel models are presented in Figure 1.

Figure 1

Unstandardized Parameter Estimates for the Multilevel Moderation Model (Panel A) and the Multilevel Mediated Moderation Model (Panel B) in Study 2

A

B

Note. Circles represent latent and rectangles observed variables. Numbers without brackets ascribed to single-headed arrows are path coefficients, numbers without brackets ascribed to double-headed arrows are covariances, and numbers in square brackets are 95% credible intervals. Darkened circles represent random slopes which are also depicted on Level 2. IAT = Implicit Association Test; Log IAT variance = Log IAT true-score variance; att = attitude; Indiv = individual.

[a] Mplus does not provide the unstandardized within effect across all cluster (Muthén & Muthén, 1998-2017; Schuurman et al., 2016).

[b] The indirect effect of IAT difficulty on the random slope via log IAT true-score variance was significant, $B = -0.56$, p$SD = 0.36$, 95% *C.I.* [-1.40, -0.03].

*p < .05, **p < .01, ***p < .001.

Unstandardized estimates of the most relevant variables resulting from the two main multilevel models per domain

For a better overview of the test difficulty of the identity IATs, the true-score variance of the identity IATs and their relationship with the direct attitude measures we present below the corresponding unstandardized estimates resulting from our two main multilevel models. Table 4 displays the test difficulty of the IATs and their relationship with the direct attitude measures based on the multilevel moderation model. Table 5 displays the test difficulty of the

IATs, the true-score variance of the IATs and their relationship with the direct attitude

measures based on the multilevel mediated moderation model.

Table 4

Domains and Their Corresponding IAT Test Difficulty and Slope Estimated From the

Multilevel Moderation Model (Study 2)

Domain	IAT difficulty[a]	Slope[b]
Redsox - Yankees	0.002	5.158
Hot - Cold	0.003	2.222
Julia Roberts - Meg Ryan	0.007	3.998
Numbers - Letters	0.010	2.116
Microsoft - Apple	0.012	4.809
Wisdom - Innocence	0.02	1.445
Evolution - Creationism	0.025	6.595
New York - California	0.026	4.104
Gun Rights - Gun Control	0.029	6.445
Receiving - Giving	0.032	1.456
Teen Pop - Jazz	0.034	4.970
David Letterman - Jay Leno	0.035	5.255
Reason - Emotions	0.040	2.806
Atheism - Religion	0.049	6.523
Tradition - Progress	0.054	3.351
Urban - Rural	0.062	3.209
Exercising - Relaxing	0.069	2.782
Politicians - Lawyers	0.069	1.677
Protestants - Catholics	0.069	2.317
Organized Labor - Management	0.074	3.445
Bill Clinton - Hillary Clinton	0.078	2.338
Boxers - Briefs	0.083	3.651
Denzel Washington - Tom Cruise	0.088	1.851
Carbohydrates - Protein	0.092	2.831
Coffee - Tea	0.093	4.002
Speed - Accuracy	0.094	1.529
Shaq - Kobe	0.096	2.295
Dogs - Cats	0.100	5.009
Astrology - Science	0.102	3.924
Plastic Surgery - Wrinkles	0.109	3.425
Prolife - Prochoice	0.119	7.900
West Coast - East Coast	0.121	3.498

Domain	IAT difficulty[a]	Slope[b]
Meat - Vegetables	0.123	3.841
Capital Punishment - Imprisonment	0.124	2.376
George Bush - John Kerry	0.127	6.090
Flexible - Stable	0.128	1.788
Drinking - Abstaining	0.135	4.138
Hiphop - Classical	0.140	5.095
Lord of the Rings - Harry Potter	0.140	4.718
Artists - Musicians	0.143	2.143
Past - Future	0.143	0.360
Realism - Idealism	0.159	2.904
Helpers - Leaders	0.161	1.809
Mother Teresa - Princess Diana	0.166	0.985
Pepsi - Coke	0.171	4.876
Southerners - Northerners	0.175	2.667
Television - Books	0.176	3.201
Pants - Skirts	0.189	1.167
State - Church	0.189	4.148
Rebellious - Conforming	0.199	3.388
Strong - Sensitive	0.206	2.897
Corporations - Nonprofits	0.212	3.280
Mountains - Ocean	0.212	3.549
50 Cent - Britney Spears	0.219	2.195
Poor People - Rich People	0.220	2.252
Burger King - McDonald's	0.226	4.003
Traditional Values - Feminism	0.229	5.395
Rich People - Beautiful People	0.242	0.783
Tax Reductions - Social Programs	0.243	3.994
African Americans - European Americans	0.247	2.031
Muslims - Jews	0.254	3.426
Tall People - Short People	0.257	0.987
Security - Freedom	0.261	2.063
Winter - Summer	0.264	3.021
Asians - Whites	0.277	0.626
Athletic People - Intelligent People	0.282	2.777
Gay People - Straight People	0.300	1.895
Jocks - Nerds	0.305	3.309
Japan - United States	0.317	2.956
Punishment - Forgiveness	0.321	2.462
Jews - Christians	0.325	2.142
Republicans - Democrats	0.339	5.497
Effort - Talent	0.351	2.307

Domain	IAT difficulty[a]	Slope[b]
Dramas - Comedies	0.357	2.886
Conservatives - Liberals	0.374	6.214
Friends - Family	0.376	2.227
Technology - Nature	0.383	2.425
Fat People - Thin People	0.388	1.637
Money - Love	0.394	1.275
Chaos - Order	0.404	3.247
Career - Family	0.410	2.580
National Defense - Education	0.412	1.957
Night - Morning	0.441	4.140
Determinism - Free will	0.451	2.901
Difficult - Simple	0.454	2.459
Old People - Young People	0.457	1.088
Canadian - American	0.479	2.232
Married - Single	0.481	2.332
Avoiding - Approaching	0.482	1.178
Skeptical - Trusting	0.487	2.969
Manufactured - Natural	0.504	2.266
Foreign Places - American Places	0.531	2.454
Public - Private	0.581	0.747
Companionship - Solitude	0.643	1.491
Team - Individual	0.927	1.099

Note. IAT = Implicit Association Test.

[a] IAT difficulty corresponds to the latent mean IAT scores per domain.

[b] The slope corresponds to the within effect of attitude IAT scores on direct attitude scores per domain.

