Rank-Order Stability of Domain-Specific Self-Esteem:

A Meta-Analysis

Laura C. Dapp and Ulrich Orth

SUPPLEMENTAL MATERIALS

Supplemental Table S1

Mean Reliability of Rank-Order Stability of Domain-Specific Self-Esteem

	Harter		Marsh	
Domain	$k_{samples}$	Mean reliability	$k_{samples}$	Mean reliability
Academic	23	.78	12	.83
Appearance	17	.82	5	.87
Athletic	14	.80	3	.80
Morality	13	.77	2	.79
Romantic	6	.73	4	.66
Social	39	.76	7	.77
Mathematics	—	—	23	.89
Verbal	—	—	20	.86

Note. Mean reliability was computed by averaging reliability coefficients reported in primary studies separately for domains and separately for Harter's and Marsh's measures. Dash indicates that the domain is not covered by measures by Harter. $k_{samples}$ = number of samples.

Supplemental Table S2

Estimates of Rank-Order Stability of Domain-Specific Self-Esteem When Excluding Outliers (Sensitivity Analyses)

Domain	$k_{samples}$	k_{ES}	N	Weighted mean effect size	95% CI	Q	Variances σ_1^2	Variances σ_2^2
Academic	53	142	50,758	.693*	[.646, .734]	6376.59*	.074	.031
Appearance	–	–	–	–	–	–	–	–
Athletic	–	–	–	–	–	–	–	–
Morality	25	79	12,474	.670*	[.604, .727]	1982.79*	.058	.034
Romantic	–	–	–	–	–	–	–	–
Social	66	205	57,514	.682*	[.643, .718]	8945.62*	.050	.061
Mathematics	27	74	53,282	.680*	[.632, .722]	5227.34*	.037	.023
Verbal	21	64	48,268	.661*	[.608, .708]	4605.48*	.012	.077

Note. The analyses were based on multilevel random-effects models. $k_{samples}$ = number of samples; k_{ES} = number of effect sizes; N = number of participants (overall sample size); weighted mean effect size = disattenuated test-retest correlation coefficient, indicating the rank-order stability of domain-specific self-esteem; CI = confidence interval; Q = statistic used to test residual heterogeneity; σ_1^2 = variance component corresponding to the level of the grouping variable (i.e., between samples); σ_2^2 = variance component corresponding to the level nested within the grouping variable (i.e., within samples). Dash indicates that there were no outliers for the domain.

* $p < .05$.

Supplemental Table S3

Sample-Level Tests of Publication Bias in Rank-Order Stability of Domain-Specific Self-Esteem

	Egger's regression test		
Domain	*k*	*z*	*p*
Academic	53	0.613	.540
Appearance	39	-1.660	.097
Athletic	30	-0.638	.524
Morality	25	-0.007	.995
Romantic	19	0.974	.330
Social	67	1.269	.204
Mathematics	28	0.917	.359
Verbal	21	0.679	.497

Note. k = number of effect sizes.

