SUPPLEMENTAL MATERIALS

Title: Implications of Item Keying and Item Valence for the Investigation of Construct Dimensionality

By Chester Chun Seng Kam and John P. Meyer

Additional Multitrait-Multimethod Confirmatory Factor Analysis

Before concluding that item valence is the underlying cause for our results and that $M_{2 \text{ valence}}$ is our final model, we also considered another alternative explanation for our findings. This alternative explanation is that participants have an identical response style among all items, regardless of whether the items are positively-valenced or negatively-valenced. We tested this explanation with the MTMM technique. To examine this alternative explanation, we constructed a model (M_{common}) in which all items are loaded on a common method factor regardless of whether these items are valenced positively or negatively. M_{baseline} is nested within M_{common} and M_{common} is nested within $M_{2 \text{ valence}}$. Therefore, chi-square difference tests (Satorra & Bentler, 2001; 2010) will be used for the comparisons (see Supplementary Table S3). M_{common} fit better than M_{baseline} but worse than $M_{2 \text{ valence}}$ in both samples ($ps < .001$). Not only do participants' response styles exist for both positively- and negatively-valenced items, but our result also suggests that the response patterns *differ* for the two types of items.

Readers may notice that the average loading for SDO items on the negative valence factor is much higher than that for the other items in Table S3. Similarly, the average loading for the agreeableness items on the positive factor is much higher than that for other items. Interestingly, the larger method effect for the constructs agreeableness and SDO here mirrors our earlier results that these two constructs are more likely to show a two-factor solution than other constructs. Our MTMM results were thus able to detect that SDO and agreeableness have a stronger item valence effect than other constructs. To ensure that our results are not simply dominated by agreeableness or SDO items, we conducted another MTMM analysis excluding agreeableness and SDO items in the model. The analysis continued to suggest the

existence of two valence factors (see Supplementary Table S5). Therefore, the existence of item valence factors is not simply due to a few specific constructs in our investigation.

Additional Nomological Network Analyses and Correlational Analyses

Our analysis the main manuscript showed that positive and negative self-esteem has distinct nomological networks. Afterwards, we further added in full-scale scores of self-esteem from the Sample 1 and Sample 2 data and compared how half-scale positive self-esteem, full-scale self-esteem, and negative self-esteem scores correlate with other constructs. By comparing the pattern of correlations between full-scaled scores and half-scaled scores, we can examine how correlations can be inflated or deflated by construct valence.¹ If scores sharing the same valence direction have an inflated correlation and scores with opposing valence directions have a deflated correlation, we would expect that in a majority of cases the positively-valenced variables would correlate most highly with the positive self-esteem score, followed by the full-scaled score and finally by the negative self-esteem score. The results are shown in the middle panel of Supplementary Table S6.² Indeed, for Sample 1 (see the left side of the analyses), when other variables in the investigations were positively-valenced, the expected pattern of correlations was discovered all seven times; see the middle left panel). Conversely, we predicted that negatively-valenced variables would correlate most highly with the negative self-esteem score, followed by the full-scaled score and finally by the positive self-esteem score. The result is shown in the bottom panel of Supplementary Table S6. Again, the expected correlational pattern was found five out of seven times (see the bottom left panel). A similar pattern of results was found for the self-esteem data in Sample 2 (see the right panel).

¹ We assume a full-scaled score as the standard of comparison because it gives a more balanced representation of a bipolar construct than any half-scaled scores. For example, a half-scaled score of introversion may be unable to fully represent its opposite pole of extraversion.

² Belief in Zero-Sum Resources (BZSR) was not included in the analysis because the unreliability of its single positively-valenced item could not be corrected by the latent variable modeling procedure.

Our nomological network analyses in the manuscript clearly demonstrated that item valence can affect a researcher's substantive conclusion regarding the dimensionality of self-esteem. We further examined the generalizability of this valence effect with a common personality dimension – extraversion. The basic pattern of the results for extraversion (see Supplementary Tables S7 and S8) were similar to the pattern found for self-esteem, with the exception that extraversion seemed to have more advantage over introversion in correlating with other full-scaled variables in Sample 2 (see the top right panel of Supplementary Table S8). Nevertheless, the noteworthy results for Sample 2 here were that the correlations favored extraversion even more when the other variables were positively valenced (see the middle right panel of Table 8). Conversely, the correlations favored extraversion less when other variables were negatively valenced (see the bottom right panel of Supplementary Table S8).

