Using the inCLASS to Measure Preschool Children's Engagement with Teachers, Peers and Tasks:
Examining Measurement Invariance Across Gender, Ethnicity, and Poverty Status in Three Samples

Natalie L. Bohlmann, Jason T. Downer, Leslie M. Booren, Michelle F. Maier, and Amanda P. Williford
University of Virginia, Center for the Advanced Study of Teaching and Learning (CASTL)

UNIVERSITY of VIRGINIA
CASTL
CURRY SCHOOL OF EDUCATION

inCLASS OBSERVATION
Individualized Classroom Assessment Scoring System

Background

A child's readiness for school may best be understood through the nature and quality of their engagement with adults, peers, and learning activities in the classroom (Ladd et al., 2006). In early childhood, positive engagement with adults predicts educational achievement (Burchinal et al., 2002), children's peer interactions forecast social competence and adjustment to kindergarten (Fantuzzo et al., 2005), and positive engagement with learning tasks is associated with higher academic achievement (Stipek, et al., 2010).

To increase our understanding of children's engagement within early childhood classrooms, there is a need for more research that focuses on child-level interactions. The *Individualized Classroom Assessment Scoring System* (inCLASS) is an observational tool designed to assess children's engagement with teachers, peers, and tasks. In order to use the instrument to understand patterns of classroom engagement and its role in facilitating school readiness, it is necessary to confirm the observation system's utility in a large, diverse sample of preschoolers that reflect the ethnic and socioeconomic landscape of American families.

Research Aims

The goals of the current study were to:

1) examine measurement invariance of a 4-factor model of the inCLASS (positive engagement with teachers, peers, tasks, and negative classroom engagement) across three large samples with different demographic characteristics

2) confirm this 4-factor model using the full, combined set of children across all three samples

3) establish measurement invariance in this combined sample across demographic subgroups (gender, poverty, ethnicity) to allow examination of mean differences in classroom engagement across these demographic categories

Observing Children's Engagement

The *Individualized Classroom Assessment Scoring System* (inCLASS) measures an individual child's engagement in the classroom using 4 domains and 10 dimensions (Downer et al., 2010; see Figure 1).

Each dimension was rated on 7-pt scale (1=low, 7=high) using behavioral markers in a standardized manual. Dimension scores were averaged across all cycles.
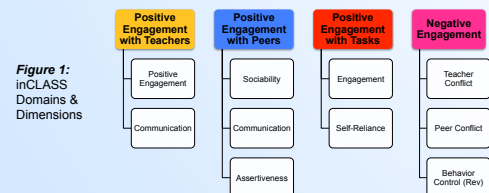
Figure 1: inCLASS Domains & Dimensions

Positive Engagement with Teachers	Positive Engagement with Peers	Positive Engagement with Tasks	Negative Engagement
Positive Engagement	Sociability	Engagement	Teacher Conflict
Communication	Communication	Self-Reliance	Peer Conflict
	Assertiveness		Behavior Control (Rev)

Methods

Participants: 925 children (466 girls and 459 boys, 44% African American, 37% Hispanic, 9% White, *M=54.89, SD=6.94* age in months) from 305 preschool classrooms.
  Sample 1: 341 children (169 girls and 172 boys, 46% African American, 32% Hispanic)
  Sample 2: 352 children (180 girls and 172 boys, 72% Hispanic, 15% White)
  Sample 3: 232 children (117 girls and 115 boys, 76% African American, 19% Hispanic)

Procedure: Data were collected as part of three separate research studies. Observations were conducted alternating four cycles (10 minutes observing, 5 minutes coding) across the morning on randomly selected children (*M=3.03, SD=1.85*) in each classroom.

Results: Measurement Invariance

Aims: 1) Examine measurement invariance of a 4-factor model of the inCLASS across three samples, 2) confirm the 4-factor model in a combined sample, 3a) establish invariance in the combined sample across demographic groups (see Table 1).

Multi-group CFA following a step-wise process of testing *configural, weak, strong,* and where needed *partial intercept* invariance was used to examine measurement invariance across three samples and across demographic groups. *Weak* factorial invariance (factor loadings invariant across groups), followed by *strong* invariance (intercepts of measured variables equal across groups) was tested (Meredith, 1993). Where *strong* invariance was not met, *partial intercept* invariance (some, but not all, intercepts invariant across groups) was then explored (Reise, Widaman, & Pugh, 1993).

