

Table 3

*Original and replication results*

Effect	Original Study			Unweighted			Weighted			Null Hypothesis Significance Tests by Sample			Null Hypothesis Significance Tests of Aggregate		
	ES stat	ES	95% CI Lower, Upper	Median Replication ES	Replication ES	95% CI Lower, Upper	Replication ES	95% CI Lower, Upper	Proportion <0 (p<.05)	Proportion >0 (p<.05)	Proportion <i>ns</i>	Key statistics	df	N	p
Stroop Task	<i>d</i>	various	n/a	0.89	0.91	0.84, 0.98	0.88	0.84, 0.92	0.00	1.00	0.00	t = 49.795	3336	3337	< .001
Metaphoric Restructuring	<i>d</i>	0.63	0.07, 1.20	0.19	0.29	0.17, 0.42	0.26	0.15, 0.37	0.00	0.25	0.75	$\chi^2 = 21.90$	1	1335	< .001
Availability Heuristic	<i>d</i>	0.82	0.47, 1.17	0.07	0.09	0.02, 0.16	0.09	0.02, 0.16	0.00	0.14	0.86	PSdep = .522	N/A	3088	0.015
Persistence and Conscientiousness	<i>d</i>	1.04	0.64, 1.43	0.000	0.03	-0.04, 0.10	0.05	-0.02, 0.12	0.00	0.00	1.00	r = .027	3191	3193	0.134
Power and Perspective	<i>d</i>	0.77	0.12, 1.41	0.03	0.03	-0.04, 0.11	0.03	-0.04, 0.10	0.00	0.05	0.95	t = .89	2967	2969	0.37
Weight Embodiment	<i>d</i>	0.59	0.01, 1.16	0.05	0.03	-0.05, 0.11	0.03	-0.06, 0.11	0.00	0.00	1.00	t = .61	2283	2285	0.543
Warmth Perceptions	<i>d</i>	0.86	0.40, 1.33	0.06	0.01	-0.08, 0.06	0.01	-0.06, 0.08	0.05	0.00	0.95	t = .22	3117	3119	0.827
Elaboration Likelihood	$\eta p^2$	0.17	0.06, 0.29	0.002	0.003	0.000, 0.009	0.00005	0.000, 0.002	0.00	0.00	1.00	F = .129	1, 2361	2365	0.72
Self-Esteem and Subjective Distance	$\eta p^2$	n/a	n/a	0.004	0.006	0.002, 0.013	0.001	0.000, 0.004	0.05	0.10	0.86	F = 1.98	1, 3131	3136	0.16
Credentials and Prejudice	$\eta p^2$	0.04	0, 0.09	0.005	0.005	0.001, 0.011	0.000	0.000, 0.000	0.00	0.10	0.90	F = .0004	1, 3130	3134	0.985
<i>Added Effects</i>															
*Elaboration	$\eta p^2$	0.59	0.47, 0.67	0.031	0.034	0.022, 0.049	0.033	0.020, 0.048	0.00	0.45	0.55	F =	1, 2361	2365	<

Likeihood - Main Effect												79.925				.001
*Self-Esteem and Subjective Distance - Main Effect	$\eta p^2$	n/a	n/a	0.011	0.016	0.009, 0.026	0.011	0.005, 0.019	0.00	0.24	0.76	F = 7.97	1, 3131	3136		<
*Credentials and Prejudice - Main Effect	$\eta p^2$	0.043	0.002, 0.103	0.006	0.010	0.005, 0.019	0.005	0.001, 0.012	0.00	0.14	0.48	F = 17.01	1, 3130	3134		<.001

*Note.* Weighted statistics are computed on the whole aggregated dataset; Unweighted statistics are computed on the disaggregated dataset (N = 20 or 21). 95% CI's for original effect sizes used cell sample sizes when available and assumed equal distribution across conditions when not available. Confidence intervals around the unweighted mean are based on the central normal distribution. Confidence intervals around the weighted effect size are based on non-central distributions. \* For three experiments, reliable main effects were added after observing the aggregate outcomes to have more effects to then test for variation across the academic semester. Credentials and Prejudice interaction effect size was estimated as 1.28 e-6, the weighted upper bound of the 95% CI was too small to compute with the statistical software.  $\eta p^2$ 's were not available for the original self-esteem and subjective distance effects. The Cohen's d estimates are 0.21 (95% CI 0.001, 0.418) and 0.39 (95% CI 0.18, 0.60) respectively.