Supplemental Table S4

Intercorrelations Among Moderators

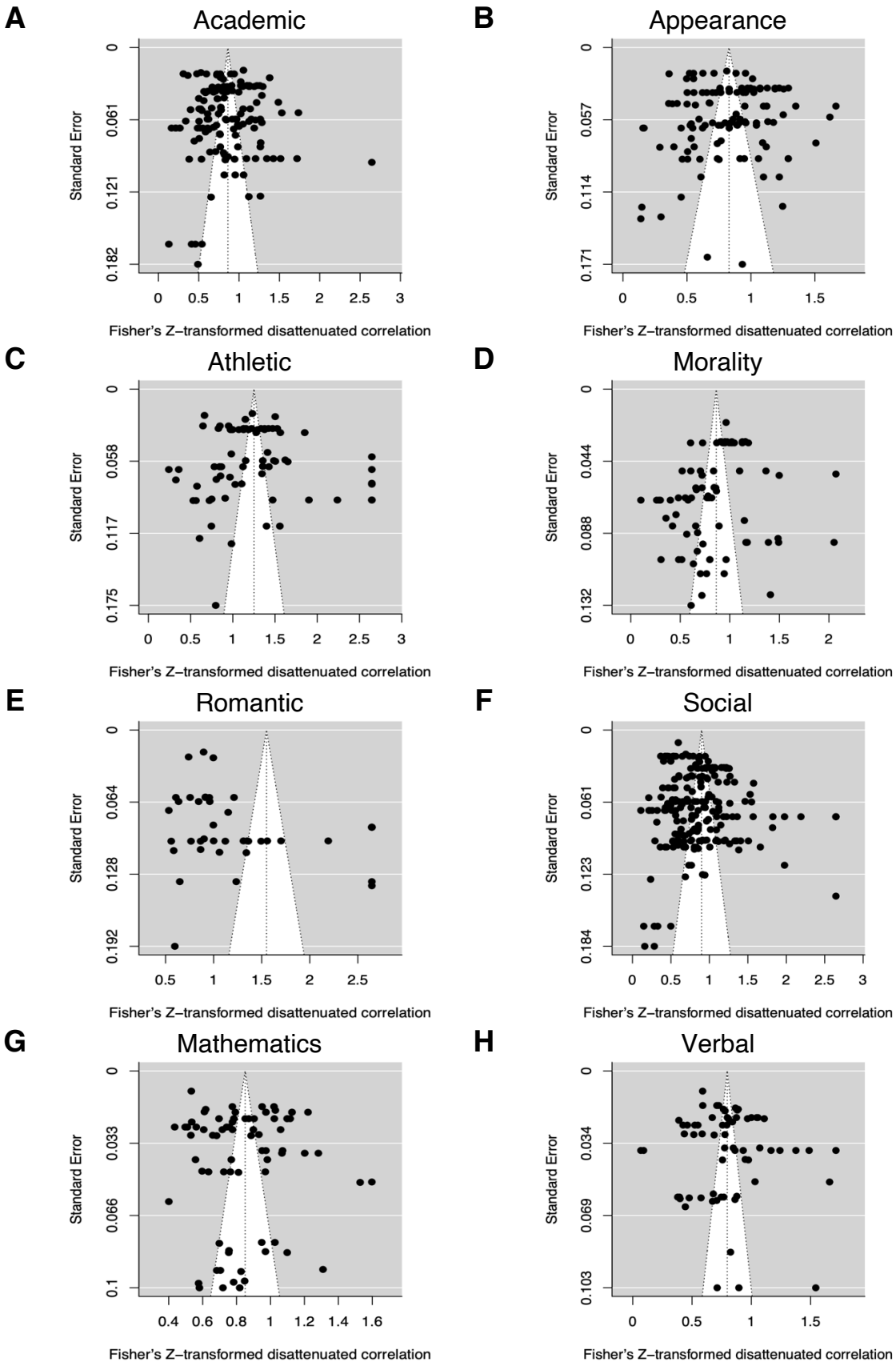
Variable	1	2	3	4
1. Age	—			
2. Time lag	**-.167**	—		
3. Female (%)	**.186**	-.065	—	
4. Measure[a]	**.094**	**-.107**	-.107	—

Note. All correlations were computed on the level of effect sizes, except for the correlation between female and measure, which was computed on the sample level (i.e., these variables did not vary across effect sizes from the same sample). Age = age at beginning of the interval on which effect size was based. Time lag = interval between assessments on which effect size was based. Values in bold are significant at $p < .05$.

[a] 0 = Harter; 1 = Marsh.