Table 5

Domains and Their Corresponding IAT Difficulty, IAT True-Score Variance and Slope

Estimated From the Multilevel Mediated Moderation Model (Study 2)

Domain	IAT difficulty[a]	Log IAT variance[b]	Slope[c]
Julia Roberts - Meg Ryan	-0.001[d]	-1.779	3.313
Boxers - Briefs	0.002	-1.335	3.489
Wisdom - Innocence	0.009	-1.550	0.994
Evolution - Creationism	0.011	-1.290	7.295
David Letterman - Jay Leno	0.020	-1.992	5.346

Domain	IAT difficulty[a]	Log IAT variance[b]	Slope[c]
Numbers - Letters	0.026	-1.569	2.534
Hot - Cold	0.031	-1.463	1.736
Reason - Emotions	0.031	-1.619	2.961
Protestants - Catholics	0.033	-1.406	2.163
Receiving - Giving	0.034	-2.017	1.450
Teen Pop - Jazz	0.034	-1.493	4.421
Tradition - Progress	0.039	-1.597	2.515
Shaq - Kobe	0.044	-2.117	2.011
Microsoft - Apple	0.047	-1.423	4.889
Gun Rights - Gun Control	0.052	-1.766	6.440
New York - California	0.055	-1.583	3.294
Atheism - Religion	0.057	-1.298	6.336
Astrology - Science	0.058	-1.466	3.273
Politicians - Lawyers	0.064	-2.074	2.356
Urban - Rural	0.064	-1.428	3.604
Speed - Accuracy	0.065	-1.723	1.230
Coffee - Tea	0.067	-1.505	4.433
Organized Labor - Management	0.072	-1.642	3.787
Bill Clinton - Hillary Clinton	0.073	-1.869	2.056
Exercising - Relaxing	0.073	-1.501	2.799
Carbohydrates - Protein	0.080	-1.492	2.839
Plastic Surgery - Wrinkles	0.081	-1.907	3.930
Meat - Vegetables	0.085	-1.491	3.540
Redsox - Yankees	0.094	-1.435	4.707
George Bush - John Kerry	0.100	-1.621	6.363
Capital Punishment - Imprisonment	0.112	-2.027	0.701
Denzel Washington - Tom Cruise	0.114	-1.781	3.145
West Coast - East Coast	0.114	-1.201	3.836
Mother Teresa - Princess Diana	0.119	-1.903	1.486
Hiphop - Classical	0.131	-1.508	5.259
Drinking - Abstaining	0.132	-1.569	4.353
Dogs - Cats	0.133	-1.506	5.333
Prolife - Prochoice	0.133	-1.858	7.628
Television - Books	0.137	-1.466	3.160
Lord of the Rings - Harry Potter	0.140	-1.394	4.621
Past - Future	0.143	-1.775	1.874
Realism - Idealism	0.151	-1.449	3.138
Flexible - Stable	0.159	-1.580	1.957
Artists - Musicians	0.162	-1.732	1.888
Pants - Skirts	0.164	-1.458	1.154

Domain	IAT difficulty[a]	Log IAT variance[b]	Slope[c]
50 Cent - Britney Spears	0.169	-2.016	1.844
Pepsi - Coke	0.169	-1.538	5.087
Southerners - Northerners	0.172	-1.310	2.746
Helpers - Leaders	0.179	-1.530	2.464
State - Church	0.187	-1.486	3.498
Rebellious - Conforming	0.196	-1.487	3.122
Muslims - Jews	0.209	-1.974	3.270
Poor People - Rich People	0.210	-1.670	2.634
Corporations - Nonprofits	0.225	-1.534	3.232
Winter - Summer	0.225	-1.556	4.048
Burger King - McDonald's	0.228	-2.108	3.378
Tall People - Short People	0.230	-1.355	0.910
African Americans - European Americans	0.237	-1.596	2.277
Security - Freedom	0.239	-1.923	2.395
Traditional Values - Feminism	0.239	-1.466	5.057
Tax Reductions - Social Programs	0.244	-1.716	3.451
Strong - Sensitive	0.245	-1.476	2.671
Mountains - Ocean	0.248	-1.759	3.614
Rich People - Beautiful People	0.260	-2.149	0.411
Jocks - Nerds	0.261	-1.695	3.588
Asians - Whites	0.297	-1.587	1.117
Athletic People - Intelligent People	0.297	-1.626	3.526
Jews - Christians	0.306	-1.733	2.069
Gay People - Straight People	0.313	-1.564	2.149
Republicans - Democrats	0.315	-1.404	6.104
Japan - United States	0.329	-1.805	2.661
Punishment - Forgiveness	0.329	-2.163	2.996
Conservatives - Liberals	0.352	-1.260	5.659
Effort - Talent	0.353	-1.922	2.319
Dramas - Comedies	0.357	-1.781	2.698
Friends - Family	0.385	-1.617	1.768
Money - Love	0.387	-1.875	1.713
Chaos - Order	0.398	-1.731	3.996
Technology - Nature	0.409	-1.839	2.463
Fat People - Thin People	0.410	-1.805	2.012
National Defense - Education	0.424	-1.868	2.054
Career - Family	0.433	-1.575	2.538
Night - Morning	0.442	-1.720	3.259
Old People - Young People	0.443	-1.978	1.502
Canadian - American	0.450	-1.484	2.005

Domain	IAT difficulty[a]	Log IAT variance[b]	Slope[c]
Avoiding - Approaching	0.466	-1.733	0.801
Married - Single	0.478	-1.477	1.475
Determinism - Free will	0.485	-2.149	3.062
Difficult - Simple	0.486	-1.789	2.882
Manufactured - Natural	0.501	-2.114	1.712
Skeptical - Trusting	0.510	-1.719	2.630
Foreign Places - American Places	0.520	-1.596	2.578
Public - Private	0.544	-1.640	1.511
Companionship - Solitude	0.652	-1.663	1.266
Team - Individual	0.968	-2.155	1.636

Note. IAT = Implicit Association Test.

[a] IAT difficulty corresponds to the latent mean IAT scores per domain.

[b] Log IAT variance corresponds to the log true-score variance of the IAT scores per domain; the closer log IAT true-score variance is to zero the higher the true-score variance.

[c] The slope corresponds to the within effect of attitude IAT scores on direct attitude scores per domain.

[d] The latent mean IAT score was estimated to be negative, which is due to centering and therefore does not affect the interpretation of the relationship between IAT difficulty and the unstandardized slopes.

Meta-Analytical Results

For better comparisons with previous studies testing moderators of IAT relations we additionally conducted a meta-analyses.

