The Strengths and Difficulties Questionnaire: An Examination of Factorial, Convergent, and Discriminant Validity using Multitrait-Multirater Data

Supplemental Material

All the analyses were repeated using a robust weighted least square estimation method (WLSMV) for analyzing categorical data. We assumed item 3 of each factor as categorical and specified these items' thresholds to represent an underlying continuous response variable. These analyses were run using the Mplus program (Muthén & Muthén, 2010). We used delta parameterization, meaning that scale factors, but not residual variances, were allowed to be parameters in the model. We computed the chi-square difference test using the Mplus Difftest for WLSMV estimation. Standard errors were adjusted using the complex command in Mplus to further account for the clustered nature of data (children nested within classrooms).

As depicted in Table S1, the five-factor CFA solution with three indicators for each trait provided an acceptable overall fit to the data from teachers, $\chi^2(75) = 103.10$, $p = .002$, RMSEA = .03 (90% CI [.02; .05]), CFI = .94, TLI = .91, WRMSR = .73, mothers, $\chi^2(75) = 105.55$, $p < .012$, RMSEA = .03 (90% CI [.02; .05]), CFI = .95, TLI = .93, WRMSR = .88, and fathers, $\chi^2(75) = 116.35$, $p = .017$, RMSEA = .04 (90% CI [.03; .05]), CFI = .96, TLI = .94, WRMSR = .88.

Table S1 about here

The tests of Measurement Invariance (MI) across raters indicated that the changes in model fit between the configural model (Model A), $\chi^2(780) = 863.13$, $p = .020$, RMSEA = .02 (90% CI [.01; .02]), CFI = .94, TLI, = .92, WRMSR = .91, and the weak MI model (Model B, equality of loadings), $\chi^2(800) = 890.61$, $p = .014$, RMSEA = .02 (90% CI [.01; .03]), CFI = .93, TLI, = .91, WRMSR = .97, was negligible (Cheung & Rensvold, 2002), ΔCFI = 0.006, thereby establishing metric invariance.

The assumption of strong MI (equal loadings and intercepts/thresholds) across all three rater types did not hold as indicated by the significant decline in model fit when comparing the Models B (weak MI model) and C (strong MI model), $\Delta\chi^2(40) = 234.16$, $p < .001$, $\Delta\text{CFI} = 0.047$. We then tested a partial MI model (Model D) by removing the imposed equality constraints on the teacher report intercepts and thresholds of the items measuring children's Conduct Problems (CP) and Hyperactivity (Hy). Although the chi-square difference test was significant, $\Delta\chi^2(32) = 80.40$, $p = .008$, there was a minimal change of CFI between Model D and Model B (weak MI model), $\Delta\text{CFI} = 0.009$, suggesting that strong MI could be assumed across mothers and fathers ratings. Strong MI across parents (mothers and fathers) and teachers ratings was only established for the Emotional Symptoms (ES), Peer Problems (PP), and Prosocial Behaviors (PB) factors. Table S2 includes the factor loadings and items intercepts/thresholds in the measurement model with partial strong invariance.

Table S2 about here

Table S3 displays the latent means, variances, and correlations estimated in the CFA model with partial strict MI. The average latent correlation between teacher, mother, and father reports of the same trait (convergent validity correlations) was .48, ranging from .32 to .71. The latent correlations between the same traits reported by mothers and fathers ranged from .49 to .71, with an average correlation of .60, whereas the correlations between parents (mothers and fathers) and teacher reports of the same trait ranged from .32 to .62, averaging .43. The average absolute correlation between different traits reported by the same rater type (discriminant validity correlation) was .41 (range = .03; .90).

Table S3 about here

The CT-C(M – 1) model fit the data adequately, $\chi^2(620) = 653.11$, $p = .173$, RMSEA = .01 (90% CI [.00;.02]), CFI = .98, TLI = .96; WRMSR = .69. Table S4 shows the estimated factor loadings for the reference and method factors. The standardized teacher reference trait factor loadings for the parcels pertaining to mother and father reports were modest, ranging from .17 to .46. This reflected the rather modest convergence of parent and teacher reports. In contrast, parent-reported loadings on their respective method factors were larger than the reference factor loadings. Taken together, this indicated that there was more method specificity than convergent validity of parent reports relative to teacher reports. Father report indicators had similar loadings on the mother method factor (range = .27; .61) and on the method factor unique to father ratings (range = .29; .61).

Table S4 about here

Consistency, method specificity, and reliability coefficients are presented in Table S5. The consistency coefficients were very low for all traits (range = .03; .21). On average, the teacher reference trait factors explained less than 10% of the variance in the indicators of mothers and fathers, indicating low convergent between parents, and teacher reports.