Table 1: *Fit Indices for the Combined Sample and for the Baseline, Weak, Strong, and Partial Intercept Invariance Models for Each Multi-group Comparison*

Sample	$X^2(df)$	CFI	RMSEA (CI)	SRMR	X^2diff^a
Aim1: Three Samples					
Separate Groups					
Sample 1 (n = 341)	96.38 (27)	.94	.09 (.07, .11)	.05	
Sample 2 (n = 352)	110.47 (27)	.91	.09 (.07, .11)	.05	
Sample 3 (n = 232)	63.97 (27)	.93	.08 (.05, .10)	.06	
Configural (baseline -all parameters free)	264.59 (81)	.93	.09 (.07, .10)	.05	
Weak Invariance	263.17 (93)	.93	.08 (.07, .09)	.06	13.66(12); ns
Strong Invariance	587.34 (105)	.81	.12 (.11, .13)	.12	292.82(12); p < .05
Partial Intercept Invariance	269.75(99)	.93	.08 (.07, .09)	.07	8.61(6); ns
Aim 2: Four-factor Model Combined Sample					
Combined Sample (N = 925)	154.90(27)	.95	.07 (.06, .08)	.06	
Aim 3a: Demographic Groups					
Gender					
Separate Groups					
Boys (n = 459)	103.02(27)	.94	.08 (.06, .10)	.07	
Girls (n = 466)	101.28(27)	.95	.08 (.06, .09)	.06	
Configural (baseline -all parameters free)	204.26(54)	.94	.08 (.07, .09)	.07	
Weak Invariance	212.74(60)	.94	.07 (.06, .09)	.07	10.21(6); ns
Strong Invariance	227.09(66)	.94	.07 (.06, .08)	.07	11.44(6); ns
Poverty Status					
Separate Groups					
Poor (n = 524)	93.85(27)	.95	.07 (.05, .08)	.06	
Non-poor (n = 312)	107.92(27)	.94	.10 (.08, .12)	.07	
Configural (baseline -all parameters free)	199.79(54)	.94	.08 (.07, .09)	.07	
Weak Invariance	205.27(60)	.94	.07 (.06, .09)	.08	10.85(6); ns
Strong Invariance	248.16(66)	.93	.08 (.07, .09)	.08	42.89(6); p < .05
Partial Intercept Invariance	207.28(63)	.94	.07 (.06, .09)	.07	0.54(3); ns
Ethnicity					
Separate Groups					
African American (n = 342)	85.55(27)	.94	.08 (.06, .10)	.07	
Hispanic (n = 406)	79.72(27)	.96	.07 (.05, .09)	.06	
Configural (baseline -all parameters free)	165.14(54)	.95	.07 (.06, .09)	.07	
Weak Invariance	162.97(60)	.95	.07 (.06, .08)	.07	2.28(6); ns
Strong Invariance	281.13(66)	.90	.09 (.08, .11)	.08	137.03(6); p < .05
Partial Intercept Invariance	167.94(62)	.95	.07 (.06, .08)	.07	4.97(2); ns

Note: Incremental fit indices (CFI) above 0.90 signify good model fit (Browne & Cudeck, 1992). RMSEA values between .05 and .08 signify acceptable model fit, and values between .08 and .10 as mediocre fit (MacCallum, Browne, & Sugawara, 1996). SRMR values lower than .08 are considered a good fit (Hu & Bentler, 1998).

Results: Demographic Comparisons

Aim 3b: Group comparisons suggest some differences in children's observed engagement across demographic groups:

Positive Engagement with Teachers

Positive Engagement with Peers

Positive Engagement with Tasks

Negative Engagement

• Girls were more positively engaged with tasks (t(923)= 3.11, p = .002) and had less negative classroom engagement (t(923)= -4.77, p = .001) than boys.

• Children from impoverished families were less positively engaged with tasks (t(834) = -2.34, p = .02) and more negatively engaged in the classroom (t(834) = 4.41, p = .001) than non-poor children.

• African Americans had more positive teacher engagement (t(746) = 4.18, p = .001), peer engagement (t(746) = 3.97, p = .001), and task engagement (t(746) = 5.69, p = .001) than Hispanics.

Discussion

Establishing measurement invariance is critical for measures designed for use with diverse populations as it provides evidence that an assessment taps into the same underlying structure across groups and lends support for the measure's utility in practical applications (Drasgow & Kanfer, 1985; Horn & McArdle, 1992).

• Consistent with findings from initial validation work on the inCLASS (Downer et al., 2010), a four-factor model, namely positive engagement with teachers, peers, and tasks and negative classroom engagement, fit the data well within all three samples.

• When making comparisons across the three large diverse samples, criteria for *weak* invariance, as well as *partial intercept* invariance, were met. These results suggest that the inCLASS is applicable and operates similarly across many different preschool classrooms.

• Four-factor structure held across gender, poverty status, and ethnicity. *Strong* measurement invariance was demonstrated across boys and girls, and *partial intercept* invariance was met for poverty status and ethnicity (Hispanic vs. Black) suggesting that the inCLASS maintains similar measurement properties across different demographic groups.

• Group comparisons indicated some differences in children's engagement across demographic groups in ways that parallel findings reported in the literature (Burchinal et al., 2000; Coolahan et al., 2000)

Implications & Future Directions

Implications
•Results demonstrate invariant measurement properties on the inCLASS across ethnically and economically diverse samples, providing validation for the observation system's utility in examining individual children's experiences in preschool.

• The inCLASS is poised to help the field better understand the role that a child's engagement in the classroom may play in their preparation for school.

Limitations and Future Directions:
• Demographic group comparisons did not consider the role of classroom context and therefore may not accurately represent population differences in children's engagement.

• Future work should examine children in context to determine the degree to which observed differences are attributable to contextual factors before any conclusions are drawn.

Acknowledgements