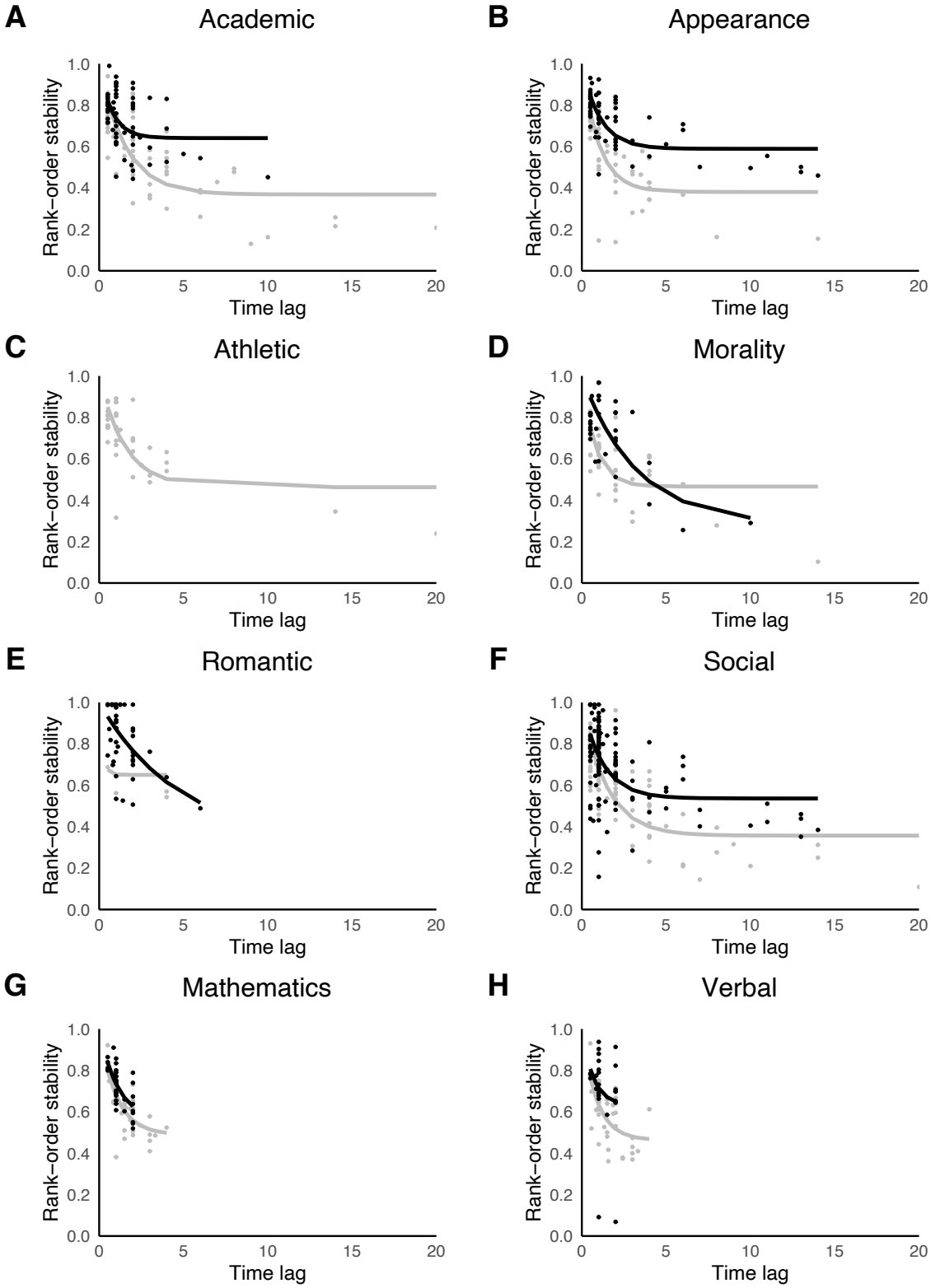
Supplemental Figure S1

Funnel Graphs Displaying the Relation Between Standard Error and Effect Size of Rank-Order Stability



Supplemental Figure S2

Scatterplots Displaying the Relation Between Rank-Order Stability of Domain-Specific Self-Esteem and Time Lag for Younger and Older Samples, Including the Estimated Function of Exponential Decay



Note. For each domain, the function is shown for the range of time lags for which data were available. Gray = younger samples; black = older samples. For the older samples in the athletic domain, the model did not converge.