Table S5 about here

Method-specificity coefficients were high for most of the mother ratings (range = .31; .75), indicating that a large proportion of variance in parent reports was not shared with teacher reports. On average, father ratings shared 24% (range = 7%; 37%) of their variance with mother ratings. On the other hand, on average, 26% (range = 8%; 37%) of the variance in father ratings was unique to fathers in that it was neither shared with teachers nor with mothers. The reliability coefficients (i.e., the proportion of variance in each observed indicator that represented true score variance and was not due to measurement error) were moderate to high for all indicators, ranging from .37 to .96 (*M*

= .64). The average reliability coefficient was higher for mother indicators (M = .66) than for father

indicators (M = .59).

References

Muthén, L., & Muthén, B. (2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén &

Muthén.

Table S1

Model Fit Information for Different CFA Models.

Model tested	Goodness of Fit					model comparison		
	X^2(df)	RMSEA (90% CI)	CFI	TLI	WRMSR	Compared Model	$\Delta X^2(\Delta df)$	ΔCFI
CFA Model								
Teacher report	103.10 (75)[**]	.03 (.02; .05)	.94	.91	.73	---	---	---
Mother report	105.55 (75)[*]	.03 (.02; .05)	.95	.93	.88	---	---	---
Father report	116.35 (75)[*]	.04 (.03; .05)	.96	.94	.88	---	---	---
MI Model								
Configural MI (A)	863.13 (780)[*]	.02 (.01; .02)	.94	.92	.91	---	---	---
Metric MI (B)	890.61 (800)[*]	.02 (.01; .03)	.93	.91	.97	A	36.81 (20)[*]	.006
Strong MI (C)	991.66 (840)[**]	.02 (.02;.03)	.88	.86	1.13	B	234.16 (40)[**]	.047
Partial Strong MI (D)	933.67 (832)[**]	.02 (.01;.03)	.92	.91	1.01	B	80.40 (32)[**]	.009

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

Table S2

CFA (Partial Strong MI model) Factor Loadings and Item's Intercepts.

	Factor loadings (US/S)			Item intercepts/thresholds (US)		
	Item 1[a]	Item 2	Item 3	Item 1	Item 2	Item 3
Teacher						
ES	1/.63	0.68/.52	1.37/.60	1.15	0.38	0.22; 1.74
PP	1/.59	0.92/.74	2.00/.63	0.61	0.45	0.95; 2.24
CP	1/.72	1.21/.65	1.77/.67	0.50	0.59	1.12; 2.11
Hy	1/.70	1.16/.73	1.27/.63	1.09	1.15	0.10; 1.47
PB	1/.67	1.26/.82	2.63/.85	3.51	3.03	-3.08; -0.58
Mother						
ES	1/.54	0.68/.52	1.37/.58	1.15	0.38	0.22; 1.74
PP	1/.45	0.92/.40	2.00/.55	0.61	0.45	0.95; 2.24
CP	1/.69	1.21/.74	1.77/.69	1.03	0.99	-0.41; 1.55
Hy	1/.69	1.16/.76	1.27/.68	1.55	1.96	-0.28; 1.59
PB	1/.56	1.26/.58	2.63/.71	3.51	3.03	-3.08; -0.58
Father						
ES	1/.67	0.68/.56	1.37/.63	1.15	0.38	0.22; 1.74
PP	1/.52	0.92/.55	2.00/.65	0.61	0.45	0.95; 2.24
CP	1/.62	1.21/.77	1.77/.67	1.03	0.99	-0.41; 1.55
Hy	1/.67	1.16/.80	1.27/.67	1.55	1.96	-0.28; 1.59
PB	1/.56	1.26/.64	2.63/.74	3.51	3.03	-3.08; -0.58

Note. ES = Emotional Symptoms; PP = Peer Problems; CP = Conduct Problems; Hy = Hyperactivity; PB = Prosocial Behaviors; US = Unstandardized Estimates; S = Standardized Estimates; [a] Parameter fixed for identification.

Table S3

CFA (Partial Strong MI model) Correlations.