The slight difference between the Sample 1 results and the Sample 2 results for self-esteem or extraversion should not be alarming because it can simply be the outcome of sampling errors (see Schmidt & Hunter, 1996). Random sampling errors can cause the observed statistical results to fluctuate across samples (Schmidt & Hunter, 1996). The overall conclusion drawn from the extraversion measure is the same as that drawn from the self-esteem measure.

Additional Analyses on Variance Explained by Item Valence

In addition to the CTCM analysis in the manuscript, a specific MTMM technique, correlated-trait (correlated-trait method minus one) [CT(M-1)] model, allowed us to directly examine the unique variance captured by a specific method factor such as negative valence (Eid, Lischetzke, Nussbeck, & Trierweiler, 2003; Geiser et al., 2012). Based on this method, only one method factor was modeled in the current study. Variance could be decomposed into (1) variance of negatively-valenced items that was shared with positively-valenced items, (2) unique variance accounted for by negatively-valenced items, and (3) indicator-specific variance (i.e., residual variance). The two methods (positively-valenced items and negatively-valenced items) are thus theoretically uncorrelated in this model to capture the unique variance captured by negatively-valenced items. Results showed that among negatively-valenced items in Sample

1 (Sample 2), 50.56% (46.04%) of the variance was shared with positively-valenced variance, 8.79% (17.15%) was uniquely captured by negatively-valenced items, and 30.82% (36.81%) was residual variance. The ratio of unique versus shared variance was estimated to be 1:5.75 in Sample 1 and 1:2.68 in Sample 2. Therefore, negatively-valenced items capture substantive amount of variance that is not shared with positively-valenced items. The use of only one specific valence type (e.g., positively-valenced items only) can clearly lead to bias in construct measurement.

References

- Eid, M, Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple indicator CT-C(M-1) model. *Psychological Methods*, 8, 38-60.
- Geiser, C., Eid, M., West, S. G., Lischetzke, T., & Nussbeck, F. W. (2012). A comparison of method effects in two confirmatory factor models for structurally different methods. *Structural Equation Modeling*, 19, 409-436.
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507-514.
- Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, 75, 243-248.
- Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, 1, 199-223.

Supplementary Table S1
Scale Reliabilities

	Full scale	Positively-valenced scale	Negatively-valenced scale
Sample 1			
Conscientiousness	.80	.67	.73
Extraversion	.87	.81	.77
Agreeableness	.76	.67	.70
Openness	.76	.52	.64
Neuroticism	.82	.66	.73
Self-Esteem	.88	.80	.82
SDO	.91	.87	.86
BDW	.81	.71	.72
BZSR	.91	-	.93
Average reliability	.84	.71	.77
Sample 2			
Conscientiousness	.77	.65	.70
Extraversion	.87	.81	.78
Agreeableness	.76	.64	.72
Openness	.75	.53	.59
Neuroticism	.84	.67	.78
Self-Esteem	.90	.84	.84
SDO	.94	.87	.87
Average reliability	.83	.72	.75
Social Desirability	.77	-	-
Impression Management	.76	-	-
Self-Deception	.66	-	-

Note. No reliability (internal consistency) was estimated for the single positively-valenced item of BZSR.

Supplementary Table S2
Basic Confirmatory Factor Analysis Models in Sample 1

	Absolute Fit Indices							Predictive Fit Indices		Factor Correlation
	χ^2	<i>df</i>	<i>p</i>	TLI	CFI	RMSEA	SRMR	AIC	SABIC	
Self-Esteem										
One-Factor	365.82	35	< .001	.86	.89	.09	.05	19530.01	19584.51	-.87
Two-Factor	302.27	34	< .001	.88	.91	.09	.05	19424.35	19480.67	
Extraversion										
One-Factor	293.47	35	< .001	.89	.91	.08	.04	28866.53	28921.12	-.84
Two-Factor	187.34	34	< .001	.93	.95	.06	.04	28732.04	28788.45	
Conscientiousness										
One-Factor	436.53	35	< .001	.75	.80	.10	.06	28771.58	28826.14	-.72
Two-Factor	308.69	34	< .001	.82	.87	.09	.06	28621.19	28677.57	
Agreeableness										
One-Factor	316.47	35	< .001	.79	.84	.09	.06	28174.93	28229.51	-.67
Two-Factor	136.05	34	< .001	.92	.94	.05	.04	27970.82	28027.23	
Openness										
One-Factor	915.60	35	< .001	.46	.59	.15	.11	29504.25	29560.65	-1.00
Two-Factor	872.92	34	< .001	.48	.61	.15	.11	29504.25	29560.66	
Neuroticism										
One-Factor	562.81	35	< .001	.75	.81	.12	.06	30415.04	30471.44	-.99
Two-Factor	562.75	34	< .001	.75	.81	.12	.06	30414.83	30471.24	
SDO										
One-Factor	1288.90	104	< .001	.77	.80	.10	.07	56394.38	56481.63	-.74
Two-Factor	607.17	103	< .001	.90	.92	.07	.05	55484.95	55574.01	
BDW										
One-Factor	416.17	54	< .001	.79	.83	.08	.06	51055.75	51121.09	-.84
Two-Factor	368.96	53	< .001	.85	.81	.07	.05	50995.90	51063.06	

