

APPENDIX 3: inCLASS Technical Manual

INTRODUCTION

The Individualized Classroom Assessment Scoring System (inCLASS) is an observational instrument developed to assess children's competence in classroom interactions with adults, peers, and learning activities. This technical manual provides information about the psychometric properties of the inCLASS, including descriptive statistics and information about reliability and validity. The data presented are from a pilot study in 44 preschool classrooms, field study in 104 preschool classrooms, and the Hands on Science project in 30 Head Start classrooms. In addition, the intent of this technical manual is to provide information about the conceptual framework for the inCLASS and the empirical underpinnings. The inCLASS is still in its early development.

CONCEPTUAL FRAMEWORK

For children between the ages of 3 and 5, their "readiness" to function competently in school is best understood in terms of the nature and quality of their behavioral, social, and language-based interactions in preschool classrooms with adults, with peers, and with learning/instructional activities. Classroom interactions provide perhaps the most valid indicator of the manner in which children make use of the learning and social opportunities provided in school, and observationally based information on children's competence in classroom interactions adheres to key professional standards for assessment (Neisworth & Bagnato, 2004), including usefulness for identification, intervention instructional improvement, evaluation, and even accountability (Shepard, Kagan, & Wurtz, 1998).

However, there is a shortage of psychometrically sound measurement tools that tap into these classroom context-specific, relational processes. Meisels and Atkins-Burnet (2006) indicate that the key aspects of classroom functioning that are relevant for the purposes of improving instruction and tailoring it to individual differences are *observations of patterns* of level of engagement, *qualities* of child behavior (such as fluidity and control) and their *interactions* with peers and adults. More specifically, they note that it is "insufficient to record that a child did or did not exhibit particular behaviors" (p. 11); rather, observation should seek to detect patterns of interaction and competence in response to contextual supports, constraints, and challenges (Meisels & Atkins-Burnet, 2006). For this reason, our approach to conceptualizing domains of development is to focus on global integrative units of analysis that reflect patterns of adaptation to core developmental tasks: competent interaction with adults, peers, and activities. These developmental tasks in turn relate to building effective social relationships and acquiring skills/knowledge through instructional opportunities (see Bruner, 1966; Pianta, 1999; or Sroufe, 1996 for a discussion of developmental tasks and global patterns of function).

OVERVIEW OF THE inCLASS

In its current format, the inCLASS is conceptually organized into three *domains*, examining interactions with Teacher, Peer, and Tasks (see Table 1). Within each domain, there are dimensions (e.g., Peer Assertiveness) that receive global scores on a 7-pt. scale, based on the observation of specific behavioral markers (e.g., initiation, leadership). These dimensions and behavioral markers were identified based on a review of the literature as well as currently utilized observational systems and teacher ratings of young children's behavior (NICHD ECCRN, 2008; Pianta, LaParo, & Hamre, 2008). We now present a brief overview of these core domains of development and adaptation that are the inCLASS observational assessment as it was originally conceptualized. Please note that this measure is still in early development and data is being collected, analyzed, and summarized.

Literature supporting the three domains of interaction is briefly summarized herein. Children's interactions in early education classroom settings have reliable and detectable effects on achievement and social competence both concurrently and in elementary school (e.g., Burchinal et al., 2002; Howes, Phillipsen, & Peisner-Feinberg, 2000; Meyer, Waldrop, Hastings, & Linn, 1993; NICHD ECCRN, 1996; NICHD ECCRN, 2003; NICHD ECCRN, 2004; Peisner-Feinberg & Burchinal, 1997). Characteristics of child-teacher interactions, such as use of the teacher as a source of support and help, are related to gains in children's performance in early childhood classrooms (Matsumura, Patthey-Chavez, Valdes, & Garnier, 2002; Nelson-LeGall & Resnick, 1998; NICHD ECCRN, 2005a). Children's initiation and maintenance of positive peer relationships is also a critical, developmentally salient task during early childhood (Guralnick, 1993; Ladd, 2005): early peer functioning (e.g., social acceptance) is linked to a range of developmental outcomes, including academic achievement (Gifford-Smith and Brownell, 2003).