Table 4

*Heterogeneity in effect sizes by data collection site*

Effect	Heterogeneity tests			
	Q	df	p-value	I <sup>2</sup>
Stroop Task	15.1883	20	0.7655	4.05%
Metaphoric Restructuring	21.9213	19	0.2882	18.23%
Availability Heuristic	19.9805	20	0.4591	1.21%
Persistence and Conscientiousness	22.4037	20	0.319	1.40%
Power and Perspective	19.7975	20	0.4707	0.01%
Weight Embodiment	12.2518	19	0.8746	<0.005%
Warmth Perceptions	16.9429	20	0.6567	<0.005%
Elaboration Likelihood	3.037	19	1	<0.005%
Self-Esteem and Subjective Distance	16.9471	20	0.6564	<0.005%
Credentials and Prejudice	14.8654	20	0.7841	<0.005%
*Elaboration Likelihood - Main Effect	12.6038	19	0.8582	<0.005%
*S-E and Subjective Distance - Main Effect	15.2059	20	0.7645	<0.005%
*Credentials and Prejudice - Main Effect	18.9273	20	0.5266	3.04%

*Note.* Effects were ordered from largest to smallest observed effect size (see Table 3). Heterogeneity tests conducted with R-package metafor. REML was used for estimation for all tests.

Table 3

Order effects by task order

Effect	F (linear)	df (interaction)	df (residuals)	p-value	$\eta_p^2$	95% CI	F (quadratic)	df (interaction)	df (residuals)	p-value	$\eta_p^2$	95% CI
Stroop Task	1.60	1	3278	0.21	0.00049	0, 0.003	2.38	1	3278	0.12	0.00073	0, 0.004
Persistence and Conscientiousness	0.02	1	3189	0.88	0.00001	0, 0.001	0.06	1	3189	0.81	0.00002	0, 0.001
Power and Perspective	1.17	1	2965	0.28	0.00039	0, 0.003	0.54	1	2965	0.46	0.00018	0, 0.002
Weight Embodiment	0.02	1	2066	0.88	0.00001	0, 0.002	0.00	1	2066	0.95	0.00000	0, 0.001
Warmth Perceptions	0.05	1	3115	0.82	0.00002	0, 0.001	0.12	1	3115	0.72	0.00004	0, 0.002
Elaboration Likelihood Self-Esteem and Subjective Distance	0.35	1	2357	0.56	0.00015	0, 0.003	0.04	1	2357	0.84	0.00002	0, 0.002
	0.42	1	3127	0.52	0.00013	0, 0.002	0.60	1	3127	0.44	0.00019	0, 0.002
Credentials and Prejudice	0.02	1	3126	0.90	0.00000	0, 0.001	0.08	1	3126	0.78	0.00003	0, 0.001
*Elaboration Likelihood - Main Effect	2.04	1	2357	0.15	0.00087	0, 0.005	2.08	1	2357	0.15	0.00088	0, 0.005
*S-E and Subjective Distance - Main Effect	0.09	1	3127	0.76	0.00003	0, 0.001	0.43	1	3127	0.51	0.00014	0, 0.002
*Credentials and Prejudice - Main Effect	0.33	1	3126	0.57	0.00010	0, 0.002	0.10	1	3126	0.76	0.00003	0, 0.001
Averages	0.56			0.59	0.00020		0.59			0.59	0.00020	
Binomial Outcomes	Likelihood Chi-Square (linear)	p-value	d	95% CI	Likelihood Chi-Square (quadratic)	p-value	d	95% CI				
Metaphoric Restructuring	0.03	0.87	0.01	-0.10, 0.12	0.17098	0.68	0.02	-0.09, 0.14				
Availability Heuristic	3.11	0.08	0.06	-0.01, 0.13	2.16960	0.14	0.05	-0.02, 0.12				
Averages	1.57	0.4754	0.035		1.17030	0.41	0.04					

