Evaluating Group and Longitudinal Measurement Equivalence in a Battery of Cognitive Self-Regulation Measures for Preschoolers

VANDERBILT UNIVERSITY

Mary Wagner Fuhs & Kimberly A. Turner

PEABODY research institute

INTRODUCTION

- Self-regulation research largely focuses on cross-sectional analyses of laboratory-based measures that may not be optimal for longitudinal field-based research in schools.
- Establishing group and longitudinal ME is a prerequisite to drawing conclusions about latent mean gender differences and latent mean changes in self-regulation across time.
- The purpose of this study was to evaluate whether or not ME can be confirmed between males and females as well as longitudinally in a battery of cognitive self-regulation measures suitable for field-based research.

METHOD

Participants

535 preschoolers (48% female), mean age was 54 months (*SD* = 3 months; range 46 – 65 months). Children were recruited from 58 ethnically and economically diverse prekindergarten classrooms.
All participants were native English speakers (e.g., passed Pre-LAS)[1].

Measures and Procedure

Head Toes Knees Shoulders (HTKS)[2]
Peg Tapping (PT)[3]
Copy Design (CD)4
KRISP[5]
Dimensional Change Card Sort (DCCS)[6]
Backwards Digit Span[7]

Preschoolers completed the battery of measures in the fall (T1) and spring (T2) of preschool in a fixed order.

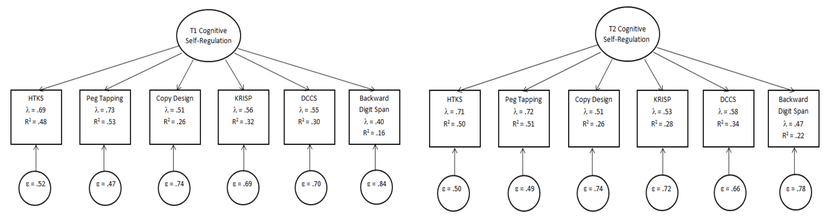
ACKNOWLEDGEMENTS

The project from which these data are drawn is funded by the Institute of Education Sciences (R305A090079; PIs: Mark W. Lipsey & Dale C. Farran). The first author is supported by an Institute of Education Sciences Postdoctoral Fellowship (R305B100016).

RESULTS

Factor Structure Summary

ME was evaluated using CFA in Mplus V. 6.11[8]. At T1 ($\chi^2(9) = 29.53$, $p < .001$; RMSEA = .065; CFI = .967; SRMR = .031) and T2 ($\chi^2(9) = 33.57$, $p < .001$; RMSEA = .071; CFI = .962; SRMR = .033), one-factor models were preferred over alternative models.



Measurement Equivalence Summary

Results with PT as the referent indicator supported strong ME for gender at T1 and partial ME for gender at T2. Model fit was significantly worse when constraining latent means to be equal at both time points. With girls as the referent group having a latent mean of 0, latent means for boys were significantly worse at T1 (- .36, *SE* = .10) and T2 (-.47, *SE* = .10). Tests of longitudinal ME revealed that weak ME did not hold.

Tests of Measurement Equivalence

Model Description	Chi-Square	DF	*p*	RMSEA	CFI	SRMR
Tests of Measurement Equivalence for Gender at T1						
T1 Gender ME BASELINE	33.21	18	0.016	0.056	0.974	0.033
T1 Gender ME LOADINGS constrained	36.15	23	0.039	0.046	0.978	0.039
T1 Gender ME INTERCEPTS constrained	45.49	28	0.019	0.048	0.97	0.048
T1 Gender ME Latent Means constrained	59.41*	29	0.001	0.063	0.948†	0.076
Tests of Measurement Equivalence for Gender at T2						
T2 Gender ME BASELINE	37.26	18	0.005	0.063	0.968	0.035
T2 Gender ME LOADINGS constrained	48.32	23	0.002	0.064	0.958	0.078
T2 Gender ME INTERCEPTS constrained	73.17*	28	<.001	0.078	0.925†	0.114
T2 Gender ME INTERCEPTS KRISP Backward Digit Span free	49.17	26	0.004	0.058	0.961	0.078
T2 Gender ME Latent Means Constrained	73.26*	27	< .001	0.08	0.923†	0.113
Tests of Longitudinal Measurement Equivalence						
Longitudinal ME BASELINE	88.43	47	<.001	0.041	0.981	0.034
Longitudinal ME LOADINGS constrained	163.21*	52	<.001	0.063	0.949†	0.084

* Denotes significantly worse fit compared to previous model using a Likelihood Ratio Test. † Denotes significantly worse fit compared to previous model using alternative ΔCFI criteria proposed by Cheung & Rensvold (2002).