Table 1

The inCLASS Domains, Dimensions, and Definitions

Domain	Dimension	Definition
Teacher Interactions	Positive Engagement with the Teacher	Measures the degree to which the child is emotionally connected to the teacher(s) and adults, including seeking and enjoying interactions with them, and using them as a secure base.
	Teacher Communication	Measures the degree to which the child initiates and maintains conversation with the teacher(s) and adults while using language as a functional tool to make needs, emotions, and opinions known (e.g., requesting, commenting, and questioning).
	Teacher Conflict	Measures the degree to which the child's interactions with the teacher(s) and adults are characterized by tension, resistance, and negativity.
Peer Interactions	Peer Sociability	Measures the degree to which the child experiences positive emotions and behaviors with other children, including the tendency to seek peer interactions, show social awareness and respond in a manner that peers react positively to.
	Peer Communication	Measures the degree to which the child initiates and maintains conversation with other children while using language as a functional tool to make needs, emotions, and opinions known (e.g., requesting, commenting, and questioning).
	Peer Assertiveness	Measures the degree to which the child uses positive strategies to initiate and lead interactions with other children, and the degree to which those strategies are successful.
	Peer Conflict	Measures the degree to which the child's interactions with other children are characterized by tension, resistance, and negativity.
Task Orientation	Engagement within Tasks	Measures the degree to which the child is consistently and actively involved in classroom tasks and activities, including the amount of time the child remains focused on any given activity, the level of intensity or enthusiasm displayed, and the proportion of time the child spends on assigned activities.
	Self-Reliance	Measures the degree to which the child takes learning into their own hands, including seeking opportunities rather than passively waiting for teacher direction, and making best use of classroom resources (including the teacher).
	Behavior Control	Measures the degree to which the child regulates movement, physical activity, and verbalizations, so that these match the expectations of the setting.

Meanwhile, children's task-oriented behavior is another crucial determinant of later school success: children with a more positive approach to learning in kindergarten (e.g., task persistence) show better pre-reading and math skills in kindergarten and first grade (Denton & West, 2002), and self-reliance is also a key mediator of children's achievement in this early schooling (NICHD ECCRN, 2005b). Finally, conflict in the classroom, though clearly not a competency, is an important indicator of future problems: teacher-child conflict in preschool and kindergarten is a strong predictor of later academic and behavioral difficulties (Howes, Phillipsen, & Peisner-Feinberg, 2000; Pianta, 2001), and peer conflict has implications for later peer relationships (Ostrov & Keating, 2004) and social adjustment (Crick, 1996).

In order to measure children's competency in these areas, as displayed in the preschool setting, the observation protocol for inCLASS uses a 15-minute cycle (10 min observing, 5 min scoring), and can rotate through four children in a classroom. Observers who stay for a full 4-hour morning can therefore complete four cycles per child (a total of 16 observations) during each visit. During the 10-min period, observers take notes on each dimension, recording the specific behavioral markers observed. During the 5-min scoring period, observers then use these notes to determine whether high, medium, or low competency was observed, consulting the inCLASS manual to assign a 1-7 score to each dimension. The inCLASS is *not a checklist* and observers are instructed to view the dimensions as holistic descriptions of children's patterns of behavior.

It is important to note that all users must obtain adequate training before attempting to use the inCLASS. The inCLASS is a multifaceted observation system that requires in-depth training for appropriate use. The level of training required depends on the intended use of the system. It is essential that all individuals interested in using the inCLASS to collect standardized data for research or other purposes attend official training workshops. There are up to date training options available on the inCLASS website (see <http://www.inclassobservation.com>).

STUDIES PROVIDING inCLASS DATA

The inCLASS data are being collected from multiple samples, each of which is described briefly below. Some of the data from these studies will be used in the statistics reported below and referred to by study name. It is important to note that the inCLASS is in early development, and many of these studies are still ongoing.

inCLASS Pilot Study

Primary Investigators: University of Virginia.

Study Summary: The purpose of this study was to pilot this new observational measure, and revise the content and/or protocol based on these findings. Nine inCLASS dimensions were included in this study.

Participants: 164 children (90 girls and 74 boys; mean age 4.10 years; 92% White), 40 teachers (all female; 95% White; 23% with BA degree) and 44 preschool classrooms (average size 15.36 children) participated.

Location: Central Virginia.

Sampling: All schools within a 45-mile radius of a small university town in central Virginia were invited to participate. Within each participating classroom, four children were randomly selected from those consented.

Time of year: Two visits, one week apart, in the fall; repeated in the spring.