Moderation hypothesis (H1): In order to conduct a meta-analysis we treated the 95 domains as 95 studies resulting in 95 IAT-direct attitude correlation coefficients following the same approach we already described for Study 1 (Supplement 3). The final results were as follows: The pooled association between identity IAT scores and direct attitude scores was $r = .39$, $p < .001$, 95% *CI* [.36, .42]. The between-study heterogeneity variance was estimated at $\tau^2 = 0.03$, 95% *CI* [0.02, 0.037], with an I^2 value of 91.6%, 95% CI [90.3%, 92.7%]. Accordingly, moderator analysis was warranted. IAT test difficulty explained 14.89% of the

heterogeneity in effect sizes. Unsurprisingly, the test of moderators was significant, $QM(1) =$ 15.84, $p < .001$, as well as the unstandardized regression weight of IAT test difficulty, $B = -0.37$, $p < .001$, 95% CI [-0.55, -0.19]. The analysis was done in R using the meta package (Schwarzer, 2007).

Mediated moderation hypothesis (H2): We proceeded as outlined above only this time we included IAT test difficulty as well as IAT total variance as moderators into the meta-regression. Accordingly, the results corresponded to the previous results except for the moderation analysis. IAT test difficulty and IAT total variance together explained 39.67% of the heterogeneity in effect sizes. Unsurprisingly, the test of moderators was significant once again, $QM(2) = 55.48$, $p < .001$. The unstandardized regression weight of IAT total variance was significant, $B = 1.81$, $p < .001$, 95% CI [1.2, 2.41], as well as the unstandardized regression weight of IAT test difficulty, $B = -0.26$, $p < .01$, 95% CI [-0.42, -0.1]. This time, compared to Study 1, the result is consistent with the multilevel approach and suggest a partial mediation. However, the meta-analytical analysis did not allow for estimating indirect effects so that conclusions concerning the mediation hypothesis should be treated with caution. Again the analysis was done in R using the meta package (Schwarzer, 2007).

Monte Carlo Simulations

We run two Monte Carlo simulation studies to evaluate the statistical performance of the multilevel models used in the application in Study 2. In the first simulation study, we simulated data based on the *multilevel moderation model* used to test the moderation hypothesis (H1) in the main text (see Figure 5A). In the second simulation study, we simulated data based on the *multilevel mediated moderation model* used to test the mediated moderation hypothesis (H2) in the main text (see Figure 5B). All data was generated based on the parameter values found in the applications. For each simulation study, we generated 500 data sets (or replications). The sample size matched the sample size in our study (i.e., 95 clusters and 460 observations per cluster, resulting in 43700 observations in total). We used

the same implementations features and the same estimates to evaluate the statistical performance of the models as for the simulations in Study 1. Table 6 shows the results of the moderation model. Table 7 shows the results of the mediated moderation model.

Table 6

Monte Carlo Simulation Study: Moderation Model

	Pop_Par	AVE_Par	PEB	STD_Pop	AVE_SE	SEB	MSE	CO	POWER
$\sigma^2_{IAT_W}$	0.20	0.20	0.00	0.00	0.00	0.00	0.00	0.95	1.00
σ^2_ε	9.49	9.49	0.00	0.06	0.06	0.03	0.00	0.95	1.00
γ_{11}	**-3.19**	**-3.15**	**-0.01**	**0.86**	**0.90**	**0.04**	**0.73**	**0.96**	**0.94**
γ_{01}	**2.86**	**2.90**	**0.01**	**0.98**	**1.00**	**0.01**	**0.97**	**0.95**	**0.84**
$\sigma_{v_{ov_1}}$	0.38	0.37	-0.02	0.26	0.28	0.05	0.07	0.95	0.31
$\mathbb{E}(IAT^B_j)$	**0.22**	**0.22**	**0.00**	**0.02**	**0.02**	**0.04**	**0.00**	**0.96**	**1.00**
γ_{00}	0.20	0.20	0.00	0.27	0.28	0.02	0.08	0.95	0.11
γ_{10}	**3.76**	**3.75**	**0.00**	**0.25**	**0.25**	**0.03**	**0.06**	**0.96**	**1.00**
$\sigma^2_{IAT_B}$	0.03	0.03	0.03	0.00	0.01	0.08	0.00	0.95	1.00
$\sigma^2_{v_o}$	2.67	2.77	0.03	0.41	0.43	0.07	0.17	0.95	1.00
$\sigma^2_{v_1}$	2.04	2.15	0.05	0.33	0.35	0.05	0.12	0.93	1.00

Note. Focal parameters of the analysis are bold. Pop_Par: population parameter; AVE_Par: average parameter estimate across all simulated data sets; PEB: relative parameter estimation bias; STD_Pop: standard deviation of the parameter estimates across all simulated data sets; AVE_SE: average standard error (here: posterior standard deviation) across all simulated data sets; SEB: standard error estimate; MSE: mean square error; CO: 95% coverage rate; POWER: statistical power from a frequentist standpoint. $\sigma^2_{IAT_W}$: within variance of the IAT measure; σ^2_ε: unexplained within variance of the direct measures; $\gamma_{11}, \gamma_{01}, \gamma_{00}, \gamma_{10}$: fixed effects (see Equations A3-A4 in Supplement 1); $\sigma_{v_{ov_1}}$: covariance of the random intercepts and random slopes; $\mathbb{E}(IAT^B_j)$: overall mean (expected value) of the latent IAT cluster means; $\sigma^2_{IAT_B}$: between variance of the IAT measure; $\sigma^2_{v_o}$: unexplained random intercept variance; $\sigma^2_{v_1}$: unexplained random slope variance.