Table 6

*Moderation of effect sizes by time of semester*

Effect	Variation in Outcome by Site (R2)	Overall model fit Time of semester	p-value	Time of semester	p-value	partial eta-sq	95% CI
Stroop Task	0.6%	$\chi^2(1, N = 2660) = 3.31$	0.069	$F(1, 2658) = 5.01$ $\chi^2(1, N = 1332) = .010$	0.025	0.002	0, 0.007
Metaphoric Restructuring	0.01%	$\chi^2(1, N = 1332) = 4.48$	0.034		0.92		
Persistence and Conscientiousness	5.0%	$\chi^2(2, N = 2624) = 4.63$	0.099				
Availability Heuristic	0.01%	$\chi^2(4, N = 2497) = 1.45$	0.228				
Power and Perspective	0.9%	$\chi^2(2, N = 2385) = .70$	0.699				
Weight Embodiment	1.7%	$\chi^2(2, N = 2279) = 3.97$	0.138				
Warmth Perceptions	22.0%	$\chi^2(2, N = 2544) = 6.04$	0.049	$F(1, 1842) = 3.83$	0.051	0.002	0, 0.008
Elaboration Likelihood	1.1%	$\chi^2(4, N = 2365) = 2.02$	0.732				
Self-Esteem and Subjective Distance	0.9%	$\chi^2(4, N = 2562) = .54$	0.969				
Credentials and Prejudice	0.4%	$\chi^2(4, N = 2571) = 4.90$	0.298				
*Elaboration Likelihood - Main Effect	1.1%	$\chi^2(2, N = 2429) = .22$	0.896				
*S-E and Subjective Distance - Main Effect	0.9%	$\chi^2(2, N = 2562) = .32$	0.851				
*Credentials and Prejudice - Main Effect	0.4%	$\chi^2(2, N = 2642) = 2.15$	0.341				
Data Quality Indicators							
Attention Check	4%	$\chi^2(1, N = 2621) = 6.75$	0.009	$r(2621) = -.08$	<.001		-.12, -.04
Reported Effort	2.5%	$\chi^2(1, N = 2628) = 17.46$	<.001	$r(2626) = -.11$	<.001		-.14, -.07
Reported Attention	1.6%	$\chi^2(1, N = 2630) = 11.60$	<.001	$r(2628) = -.08$	<.001		-.12, -.04

Demographics						
Age	2.6%	$\chi^2(1, N = 2592) = 0.05$	0.821			
Sex	3.7%	$\chi^2(1, N = 2598) = 17.57$	<.001	$r(2598) = 0.12$	<.001	.08, .16
Race/Ethnicity	1.7%	$\chi^2(1, N = 2607) = 2.38$	0.123			
Year in College	14.2%	$\chi^2(1, N = 2570) = 0.89$	0.346			
Individual Differences						
Conscientiousness	4.2%	$\chi^2(1, N = 2628) = 32.11$	<.001	$r(2626) = -.14$	<.001	-.18, -.10
Agreeableness	<0.01%	$\chi^2(1, N = 2629) = 0.005$	0.945			
Extraversion	2.2%	$\chi^2(1, N = 2625) = 2.40$	0.121			
Neuroticism	1.1%	$\chi^2(1, N = 2630) = 1.31$	0.252			
Openness to Experience	1.8%	$\chi^2(1, N = 2631) = 0.01$	0.923			
Intrinsic Motivation	1.2%	$\chi^2(1, N = 2608) = 0.14$	0.71			
Stress	1.9%	$\chi^2(1, N = 2623) = 10.08$	0.001	$r(2621) = .08$	<.001	.04, .12
Mood	1.2%	$\chi^2(1, N = 2636) = 8.03$	0.005	$r(2634) = -.07$	0.001	-.10, -.03
Self-Esteem	0.4%	$\chi^2(1, N = 2625) = 0.56$	0.456			
Need for Cognition	1.1%	$\chi^2(1, N = 2601) < 0.00005$	0.998			

*Note.* Variation by site indicates the amount of variation in the dependent variable attributable to location of data collection. Follow-up tests of time of semester predicting variation in the effect conducted for only those effects in which the overall model improved ( $p < .07$ ) by adding time of semester as a factor. Two of the outcomes, sex and attention check, had binary outcomes. Changes over time in those variables are quantified by odds ratio of a given outcome on the last day of the semester compared to the first day of the semester (odds ratio estimates taken from the mixed model).