Data collection procedure: During each visit, observers spent a full morning in the classroom; children were observed for an average of four cycles per visit (16 total cycles across the year). The teacher completed a packet of ratings (e.g., TCRS and STRS) for each child observed, and repeated these in the spring. Teachers also completed classroom and teacher demographic surveys. Parents completed a family demographics survey. See Downer et al. (2010) for more details.

inCLASS Field Study

Primary Investigators: University of Virginia and the University of California-Los Angeles.

Study Summary: The purpose of this study was to further validate the inCLASS measure by using a diverse preschool sample, and by following these children through a follow-up year of possible kindergarten entry. Ten inCLASS dimensions were included in this study, all the revised content from the Pilot Study plus one additional dimension entitled Behavior Control.

Participants: 381 children (192 girls and 189 boys; mean age 3.69 years; 66% Hispanic, 19% White), 88 teachers (all female; 63% Hispanic, 22% White; 24% with BA degree), and 104 preschool classrooms (average size 18.84 children) participated in year 1. 123 classrooms (89 kindergarten; average size 19.79 children) and 101 teachers (9 male; 45% Hispanic, 37% White; 33% with BA degree) were added in year 2.

Location: Greater Los Angeles metro area.

Sampling: Teachers who agreed to participate in a random sample of preschool classrooms in Los Angeles area. Four children were randomly selected from those consented in each class, and followed into classrooms the following year (most in kindergarten).

Time of year: Fall and spring (year 1), and following year fall visits (year 2, kindergarten entry).

Data collection procedure: Generally 3-4 observations were completed of each child during a visit using a 15-minute rotation (10 minutes of observation and 5 minutes of coding). Two children in each class were also randomly selected to complete a direct assessment battery (e.g., PPVT, etc.) for year 1, and all children followed into year 2 were assessed. The teacher also completed a packet of ratings (e.g., TCRS, STRS, etc.) on each of the children observed during each time point. Teachers also completed classroom and teacher demographic surveys. Parents completed a family demographics survey.

Early Childhood Hands on Science Study

Primary Investigators: University of Virginia and the University of Miami.

Study Summary: This sub-study was part of a larger study that was developing a new science measure. The purpose of this sub-study was to further validate the inCLASS measure by connecting it to other assessments in an ethnically diverse, at-risk preschool sample. Children were given direct assessments of school readiness (science, vocabulary, math, listening comprehension, and alphabet knowledge), receptive vocabulary, and executive functions. They were rated by teachers on classroom adjustment and approaches to learning. Ten inCLASS dimensions were included in this study, all the revised content from the Pilot Study plus one additional dimension entitled Behavior Control.

Participants: 242 children (126 girls and 116 boys; mean age 3.79 years; 75% African American), 30 teachers (all female; 53% African American, 37% Hispanic; 60% with BA degree), and 30 Head Start classrooms participated.

Location: Miami-Dade County, Miami, FL.

Sampling: Head Start centers within a 20-mile radius of the university were identified and invited to participate. Teachers were then contacted in those centers that agreed to participate and that reported that the majority of their enrolled children were English proficient. Within these classrooms, 8-10 children were randomly selected to participate in the study, and observed on different days in groups of 4.

Time of year: Fall and spring visits.

Data collection procedure: Trained coders observed three children per day for four observation cycles, using a 15-minute rotation (10 minutes of observation and 5 minutes of coding). Observations were conducted in the fall and spring. Teachers rated children's classroom adjustment in the fall. School readiness assessments were administered in the fall and spring. Receptive vocabulary was assessed mid-year. Teachers rated children's approaches to learning in the spring.

National Center for Research on Early Childhood Education (NCRECE)

Primary Investigators: University of Virginia, University of North Carolina at Chapel Hill, University of North Carolina at Greensboro, and University of California at Los Angeles.

Study Summary: A randomized, controlled evaluation of the effects of two forms of professional development support for teachers. These supports to teachers will improve the implementation of curricula and interactions with children, as well as, promote gains in children's social and academic development. These approaches will focus in particular on promoting language and literacy skills, domains of child development that operate as gatekeepers to later achievement. Ten inCLASS dimensions were included in this study, all the revised content from the Pilot Study plus one additional dimension entitled Behavior Control.

Participants: 354 children and 177 early childhood teachers/classrooms participated in Year 1. 268 children and 134 early childhood teachers/classrooms participated in Year 2.

Locations: Early childhood classrooms across 9 settings in six states.