Table 7

Monte Carlo Simulation Study: Mediated Moderation Model

	Pop_Pa	AVE_Pa	PEB	STD_Pop	AVE_SE	SEB	MSE	CO	POWER
σ_ε^2	9.50	9.50	0.00	0.07	0.06	-0.04	0.00	0.92	1.00
γ_{12}	**1.94**	**1.94**	**0.00**	**0.65**	**0.69**	**0.06**	**0.42**	**0.96**	**0.81**
γ_{11}	**-2.65**	**-2.61**	**-0.01**	**0.86**	**0.88**	**0.02**	**0.74**	**0.95**	**0.86**
γ_{21}	**-0.31**	**-0.31**	**0.00**	**0.14**	**0.14**	**0.03**	**0.02**	**0.94**	**0.58**
γ_{02}	**-1.46**	**-1.46**	**0.00**	**0.73**	**0.79**	**0.07**	**0.54**	**0.96**	**0.45**
γ_{01}	**2.40**	**2.43**	**0.01**	**0.94**	**1.00**	**0.06**	**0.88**	**0.97**	**0.68**
$\sigma_{v_{ov_1}}$	0.54	0.56	0.04	0.25	0.27	0.06	0.06	0.94	0.62
$\mathbb{E}(IAT_j^B)$	0.22	0.22	0.00	0.02	0.02	0.02	0.00	0.95	1.00
$mean\_interc$	-2.13	-2.12	0.00	1.20	1.28	0.06	1.45	0.95	0.36
$mean\_slope$	6.83	6.82	0.00	1.05	1.12	0.06	1.10	0.96	1.00
$\mathbb{E}(ln(\sigma_{IAT_j}^2))$	-1.59	-1.59	0.00	0.04	0.04	0.01	0.00	0.94	1.00
$\sigma_{IAT_B}^2$	0.03	0.03	0.04	0.00	0.01	0.06	0.00	0.94	1.00
$\sigma_{v_o}^2$	2.56	2.67	0.04	0.40	0.42	0.05	0.18	0.94	1.00
$\sigma_{v_1}^2$	1.85	1.95	0.05	0.31	0.33	0.04	0.11	0.95	1.00
σ_ϑ^2	0.05	0.05	0.02	0.01	0.01	-0.03	0.00	0.94	1.00
$\gamma_{02}\,\gamma_{21}$	**0.40**	**0.41**	**0.03**	**0.32**	**0.36**	**0.10**	**0.10**	**0.96**	**0.20**
$\gamma_{12}\gamma_{21}$	**-0.56**	**-0.57**	**0.01**	**0.35**	**0.37**	**0.07**	**0.12**	**0.95**	**0.44**

Note. Focal parameters of the analysis are bold. Pop_Par: population parameter; AVE_Par: average parameter estimate across all simulated data sets; PEB: relative parameter estimation bias; STD_Pop: standard deviation of the parameter estimates across all simulated data sets; AVE_SE: average standard error (here: posterior standard deviation) across all simulated data sets; SEB: standard error estimate; MSE: mean square error; CO: 95% coverage rate; POWER: statistical power from a frequentist standpoint. σ_ε^2: unexplained within variance of the direct measures; $\gamma_{12}, \gamma_{11}, \gamma_{21}, \gamma_{02}, \gamma_{01}$: fixed effects (see Equations A9-A10 in Supplement 1); $\gamma_{02}\,\gamma_{21}, \gamma_{12}\gamma_{21}$: indirect effects; $\sigma_{IAT_B}^2$: between variance of the IAT measure; $\sigma_{v_o}^2$: unexplained random intercept variance; $\sigma_{v_1}^2$: unexplained random slope variance; σ_ϑ^2: unexplained variability in the log-variance of the IAT measure; $\sigma_{v_{ov_1}}$: covariance of the random intercepts and random slopes; $\mathbb{E}(IAT_j^B)$: overall mean (expected value) of latent IAT cluster means; $\mathbb{E}(ln(\sigma_{IAT_j}^2))$: overall mean (expected value) of latent log-variance of the IAT

measure; $mean\_interc$: mean of the random intercepts (see Equations A9-A10 in Supplement 1); $mean\_slope$: mean of the random slopes (see Equations A9-A10 in Supplement 1).

References

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide* (8th ed.).

Schuurman, N. K., Ferrer, E., Boer-Sonnenschein, M. de, & Hamaker, E. L. (2016). How to

Compare Cross-Lagged Associations in a Multilevel Autoregressive Model.

Psychological Methods, *21*(2), 206–221. https://doi.org/10.1037/met0000062

Schwarzer, G. (2007). meta: An R package for meta-analysis. *R News*, *7*(3), 40–45.

https://cran.rstudio.org/doc/rnews/rnews_2007-3.pdf#page=40

Supplement 5

Supplementary Results for Study 3

In this Supplement, we report additional results, including a) descriptive statistics, b) model

comparisons to test for measurement invariance (MI) in our main multigroup SEM analyses,

c) multigroup SEM with three indicators for the latent IAT factor, and d) multigroup SEM

with recoding as a third variable, which we have omitted from the main text for clarity and

ease of presentation of our results.

Descriptive statistics of the observed variables

These descriptive statistics include the mean, variance, skewness, and kurtosis of the

direct attitude items (see Table 1) and the D scores (see Table 2) per group, as well as the

correlations between the direct attitude items and the D scores per group (see Table 3-5).

Table 1

Direct attitude items and Their Corresponding Mean, Variance, Skew, and Kurtosis per

Group (Study 3)

Group	Item	Mean	Variance	Skew	Kurtosis
Degradation IAT	GR env. prot.	8.93	1.49	-1.30	2.66
(n = 160)	GR env. degr.	1.56	0.95	3.51	20.31
	GR war	1.31	0.44	2.14	3.76
	GR leisure time	9.25	1.01	-1.18	0.63
	AF env. prot.	8.90	1.54	-1.84	6.00
	AF env. degr.	1.40	0.52	1.86	3.50
	AF war	1.27	0.93	5.75	42.51
	AF leisure time	9.09	1.35	-2.51	12.93
War IAT	GR env. prot.	8.78	1.74	-1.34	2.35
(n = 160)	GR env. degr.	1.82	1.94	2.58	7.81
	GR war	1.63	1.86	3.04	11.19
	GR leisure time	9.32	0.89	-1.16	0.37
	AF env. prot.	8.79	1.51	-0.92	0.33
	AF env. degr.	1.60	1.11	2.38	8.08
	AF war	1.62	1.80	3.13	12.35
	AF leisure time	9.04	1.25	-0.94	0.14
Leisure IAT	GR env. prot.	8.56	1.47	-0.56	0.13
(n = 160)	GR env. degr.	1.69	0.83	1.54	2.76

Group	Item	Mean	Variance	Skew	Kurtosis
	GR war	1.32	0.66	3.54	16.14
	GR leisure time	9.22	0.98	-1.08	0.29
	AF env. prot.	8.62	2.15	-2.09	7.61
	AF env. degr.	1.58	0.80	2.07	5.28
	AF war	1.29	0.51	3.00	9.98
	AF leisure time	9.15	1.08	-1.43	2.21

Note. IAT = Implicit Association Test; GR = gut reactions; env. prot. = environmental

protection; env. degr. = environmental degradation; AF = actual feelings.