Notes. SDO = Social Dominance Orientation; BDW = Belief in a Dangerous World.

Supplementary Table S3
Averaged Model Fit Estimate in Both Sample 1 and 2

		Absolute Fit Indices							Predictive Fit Indices	
		χ^2	df	p	TLI	CFI	RMSEA	SRMR	AIC	SABIC
		% Successful Convergence Across Datasets								
Sample 1										
Baseline model	100%	3299.21 (122.40)	666	< .001	.83 (.01)	.85 (.01)	.06 (.001)	.06 (.002)	99077.26 (622.22)	99355.79 (622.22)
M _{2 valence}	100%	1875.41 (113.32)	626	< .001	.92 (.01)	.93 (.01)	.04 (.002)	.05 (.002)	97537.14 (662.59)	97888.49 (662.59)
M _{common}	100%	2279.55 (112.91)	627	< .001	.89 (.01)	.91 (.01)	.05 (.002)	.05 (.002)	97955.08 (646.66)	98304.61 (646.66)
Sample 2										
Baseline model	100%	2869.60 (167.02)	384	< .001	.83 (.01)	.85 (.01)	.07 (.002)	.06 (.004)	78817.51 (688.74)	79387.15 (688.74)
M _{2 valence}	100%	1464.58 (142.90)	353	< .001	.92 (.01)	.93 (.01)	.05 (.003)	.05 (.004)	77339.15 (741.83)	77616.83 (741.83)
M _{common}	100%	1917.05 (160.32)	354	< .001	.88 (.01)	.91 (.01)	.06 (.003)	.05 (.004)	77783.88 (718.85)	78507.48 (718.86)

Notes. Numbers within the brackets are standard deviations of the corresponding fit indices across 100 parcel-level datasets from a random item-to-parcel allocation distribution. TLI = Tucker Lewis Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; AIC = Akaike Information Criterion; SABIC = Sample-size Adjusted Bayesian Information Criterion. Maximum likelihood estimator with robust standard errors is used in these test results.

Supplementary Table S4

Averaged Standardized Factor Loadings for the Model with Both Latent Constructs and Two Valence Method Factors ($M_{2\text{ valence}}$ Model)

construct	Sample 1			Sample 2		
	Factor loading on			Factor loading on		
	construct factor	positively-valenced factor	negatively-valenced factor	construct factor	positively-valenced factor	negatively-valenced factor
CONSC	.66	.30	.27	.64	.25	.17
EXTRA	.73	.36	.20	.74	.37	.11
AGREE	.56	.56	.26	.60	.45	.19
OPEN	.63	.31	.19	.65	.24	.10
NEURO	.70	.25	.18	.75	.22	.08
SE	.75	.32	.17	.79	.27	.10
SDO	.67	.24	.61	.66	.11	.57
BDW	.61	.18	.23			
BZSR	.76	.22	.14			
Variance Explained						
	construct	valence	residual	construct	valence	residual
	45.98%	9.69%	44.33%	49.78%	7.98%	42.24%

Notes. CONSC = Conscientiousness; AGREE = Agreeableness; EXTRA = Extraversion; OPEN = Openness to Experience; NEURO = Neuroticism; SE = Self-Esteem; SDO = Social Dominance Orientation; BDW = Belief in a Dangerous World; BZSR = Belief in Zero-Sum Resources; PL = parcels for positive-keyed constructs; NL = parcels for negative-keyed constructs

Supplementary Table S5

Averaged Model Fit Estimate in Both Sample 1 and 2 (Excluding Social Dominance Orientation and Agreeableness)