Sampling: Teachers who agreed to participate in the randomized control trial testing the effects of a professional development intervention. Out of previously consented children, two were randomly selected for inCLASS observations.

Time of year: One observation day scheduled between January and March each year.

Observation procedure: Generally 3 observations were completed of each child during a visit using a 15-minute rotation (10 minutes of observation and 5 minutes of coding), per child, and alternating with a classroom wider observation assessment (CLASS).

The Foundations of Learning Study

Primary Investigators: MDRC.

Study Summary: This study builds on growing evidence that resolving children's early problem behaviors can provide the underpinning for a high-quality and effective preschool experience. Through random assignment, this study tests an intervention based on professional development and classroom consultation for teachers. Ten inCLASS dimensions were included in this study, all the revised content from the Pilot Study plus one additional dimension entitled Behavior Control.

Location: Newark, NJ.

Sampling: Teachers who agreed to participate in 26 program sites and 25 control sites; four children randomly selected from those consented in each participating classroom.

Time of year: Spring.

Data collection procedure: Trained coders used 15-minute observation cycles, and rotated through children in the classroom during a full morning, thus averaging 4 cycles per child. Other data were collected throughout the study.

GENERAL STATISTICS

What information does the inCLASS provide about children?

Table 2 provides descriptive information on the inCLASS scores that are currently available for the inCLASS pilot study, the inCLASS Field Study, and the Early Childhood Hands on Science Study. These data are presented by conceptual domain and include domain and dimension scores across time points.

The Pilot Study results indicated that most inCLASS dimensions show good variability (except for the Conflict dimensions, which are negatively skewed and limited in range). Scores for the other dimensions were adequately distributed across the 7-point scale, although Engagement shows some positive skew, and Peer Assertiveness and Teacher Communication some negative skew. These Pilot Study data are drawn from a single, highly homogeneous sample.

For the Field Study and the Hands on Science Study, a new dimension, Behavior Control, was added to the inCLASS for a total of 10 dimensions. Behavior Control was positively skewed across both studies. Following previous trends, the conflict dimensions generally showed a negative skew. The other positive inCLASS dimensions had lower mean scores and distributions, perhaps due to the fact that these data were drawn from more at-risk, ethnically and demographically diverse samples of children and classrooms. For most dimensions, however, the range of scores does span across the scale. Because data for the Field and Hands on Science Studies were collected using the same observation procedures, it is likely that differences in mean scores are due to differences between the samples.

Table 2

Descriptive Statistics for the inCLASS

Domain	Dimension	Time	Pilot Study			Field Study			Hands on Science Study					
			M	SD	Range		M	SD	Range		M	SD	Range	
					Min.	Max.			Min.	Max.			Min.	Max.
Teacher Interactions		Fall	4.39	.58	3.11	6.28	2.18	.78	1.00	5.50	2.62	.81	1.00	5.38
		Spring	4.19	.45	3.17	5.22	2.27	.45	1.00	4.00	2.47	.70	1.13	5.00
		Fall Yr 2					2.10	.57	1.17	4.13				
	Positive Engagement	Fall	3.61	.86	1.63	6.33	2.54	.98	1.00	6.00	3.05	.94	1.00	5.75
		Spring	3.22	.69	1.50	4.67	2.49	.68	1.00	5.40	2.93	.85	1.25	5.75
		Fall Yr 2					2.68	.74	1.25	5.25				
	Teacher Communication	Fall	2.73	1.05	1.00	5.50	1.82	.71	1.00	5.00	2.19	.84	1.00	5.00
		Spring	2.47	.76	1.00	4.33	1.75	.48	1.00	3.67	2.88	.97	1.00	5.50
		Fall Yr 2					1.53	.55	1.00	3.50				
Teacher Conflict	Fall	1.19	.28	1.00	2.33	1.17	.34	1.00	3.43	1.41	.63	1.00	4.25	
	Spring	1.11	.21	1.00	2.33	1.16	.32	1.00	2.67	1.30	.48	1.00	5.00	
	Fall Yr 2					1.10	.26	1.00	3.00					
Peer Interactions		Fall	4.16	.70	2.71	5.96	2.48	.75	1.00	4.83	3.00	.85	1.17	5.50
		Spring	4.13	.70	2.79	6.00	2.46	.53	1.22	4.42	2.97	.77	1.25	5.58
		Fall Yr 2					2.31	.77	1.08	5.08				
	Peer Sociability	Fall	4.16	.89	1.67	6.17	3.28	.71	1.00	5.57	3.66	.90	1.25	6.50
		Spring	3.96	.89	2.30	6.67	3.71	.75	1.50	5.75	3.66	.84	1.50	6.25
		Fall Yr 2					3.22	.92	1.25	5.75				
	Peer Communication	Fall	3.06	1.16	1.00	6.00	2.08	.83	1.00	4.75	2.88	.97	1.00	5.50
		Spring	3.06	1.11	1.00	6.00	2.06	.62	1.00	4.50	2.80	.90	1.00	6.50
		Fall Yr 2					1.96	.83	1.00	5.25				
Peer Assertiveness	Fall	2.75	1.02	1.00	5.88	2.07	.80	1.00	4.50	2.46	.93	1.00	5.50	
	Spring	2.72	.99	1.00	5.83	1.62	.56	1.00	3.67	2.44	.85	1.00	6.00	
	Fall Yr 2					1.76	.78	1.00	4.75					
Peer Conflict	Fall	1.36	.39	1.00	3.00	1.29	.45	1.00	4.67	1.79	.85	1.00	5.75	
	Spring	1.21	.30	1.00	2.50	1.21	.36	1.00	3.67	1.55	.60	1.00	4.50	
	Fall Yr 2					1.20	.38	1.00	4.50					