Evaluating Non-ME Sets of Variables Longitudinally

Simulation studies and empirical evidence suggest that the choice of reference indicator can affect tests of ME, particularly if the referent indicator is noninvariant[9]. In the first set of ME analyses, the referent indicator was PT. An examination of loadings across time suggested several variables could be noninvariant. Thus, we used factor ratio tests (LRT) and the stepwise partitioning procedure to empirically determine an appropriate referent invariant indicator or set of indicators[10,11].

Eight pairs were noninvariant, resulting in three final invariant sets: 1) HTKS, CD, DCCS, 2) PT, KRISP, DCCS, and 3) PT, CD, DCCS.

Stepwise Partitioning Procedure

Variables	Noninvariant Sets*						
1.PT 2.HTKS 3.CD 4.DS 5.KRISP 6.DCCS	(1,2)	23456	13456				
	(1,4)	23456	3456	1356			
	(2,4)	3456	2356	3456	1356		
	(2,5)	3456	356	236	1356		
	(3,4)	456	356	236	1356		
	(3,5)	456	56	36	236	156	136
	(4,5)	56	46	36	236	156	136
	(4,6)	56	36	236	156	136	
Final Sets	236	156	136				

*Noninvariant sets determined by 15 comparisons of each variable pair with the baseline model. Models that fit worse than the baseline model according to an LRT were considered noninvariant. Because 15 tests were performed, the required *p* value was .05/15 = .003.

Using DCCS as a referent indicator and invariant Set 3, we found noninvariance using LRT critical values but strong ME using alternative CFI estimates. Children's latent means were significantly greater at T2 (referent latent mean at T1 set to 0) (.91, *SE* = .06).

Tests of Longitudinal Measurement Equivalence: PT, CD, and DCCS

	Chi-Square	DF	*p*	RMSEA	CFI	SRMR
Longitudinal ME BASELINE	88.43	47	<.001	0.041	0.98	0.03
Longitudinal ME LOADINGS constrained	95.52*	49	< .001	0.042	0.98	0.04
Longitudinal ME INTERCEPTS constrained	102.86*	51	< .001	0.044	0.98	0.04
Longitudinal ME Latent Means Constrained	500.04*	52	< .001	0.127	0.79†	0.09

* Denotes a significantly worse fit compared to the previous model using a Likelihood Ratio Test. † Denotes a significantly worse fit compared to previous model using alternative ΔCFI criteria proposed by Cheung & Rensvold (2002).

DISCUSSION

- A cognitive self-regulation battery was suitable for examining gender latent mean differences.
- Empirically identifying the best referent indicator and invariant sets using LRTs and the stepwise partitioning procedure yielded three possible invariant sets of indicators. Using set 3 with DCCS as the referent indicator, CFI criteria but not LRT criteria revealed adequate fit to test latent mean differences.
- The cognitive self-regulation construct may be changing across time and may predict children's performance on measures of cognitive self-regulation differently depending on children's age and/or maturation level.
- This study highlights the importance of explicitly testing for ME before making assumptions about the stability of a latent construct across time, particularly for young children.
- The implications of partial ME for developmental methodology, and more specifically for studying the development of critical school readiness skills such as self-regulation, are important areas for future research.

REFERENCES

1. Duncan, S. E., & De Avila (1986). *Pre-LAS user's manual.* San Rafael, CA: Linguametrics.
2. Ponitz, C. C., McClelland, M. M., Matthews, J. S., & Morrison, F. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology, 45*, 605-619. doi: 10.1037/a0015365
3. Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do". *Developmental Psychobiology, 4*, 315-334. doi: 10/1002/(SICI)1098-2302(199605)
4. Davie, R., Butler, N., & Goldstein, H. (1972). *From birth to seven: The second report of the National Child Development Study.* London: Longman.
5. Wright, J. C. (1971). *The Kansas Reflection Impulsivity Scale for Preschoolers.* St. Louis, MO: CEMREL.
6. Zelazo, P. D. (2006). The Dimensional Change Card Sort: A method of assessing executive function in children. *Nature Protocols, 1*, 297-301. doi: 10.1038/nprot.2006.46
7. Davis, H. L., & Pratt, C. (1996). The development o children's theory of mind: The working memory explanation. *Australian Journal of Psychology, 47*, 25-31.
8. Muthén, L. K., & Muthén, B. O. (2007). Mplus user's guide (5th ed.). Los Angeles: Muthén & Muthén.
9. Johnson, E. C., & Meade, A. W. (2007, April). *The role of referent indicators in tests of measurement invariance.* Paper presented at the 22nd Annual Meeting of the Society for Industrial and Organizational Psychology, New York.
10. French, B. F., & Finch, H. (2008). Multigroup confirmatory factor analysis: Locating the invariant referent sets. *Structural Equation Modeling, 15*, 96-113. doi: 10.1080/10705510701758349
11. Cheung , G. W., & Rensvold , R. B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management, 25*, 1-27.