Table 2

D Scores and Their Corresponding Mean, Variance, Skew, and Kurtosis per group (Study 3)

Group	D score	Mean	Variance	Skew	Kurtosis
Degradation IAT	D score 1	1.01	0.07	-0.34	0.69
(n = 160)	D score 2	0.90	0.09	-0.61	0.63
	D score mean	0.95	0.06	-0.51	0.59
War IAT	D score 1	0.84	0.14	-0.79	0.65
(n = 160)	D score 2	0.68	0.11	-0.69	0.63
	D score mean	0.76	0.09	-0.81	1.33
Leisure IAT	D score 1	-0.13	0.33	0.2	-0.74
(n = 160)	D score 2	-0.17	0.32	0.1	-1.01
	D score mean	-0.15	0.27	0.1	-0.84

Note. IAT = Implicit Association Test; D score 1 = IAT D score calculated via the short test

blocks; D score 2 = IAT D score calculated via the long test blocks; D score mean = IAT D

score calculated via both the short and the long test blocks.

Table 3

Correlations Between the Observed Variables in the Degradation IAT Group (Study 3)

Variable	1	2	3	4	5	6	7	8	9	10
1. GR env. prot.										
2. GR env. degr.	-.43**									
	[-.55, -.30]									
3. GR war	-.21**	.32**								
	[-.36, -.06]	[.17, .45]								
4. GR leisure time	.14	-.12	-.08							
	[-.02, .29]	[-.27, .03]	[-.23, .08]							
5. AF env. prot.	.51**	-.37**	-.20*	.08						
	[.38, .61]	[-.50, -.23]	[-.34, -.05]	[-.08, .23]						
6. AF env. degr.	-.38**	.50**	.33**	-.03	-.36**					
	[-.51, -.24]	[.38, .61]	[.19, .47]	[-.19, .12]	[-.49, -.22]					
7. AF war	-.03	.09	.46**	-.12	-.12	.13				
	[-.18, .13]	[-.07, .24]	[.33, .57]	[-.27, .03]	[-.27, .03]	[-.03, .28]				
8. AF leisure time	.09	-.12	.01	.61**	.15	-.02	-.01			
	[-.06, .24]	[-.27, .04]	[-.14, .17]	[.51, .70]	[-.00, .30]	[-.17, .14]	[-.17, .15]			
9. D score 1	-.04	-.02	.07	.06	.11	-.03	-.05	.03		
	[-.20, .11]	[-.18, .13]	[-.08, .23]	[-.10, .21]	[-.04, .26]	[-.19, .12]	[-.21, .10]	[-.13, .18]		
10. D score 2	-.11	-.03	.06	.00	.06	.07	-.08	.05	.44**	
	[-.26, .04]	[-.18, .13]	[-.09, .21]	[-.15, .16]	[-.10, .21]	[-.09, .22]	[-.23, .08]	[-.10, .21]	[.30, .55]	
11. D score mean	-.10	-.03	.08	.03	.10	.02	-.08	.05	.82**	.87**
	[-.25, .06]	[-.18, .13]	[-.08, .23]	[-.12, .19]	[-.06, .25]	[-.13, .18]	[-.23, .08]	[-.11, .20]	[.76, .86]	[.83, .91]

Note. GR = gut reactions; env. prot. = environmental protection; env. degr. = environmental degradation; AF = actual feelings; D score 1 = IAT D score calculated via the short test blocks; D score 2 = IAT D score calculated via the long test blocks; D score mean = IAT D score calculated via both the short and the long test blocks. Values in square brackets indicate the 95% confidence interval for each correlation.
* $p < .05$, ** $p < .01$.

Table 4

Correlations Between the Observed Variables in the War IAT Group (Study 3)

Variable	1	2	3	4	5	6	7	8	9	10
1. GR env. prot.										
2. GR env. degr.	-.41**									
	[-.53, -.27]									
3. GR war	-.35**	.46**								
	[-.48, -.20]	[.33, .57]								
4. GR leisure time	.16*	-.24**	-.27**							
	[.01, .31]	[-.38, -.09]	[-.40, -.12]							
5. AF env. prot.	.56**	-.39**	-.35**	.23**						
	[.44, .66]	[-.51, -.25]	[-.48, -.21]	[.07, .37]						
6. AF env. degr.	-.39**	.56**	.51**	-.18*	-.49**					
	[-.52, -.25]	[.44, .65]	[.38, .61]	[-.32, -.02]	[-.60, -.36]					
7. AF war	-.13	.30**	.69**	-.10	-.22**	.39**				
	[-.28, .02]	[.15, .44]	[.60, .76]	[-.26, .05]	[-.36, -.06]	[.25, .51]				
8. AF leisure time	.12	-.14	-.18*	.63**	.18*	-.19*	-.08			
	[-.04, .27]	[-.29, .02]	[-.33, -.03]	[.52, .71]	[.02, .32]	[-.34, -.04]	[-.23, .07]			
9. D score 1	.12	-.09	-.18*	.03	.09	-.17*	-.09	.05		
	[-.04, .27]	[-.24, .07]	[-.32, -.02]	[-.13, .18]	[-.07, .24]	[-.32, -.02]	[-.24, .06]	[-.11, .20]		
10. D score 2	.04	-.00	-.23**	.02	.08	-.13	-.18*	-.06	.37**	
	[-.11, .20]	[-.16, .15]	[-.37, -.08]	[-.14, .17]	[-.07, .24]	[-.28, .03]	[-.32, -.02]	[-.21, .10]	[.22, .49]	
11. D score mean	.10	-.06	-.25**	.03	.10	-.18*	-.16*	-.01	.85**	.80**
	[-.06, .25]	[-.21, .10]	[-.39, -.09]	[-.13, .18]	[-.05, .26]	[-.33, -.03]	[-.31, -.00]	[-.16, .15]	[.80, .89]	[.74, .85]

Note. GR = gut reactions; env. prot. = environmental protection; env. degr. = environmental degradation; AF = actual feelings; D score 1 = IAT D score calculated via the short test blocks; D score 2 = IAT D score calculated via the long test blocks; D score mean = IAT D score calculated via both the short and the long test blocks. Values in square brackets indicate the 95% confidence interval for each correlation.