		Absolute Fit Indices							Predictive Fit Indices	
	% Successful Convergence Across Datasets	χ^2	df	p	TLI	CFI	RMSEA	SRMR	AIC	SABIC
Sample 1										
Baseline model	100%	1598.59 (95.80)	356	< .001	.88 (.01)	.90 (.01)	.06 (.002)	.05 (.003)	72956.10 (543.55)	73495.74 (543.55)
M _{2 valence}	100%	953.24 (97.56)	326	< .001	.93 (.01)	.95 (.01)	.04 (.003)	.04 (.003)	72247.74 (563.15)	72937.28 (563.15)
M _{common}	100%	1148.58 (114.65)	327	< .001	.91 (.01)	.93 (.01)	.05 (.003)	.05 (.003)	72444.27 (567.67)	72693.68 (567.67)
Sample 2										
Baseline model	100%	1270.34 (128.68)	160	< .001	.88 (.01)	.90 (.01)	.07 (.004)	.06 (.01)	48018.22 (566.98)	48155.10 (566.98)
M _{2 valence}	100%	692.58 (108.05)	139	< .001	.93 (.01)	.95 (.01)	.06 (.01)	.04 (.01)	47375.93 (514.48)	47842.93 (514.48)
M _{common}	100%	859.61 (114.09)	140	< .001	.91 (.02)	.93 (.01)	.06 (.01)	.05 (.004)	47548.45 (563.92)	47724.43 (563.92)

Notes. Numbers within the brackets are standard deviations of the corresponding fit indices across 100 parcel-level datasets from a random item-to-parcel allocation distribution. TLI = Tucker Lewis Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; AIC = Akaike Information Criterion; SABIC = Sample-size Adjusted Bayesian Information Criterion. Maximum likelihood estimator with robust standard errors is used in these test results.

Supplementary Table S6

Comparing the Nomological Networks among Positive Self-Esteem, Negative Self-Esteem, and Full-scale Self-Esteem After Unreliability Correction in Sample 1 and 2

	Sample 1				Sample 2			
	Positive self-esteem (PSE)	Full self-esteem (SE)	Negative self-esteem (NSE)	Expected Pattern of Correlational Strength	Positive self-esteem (PSE)	Full self-esteem (SE)	Negative self-esteem (NSE)	Expected Pattern of Correlational Strength
Full-scale scores								
Conscientiousness	.442	.476	.475		.367	.454	.376	
Extraversion	.546	.554	.521		.472	.472	.438	
Agreeableness	.267	.266	.245		.220	.237	.236	
Openness	.145	.106	.065		.010	.021	.058	
Neuroticism	-.862	-.911	-.894		-.857	-.891	-.867	
SDO	-.128	-.128	-.118		-.042	-.053	-.059	
BDW	-.197	-.215	-.218					
Positive-valenced scores				PSE>SE>NSE?				PSE>SE>NSE?
Conscientiousness	.436	.433	.400	Yes	.367	.384	.376	No
Extraversion	.554	.533	.475	Yes	.521	.506	.452	Yes
Agreeableness	.398	.343	.264	Yes	.360	.332	.270	Yes
Openness	.315	.223	.128	Yes	.146	.152	.011	No
Neuroticism	-.900	-.884	-.864	Yes	-.930	-.928	-.856	Yes
SDO	-.131	-.121	-.097	Yes	.003	-.025	-.054	No
BDW	-.214	-.196	-.166	Yes				
Negative-valenced scores				NSE>SE>PSE?				NSE>SE>PSE?
Conscientiousness	.369	.432	.460	Yes	.350	.421	.474	Yes
Extraversion	.475	.512	.509	No	.354	.374	.368	No
Agreeableness	.121	.162	.185	Yes	.121	.158	.190	Yes
Openness	.108	.087	.065	Yes	.002	.022	.047	Yes
Neuroticism	-.796	-.913	-.900	No	-.747	-.805	-.814	Yes
SDO	-.101	-.118	-.130	Yes	-.069	-.068	-.059	No
BDW	-.152	-.196	-.224	Yes				

Notes. SE = Self-Esteem; SDO = Social Dominance Orientation; BDW = Belief in a Dangerous World. BZSR was excluded from the analysis because it only had one positively-valenced item. The corrected correlations of this item with other variables could not be examined.