	Correlations														
	Teacher					Mother					Father				
	ES	PP	CP	Hy	PB	ES	PP	CP	Hy	PB	ES	PP	CP	Hy	PB
Teacher															
ES	---														
PP	.73**	---													
CP	.03	.29**	---												
Hy	.03	.28**	.90**	---											
PB	-.29**	-.58**	-.42**	-.48**	---										
Mother															
ES	.41**	.30**	-.17*	-.06	-.08	---									
PP	.33**	.35**	-.09	.10	-.30**	.86**	---								
CP	.11	.08	.40**	.37**	-.17**	.22**	.34**	---							
Hy	.01	.11	.39**	.62**	-.19**	.24**	.25*	.67**	---						
PB	.02	-.16	-.24**	-.26**	.33**	-.17*	-.31**	-.34**	-.32**	---					
Father															
ES	.51**	.28**	-.02	-.10	.00	.63**	.49**	.10	-.01	-.16	---				
PP	.39**	.41**	-.04	-.12	-.14	.42**	.49**	.13	-.00	-.27*	.81**	---			
CP	-.03	.09	.40**	.27**	-.12	.00	.12	.65**	.49**	-.38**	.34**	.41**	---		
Hy	-.07	.13	.48**	.51**	-.13*	-.02	.06	.56**	.71**	-.28**	.24**	.23**	.79**	---	
PB	-.08	-.28**	-.17	-.06	.32**	-.10	-.36**	-.32**	-.21**	.51**	-.32**	-.49**	-.53**	-.38**	---

Note. ES = Emotional Symptoms; PP = Peer Problems; CP = Conduct Problems; Hy =

Hyperactivity; PB = Prosocial Behaviors; SE = Standard Error; *p* < .05, **p* < .01.

Table S4

Range of Factor Loadings Obtained in the CT-C(M − 1) Analysis.

	Standardized Reference Factor Loadings (Range)			Standardized Method Factor Loadings[a] (Range)		Standardized Method Factor Loadings[b] (Range)
	Teacher Report[c]	Mother Report	Father Report	Mother Report[c]	Father Report	Father Report
Emotional Symptoms	[.63, .96]	[.30, .31]	[.21, .38]	[.68, .85]	[.27, .61]	[.43, .61]
Peer Problems	[.68, .98]	[.18, .46]	[.17, .42]	[.55, .87]	[.45, .58]	[.42, .51]
Conduct Problems	[.63, .96]	[.24, .36]	[.22, .34]	[.71, .79]	[.39, .56]	[.43, .58]
Hyperactivity	[.70, .97]	[.36, .46]	[.29, .42]	[.64, .72]	[.52, .60]	[.29, .58]
Prosocial Behaviors	[.74, .82]	[.17, .40]	[.18, .37]	[.72, .80]	[.43, .50]	[.53, .60]

Note. Model used teacher ratings as indicators of the reference factor. (a) Method factor combining

items from mother and father reports; (b) Method factor including items from father report; (c)

Parameter fixed for identification; All factor loadings were significantly different from zero (p

< .05).

Table S5

Variance Components in the CT-C(M − 1) Model with Teacher Reports as Reference Method.

Trait and Method	Consistency	Method specificity[a]	Method specificity[b]	Reliability
Emotional Symptoms				
Teacher 1				.49
Teacher 2				.39
Teacher 3				.93
Mother 1	.10	.46		.56
Mother 2	.09	.62		.71
Mother 3	.10	.72		.82
Father 1	.14	.16	.18	.48
Father 2	.04	.07	.37	.48
Father 3	.11	.37	.22	.70
Peer Problems				
Teacher 1				.71
Teacher 2				.46
Teacher 3				.96
Mother 1	.06	.31		.37
Mother 2	.03	.72		.75
Mother 3	.21	.75		.96
Father 1	.04	.34	.20	.58
Father 2	.03	.20	.26	.49
Father 3	.17	.21	.18	.56
Conduct Problems				
Teacher 1				.41
Teacher 2				.61
Teacher 3				.91
Mother 1	.09	.51		.60
Mother 2	.13	.52		.65
Mother 3	.06	.61		.67
Father 1	.07	.15	.27	.49
Father 2	.05	.23	.33	.61
Father 3	.11	.32	.18	.61
Hyperactivity				
Teacher 1				.50
Teacher 2				.64
Teacher 3				.95
Mother 1	.13	.43		.56
Mother 2	.19	.45		.64
Mother 3	.21	.40		.61
Father 1	.08	.27	.30	.65
Father 2	.18	.36	.08	.62
Father 3	.15	.33	.34	.82
Prosocial Behaviors				
Teacher 1				.63
Teacher 2				.67
Teacher 3				.55
Mother 1	.07	.51		.58
Mother 2	.03	.63		.66
Mother 3	.16	.57		.73
Father 1	.03	.19	.36	.58
Father 2	.04	.19	.28	.51
Father 3	.13	.25	.30	.68

Note. (a) Method factor combining items from mothers and fathers reports; (b) Method factor including items from fathers report.