$* p < .05, ** p < .01.$

Table 5

Correlations Between the Observed Variables in the Leisure IAT Group (Study 3)

Variable	1	2	3	4	5	6	7	8	9	10
1. GR env. prot.										
2. GR env. degr	-.49**									
	[-.60, -.36]									
3. GR war	-.31**	.40**								
	[-.44, -.16]	[.26, .52]								
4. GR leisure time	.12	-.22**	-.11							
	[-.04, .27]	[-.36, -.07]	[-.27, .04]							
5. AF env. prot.	.59**	-.25**	-.18*	.12						
	[.48, .69]	[-.39, -.10]	[-.32, -.02]	[-.04, .27]						
6. AF env. degr.	-.42**	.55**	.48**	-.21**	-.34**					
	[-.54, -.28]	[.43, .65]	[.35, .59]	[-.36, -.06]	[-.47, -.19]					
7. AF war	-.26**	.34**	.75**	-.09	-.16*	.52**				
	[-.40, -.11]	[.20, .47]	[.68, .81]	[-.25, .06]	[-.30, -.00]	[.40, .62]				
8. AF leisure time	.07	-.16*	-.07	.58**	.14	-.13	-.02			
	[-.08, .23]	[-.30, -.00]	[-.23, .08]	[.47, .68]	[-.01, .29]	[-.28, .03]	[-.17, .14]			
9. D score 1	.15	-.06	-.10	-.15	.26**	-.10	-.03	-.06		
	[-.00, .30]	[-.21, .10]	[-.25, .06]	[-.30, .00]	[.11, .40]	[-.26, .05]	[-.19, .12]	[-.22, .09]		
10. D score 2	.09	-.01	-.03	-.25**	.21**	-.05	.04	-.11	.67**	
	[-.07, .24]	[-.17, .14]	[-.19, .12]	[-.39, -.10]	[.06, .36]	[-.20, .11]	[-.12, .19]	[-.26, .05]	[.58, .75]	
11. D score mean	.13	-.04	-.07	-.22**	.26**	-.08	.00	-.09	.92**	.91**
	[-.02, .28]	[-.19, .12]	[-.22, .08]	[-.36, -.07]	[.11, .40]	[-.23, .07]	[-.15, .16]	[-.25, .06]	[.89, .94]	[.88, .94]

Note. GR = gut reactions; env. prot. = environmental protection; env. degr. = environmental degradation; AF = actual feelings; D score 1 = IAT D score calculated via the short test blocks; D score 2 = IAT D score calculated via the long test blocks; D score mean = IAT D score calculated via both the short and the long test blocks. Values in square brackets indicate the 95% confidence interval for each correlation.

* $p < .05$, ** $p < .01$.

Testing measurement invariance in our main multigroup SEM analyses (two indicators for the latent IAT factor)

In Table 6 we present the model tests for the three types of MI that we tested: weak MI, strong MI, and strict MI.

Table 6

Model fit of the different measurement invariance models (Study 3)

Model	S-B χ^2 (df)	p	RMSEA	CFI	SRMR	AIC	BIC	$\Delta\chi^2$	p
Weak MI	14.61 (15)	.48	0.00	1.00	0.04	3729.8	3892.6		
Strong MI	22.99 (19)	.24	0.04	0.99	0.04	3730.6	3876.7	7.38	.12
Strict MI	40.06 (27)	.051	0.06	0.95	0.09	3761.6	3874.3	13.00	.11

Note. S-B χ^2 = Satorra-Bentler scaled χ^2; RMSEA = robust root-mean-square error of approximation; CFI = robust comparative fit index; SRMR = robust standardized root-mean-square residual; AIC = Akaike information criterion; BIC = Bayesian information criterion; MI = measurement invariance.

Multigroup SEM with three indicators for the latent IAT factor

We reanalyzed our main multigroup SEM analyses using three instead of two indicators for the latent IAT factor, hoping to test the configural MI model, which showed convergence problems when modeling the latent IAT factor with two indicators (see Figure 1 for the correlated two-factor model with three indicators for the latent IAT factor that was fitted in all groups). Descriptive statistics of the three D score indicators can be found in Table 7. Since the observed variables were not (multivariately) normally distributed (all *W*s of the Shapiro-Wilk test $\leq$ .92, all *p*s < .001; Mardia's skewness = 632.07, *p* < .001; Mardia's kurtosis = 18.87, *p* < .001), we again used the maximum likelihood mean-variance adjusted (MLMV) estimator. Unfortunately, the configural MI model still did not converge. This was not because of the number of indicators for the latent IAT factor, but because of the number of indicators for the latent direct attitude factor. We present the model tests of the configural MI model with a fixed second indicator of the latent

direct attitude factor to 1, the weak MI model, the strong MI model, and the strict MI model in Table 8. As was true for the analyses with two indicators for the latent IAT factor, strict MI could be assumed. The results with regard to our hypotheses were also very similar and accordingly supported our hypotheses. See Table 9 for the latent means, latent true-score variances and latent correlations of the three IATs. With regard to the latent means the overall test showed that the latent means differed significantly between the three IATs (see Table 10 for the model fit of the different models to test the overall differences in the estimates). Individual Wald tests of each latent mean difference showed that the leisure IAT differed significantly from the degradation IAT, $W(1) = 678.35$, $p < .001$, as well as from the war IAT, $W(1) = 432.52$, $p < .001$, and that the degradation IAT differed significantly from the war IAT, $W(1) = 42.47$, $p < .001$. With regard to the true-score variances the overall test showed that the true-score variances differed significantly between the three IATs (see Table 10). Individual Wald tests of each true-score variance difference, showed that the leisure IAT differed significantly from the degradation IAT, $W(1) = 50.08$, $p < .001$, as well as from the war IAT, $W(1) = 34.34$, $p < .001$, but that the degradation IAT and the war IAT did not differ significantly from each other, $W(1) = 1.37$, $p = .24$. With regard to the latent correlations the overall test showed that the latent correlations differed significantly between the three IATs, since the latent true-score variances differed significantly while the latent covariances did not, which suggests different latent correlations (see Table 10). The use of Baysian evaluation of informative hypotheses showed that the leisure IAT has a higher correlation with the outcome than the other two IATs, while the other two IATs have similar correlations at the same time ($\hat{r}_{\text{leisure IAT}} > \hat{r}_{\text{war IAT}} = \hat{r}_{\text{degradation IAT}}$), since the Bayes factors $BF_{.c}$ and $BF_{.u}$ were both 12.95, which indicates that our hypothesis was 13 times more likely than its complement ($BF_{.c}$) or the unconstrained hypothesis ($BF_{.u}$).