Supplementary Table S7
Comparing the Nomological Networks between Extraversion and Introversion in Sample 1 and Sample 2

	Sample 1			Sample 2		
	Half-Scale Extraversion (EXT)	Half-Scale Introversion (INT)	$\Delta\chi^2$	Half-Scale Extraversion (EXT)	Half-Scale Introversion (INT)	$\Delta\chi^2$
Full-scale scores						
Conscientiousness	.233	.262	0.73	.231	.186	1.63
Agreeableness	.253	.124	10.59** (EXT stronger)	.157	.007	18.93*** (EXT stronger)
Openness	.159	.178	0.25	.109	.111	0.26
Neuroticism	-.559	-.502	4.10* (EXT stronger)	-.507	-.376	22.24*** (EXT stronger)
SE	.532	.519	0.12	.504	.377	9.66* (EXT stronger)
SDO	-.066	-.083	1.02	-.004	-.032	0.49
BDW	-.057	-.103	2.34			
BZSR	-.067	-.030	1.74			
Positive-valenced scores						
Conscientiousness	.301	.222	5.06* (EXT stronger)	.303	.188	11.87*** (EXT stronger)
Agreeableness	.512	.258	41.35*** (EXT stronger)	.486	.229	55.75*** (EXT stronger)
Openness	.413	.388	0.47	.435	.375	1.50
Neuroticism	-.596	-.458	17.91*** (EXT stronger)	-.567	-.359	41.19*** (EXT stronger)
SE	.546	.469	5.57* (EXT stronger)	.518	.354	32.79*** (EXT stronger)
SDO	-.100	-.079	0.64	.043	-.047	0.39
BDW	.089	.081	0.05			
BZSR	.061	.113	1.47			
Negative-valenced scores						
Conscientiousness	.146	.265	13.20*** (INT stronger)	.152	.179	0.75
Agreeableness	.016	.002	0.17	.021	.098	5.81* (INT stronger)
Openness	.127	.190	2.54	.043	.081	1.67
Neuroticism	-.495	-.517	0.43	-.455	-.372	9.31*** (EXT stronger)
SE	.471	.510	1.54	.442	.363	7.86** (EXT stronger)
SDO	-.010	-.084	5.38* (INT stronger)	.037	-.017	0.32
BDW	-.009	-.095	7.32** (INT stronger)			
BZSR	.070	.034	1.15			

Notes. SE = Self-Esteem; SDO = Social Dominance Orientation; BDW = Belief in a Dangerous World; BZSR = Belief in Zero-Sum Resources. BDW and BZSR were not measured in Sample 2. * $p < .05$, ** $p < .01$, *** $p < .001$.

Supplementary Table S8

Comparing the Nomological Networks among Half-scale Extraversion, Half-scale Introversion, and Full-scale After Unreliability Correction in Sample 1 and 2

	Sample 1				Sample 2			
	Half-Scale Extraversion (EXT)	Full-Scale (FS)	Half-Scale Introversion (INT)	Expected Pattern of Correlational Strength	Half-Scale Extraversion (EXT)	Full-Scale (FS)	Half-Scale Introversion (INT)	Expected Pattern of Correlational Strength
Full-scale scores								
Conscientiousness	.233	.259	.262		.231	.220	.186	
Agreeableness	.253	.196	.124		.157	.083	.007	
Openness	.159	.174	.178		.109	.115	.111	
Neuroticism	-.559	-.562	-.502		-.507	-.471	-.376	
SE	.532	.554	.519		.504	.472	.377	
SDO	-.066	-.075	-.083		-.004	-.016	-.032	
BDW	-.057	-.082	-.103					
Positive-valenced scores				EXT>FS>INT?				EXT>FS>INT?
Conscientiousness	.301	.279	.222	Yes	.303	.264	.188	Yes
Agreeableness	.512	.408	.258	Yes	.486	.371	.229	Yes
Openness	.413	.416	.388	No	.435	.428	.375	Yes
Neuroticism	-.596	-.564	-.458	Yes	-.567	-.502	-.359	Yes
SE	.546	.537	.469	Yes	.518	.471	.354	Yes
SDO	-.100	-.092	-.079	Yes	.043	.003	.047	No
BDW	.089	-.090	.081	No				
Negative-valenced scores				INT>FS>EXT?				INT>FS>EXT?
Conscientiousness	.146	.210	.265	Yes	.152	.173	.179	Yes
Agreeableness	.016	.009	.002	No	.021	.063	.098	Yes
Openness	.127	.162	.190	Yes	.043	.065	.081	Yes
Neuroticism	-.495	-.531	-.517	No	-.455	-.438	-.372	No
SE	.471	.514	.510	No	.442	.429	.363	No
SDO	-.010	-.044	-.084	Yes	.037	.027	-.017	Yes
BDW	-.009	-.052	-.095	Yes				

Notes. SE = Self-Esteem; SDO = Social Dominance Orientation; BDW = Belief in a Dangerous World. BZSR was excluded from the analysis because it only had one positively-valenced item. The corrected correlations of this item with other variables thus could not be examined.