Figure 1

Basic correlated two factor model with three indicators for the latent IAT factor that was fitted in

all groups
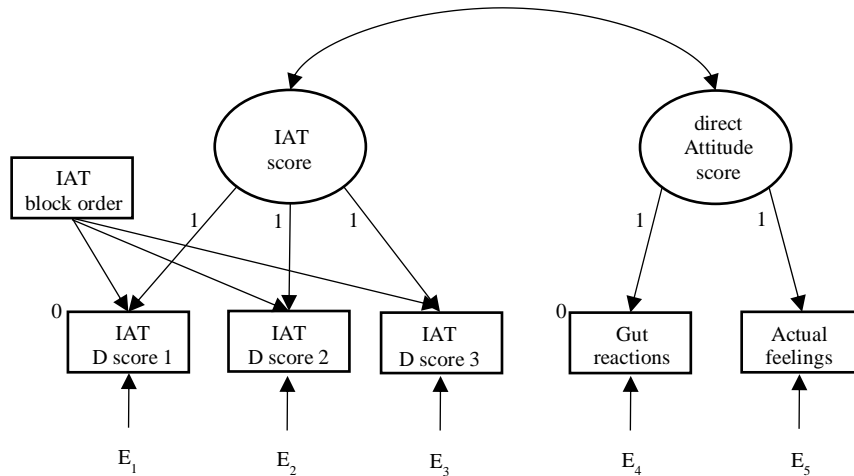
Note. Circles represent latent and rectangles observed variables. IAT = implicit association test;

IAT D score 1 = First IAT D score parcel calculated via trials; IAT D score 2 = Second IAT D

score parcel calculated via trials; IAT D score 3 = Third IAT D score parcel calculated via trials.

Table 7

D Scores and Their Corresponding Mean, Variance, Skew, and Kurtosis per group (Study 3 with

three indicators for the latent IAT factor)

Group	D score	Mean	Variance	Skew	Kurtosis
Degradation IAT	D score 1	0.91	0.08	-0.74	1.23
(n = 160)	D score 2	0.92	0.09	-0.65	0.18
	D score 3	0.95	0.09	-0.84	0.73
War IAT	D score 1	0.73	0.11	-0.60	0.82
(n = 160)	D score 2	0.74	0.12	-0.68	0.87
	D score 3	0.74	0.13	-0.83	1.20
Leisure IAT	D score 1	-0.14	0.30	0.16	-0.93
(n = 160)	D score 2	-0.15	0.33	0.13	-0.71
	D score 3	-0.16	0.31	0.18	-0.91

Note. IAT = implicit association test; D score 1 = First IAT D score parcel calculated via trials; D score 2 = Second IAT D score parcel calculated via trials; D score 3 = Third IAT D score parcel calculated via trials.

Table 8

Model fit of the different measurement invariance models (Study 3 with three indicators for the latent IAT factor)

Model	S-B χ^2 (df)	p	RMSEA	CFI	SRMR	AIC	BIC	$\Delta\chi^2$	p
ConFix MI	21.34 (21)	.44	0.01	1.00	0.04	3775.8	4001.2		
Weak MI	31.42 (27)	.25	0.03	0.99	0.05	3775.3	3975.6	8.83	.18
Strong MI	34.91 (33)	.38	0.02	1.00	0.05	3766.0	3941.3	3.71	.72
Strict MI	52.75 (43)	.15	0.04	0.98	0.06	3788.1	3921.6	14.07	.17

Note. S-B χ^2 = Satorra-Bentler scaled χ^2; RMSEA = robust root-mean-square error of approximation; CFI = robust comparative fit index; SRMR = robust standardized root-mean-square residual; AIC = Akaike information criterion; BIC = Bayesian information criterion; ConFix = configural with a fixed second indicator of the latent direct attitude factor to 1; MI = measurement invariance.

Table 9

Latent means, latent true-score variances, latent correlations, R^2, and reliabilities of the latent IAT factor for the three groups in the strict invariance model (Study 3 with three indicators for the latent IAT factor)

Group	L mean (*SE*)	L variance (*SE*)	L correlation (*CI*)	R^2	Reliability
Degradation IAT	0.92 (.02)	0.04 (.01)	.01 (-.20, .25)	0.00	0.69
War IAT	0.73 (.02)	0.06 (.01)	.20 (-.02, .40)	0.04	0.74
Leisure IAT	-0.15 (.04)	0.19 (.02)	.36 (.18, .52)	0.13	0.84

Note. L = latent; CI = bootstrap-bias-corrected confidence intervals; IAT = implicit association test.

Table 10

Model fit of the different models to test the overall manipulation hypotheses (Study 3 with three indicators for the latent IAT factor)

Model	S-B χ^2 (df)	p	RMSEA	CFI	SRMR	AIC	BIC	$\Delta\chi^2$	p
Strict MI	52.75 (43)	.15	0.04	0.98	0.06	3788.1	3921.6		
Means	215.58 (47)	<.001	0.15	0.60	0.60	4095.0	4211.9	221.41	<.001
Variances	362.13 (51)	<.001	0.20	0.27	1.04	4363.5	4463.6	160.25	<.001
Covariances	368.04 (53)	<.001	0.19	0.26	1.04	4359.9	4451.8	1.49	.47

Note. S-B χ^2 = Satorra-Bentler scaled χ^2; RMSEA = robust root-mean-square error of approximation; CFI = robust comparative fit index; SRMR = robust standardized root-mean-square residual; AIC = Akaike information criterion; BIC = Bayesian information criterion; MI = measurement invariance; Means = strict measurement invariance model plus equal group means; Variances = strict measurement invariance model plus equal group means and variances; Covariances = strict measurement invariance model plus equal group means, variances, and covariances.

**Multigroup SEM with recoding as a third variable (two indicators for the latent IAT factor)**

To examine the role of recoding, we ran the same multigroup SEM analyses that we report in the main text to test our hypotheses, except this time we included an indicator of recoding as a third variable (see the discussion in the main text on recoding for a better understanding of the recoding indicator; p. 70). Figure 8 displays the correlated two-factor model including the variable recoding that was fitted in all groups. The observed variables were not (multivariately) normally distributed (all Ws of the Shapiro-Wilk test $\leq .92$, all ps $< .001$; Mardia's skewness = 580.93, $p < .001$; Mardia's kurtosis = 24.00, $p < .001$) and we used the maximum likelihood mean-variance adjusted (MLMV) estimator accordingly. We tested weak MI, strong MI, and strict MI, while we could not test configural MI for the same reasons already

outlined in the main text (see Table 11 for the model tests). As was true for our main analyses without recoding, strict MI could be assumed. However, the results related to our hypotheses changed. While the latent means and the latent true-score variances still differed significantly between the groups, the latent correlations did not. See Table 12 for the latent means, latent true-score variances and latent correlations of the three IATs. With regard to the latent means the overall test showed that the latent means differed significantly between the three IATs (see Table 13 for the model fit of the different models to test the overall differences in the estimates). Individual Wald tests of each latent mean difference showed that the leisure IAT differed significantly from the degradation IAT, $W(1) = 390.43$, $p < .001$, as well as from the war IAT, $W(1) = 281.93$, $p < .001$, and that the degradation IAT differed significantly from the war IAT, $W(1) = 12.14$, $p < .001$. With regard to the true-score variances the overall test showed that the true-score variances differed significantly between the three IATs (see Table 13). Individual Wald tests of each true-score variance difference, showed that the leisure IAT differed significantly from the degradation IAT, $W(1) = 35.00$, $p < .001$, as well as from the war IAT, $W(1) = 17.11$, $p < .001$, but that the degradation IAT and the war IAT did not differ significantly from each other, $W(1) = 2.28$, $p = .13$ (note, however, that the true-score variance of the leisure IAT was nevertheless strongly reduced in comparison to our main results without recoding). With regard to the latent correlations the overall test still suggests that the latent correlations differed significantly between the three IATs, since the latent true-score variances differed significantly while the latent covariances did not (see Table 13), however, none of the correlations are individually significantly different from zero. This latter result is also strongly supported by the use of Baysian evaluation of informative hypotheses, which suggests that the leisure IAT, the degradation IAT, and the war IAT all have a correlation with the outcome of zero ($\hat{r}_{\text{leisure IAT}} = \hat{r}_{\text{war}}$

$_{IAT} = \hat{r}_{degradation\ IAT} = 0$), since the Bayes factors $BF_{.c}$ and $BF_{.u}$ for this hypothesis were both

250.61.

Figure 2

Basic correlated two factor model including recoding as a third variable that was fitted in all
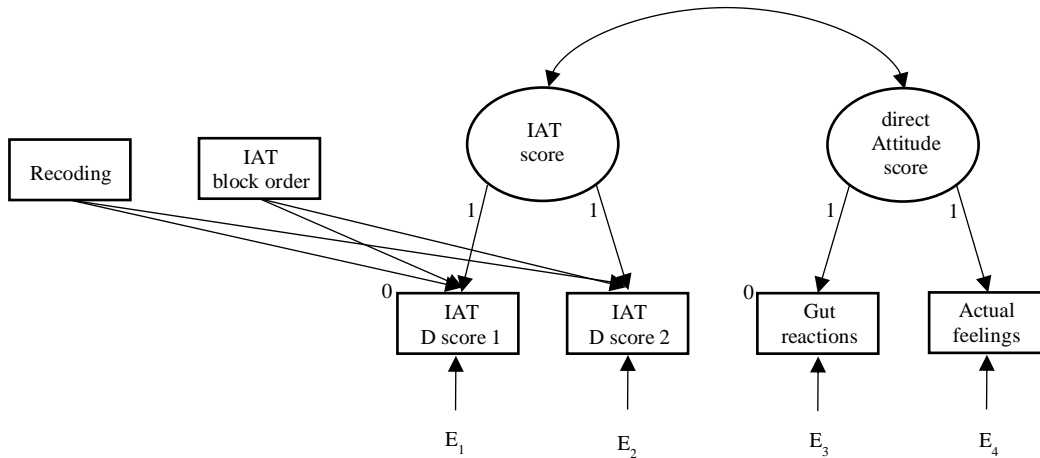
groups

Note. Circles represent latent and rectangles observed variables. IAT = implicit association test;

IAT D score 1 = IAT D score calculated via the short test blocks; IAT D score 2 = IAT D score

calculated via the long test blocks.

Table 11

Model fit of the different measurement invariance models (Study 3 including recoding as a third

variable)

Model	S-B χ^2 (df)	p	RMSEA	CFI	SRMR	AIC	BIC	$\Delta\chi^2$	p
Weak MI	35.05 (21)	.03	0.07	0.96	0.05	3684.5	3872.4		
Strong MI	39.04 (25)	.04	0.06	0.96	0.05	3679.8	3851.0	3.61	.46
Strict MI	48.66 (33)	.04	0.05	0.95	0.08	3705.4	3843.1	11.90	.16

Note. S-B χ^2 = Satorra-Bentler scaled χ^2; RMSEA = robust root-mean-square error of

approximation; CFI = robust comparative fit index; SRMR = robust standardized root-mean-

square residual; AIC = Akaike information criterion; BIC = Bayesian information criterion; MI = measurement invariance.

Table 12

Latent means, latent true-score variances, latent correlations, R^2, and reliabilities of the latent IAT factor for the three groups in the strict invariance model (Study 3 including recoding as a third variable)

Group	L mean (*SE*)	L variance (*SE*)	L correlation (*CI*)	R^2	Reliability
Degradation IAT	0.95 (.04)	0.02 (.01)	-.03 (-.33, .32)	0.00	0.39
War IAT	0.77 (.04)	0.04 (.01)	.20 (-.06, .47)	0.04	0.52
Leisure IAT	-0.04 (.04)	0.12 (.02)	.16 (-.08, .39)	0.03	0.57

Note. L = latent; CI = bootstrap-bias-corrected confidence intervals; IAT = implicit association test.

Table 13

Model fit of the different models to test the overall manipulation hypotheses (Study 3 including recoding as a third variable)

Model	S-B χ^2 (df)	*p*	RMSEA	CFI	SRMR	AIC	BIC	$\Delta\chi^2$	*p*
Strict MI	48.66 (33)	.04	0.05	0.95	0.08	3705.4	3843.1		
Means	179.71 (37)	<.001	0.16	0.58	0.34	3972.4	4093.5	199.32	<.001
Variances	235.13 (41)	<.001	0.17	0.43	0.59	4070.7	4175.1	61.31	<.001
Covariances	238.42 (43)	<.001	0.17	0.42	0.59	4067.6	4163.6	1.36	.51

Note. S-B χ^2 = Satorra-Bentler scaled χ^2; RMSEA = robust root-mean-square error of approximation; CFI = robust comparative fit index; SRMR = robust standardized root-mean-square residual; AIC = Akaike information criterion; BIC = Bayesian information criterion; MI = measurement invariance; Means = strict measurement invariance model plus equal group means; Variances = strict measurement invariance model plus equal group means and variances;

Covariances = strict measurement invariance model plus equal group means, variances, and covariances.