Table 2 cont.

Descriptive Statistics for the inCLASS

Domain	Dimension	Time	Pilot Study			Field Study			Hands on Science Study						
			<i>M</i>	<i>SD</i>	Range		<i>M</i>	<i>SD</i>	Range		<i>M</i>	<i>SD</i>	Range		
					Min.	Max.			Min.	Max.			Min.	Max.	
Task Orientation		Fall	4.61	.68	2.08	6.13	3.66	.78	1.50	5.83	3.99	.78	2.17	6.38	
		Spring	4.56	.69	2.55	6.17	3.31	.60	1.67	5.33	4.11	.84	1.25	6.38	
		Fall Yr 2					3.81	.64	2.13	5.75					
		Engagement	Fall	5.10	.73	2.50	6.54	4.53	.86	2.33	6.33	4.66	.84	2.50	7.00
			Spring	5.04	.66	3.10	6.33	4.31	.66	2.33	6.00	4.76	.87	1.25	6.75
			Fall Yr 2					4.94	.73	2.50	7.00				
		Self-Reliance	Fall	4.12	.79	1.67	5.83	2.79	.98	1.00	5.67	3.32	1.08	1.00	6.25
			Spring	4.08	.82	1.80	6.00	2.31	.74	1.00	4.67	3.47	1.11	1.00	6.50
			Fall Yr 2					2.69	.81	2.50	7.00				
		Behavior Control	Fall					5.46	.83	2.33	7.00	6.16	1.01	2.00	7.00
			Spring					5.56	.69	2.75	7.00	6.33	.91	3.75	7.00
			Fall Yr 2												

Note: The Teacher and Peer Interactions domain scores are calculated without the Conflict dimensions. The Task Orientation domain score is calculated without the Behavior Control dimension.

How do the dimensions of the inCLASS relate to one another?

The inCLASS was developed using a theoretical framework suggesting three major domains of children's competent behavior in classrooms: Teacher Interactions, Peer Interactions, and Task Orientation (see Table 1). To test the degree to which data from actual classrooms matched this theoretical framework, we began by conducting a confirmatory factor analysis on data from the inCLASS pilot study. This analysis is presented in Table 3 (also see Downer et al., 2010) and resulted in three positive-toned factors as conceptualized, with an additional negative-toned factor comprised of the two conflict dimensions.

This established factor structure was then applied to the Field Study and the Hands on Science Study. Results are presented in Table 4 and 5, respectively. The main factors found in the pilot study (teacher, peer, and task interactions) were confirmed in these studies. The additional inCLASS dimension of Behavior Control, included for the first time in these studies, factored with the two conflict dimensions to form a Conflict Interactions domain.

As presented in Tables 3, 4, and 5, the loadings for each factor were in the moderate to high range and each domain had adequate internal consistency. The only exception is the Hands on Science Study where the task interactions domain had slightly lower internal consistency. It should be noted that throughout these data the task factor had an eigenvalue under one and more cross-loadings than the other established factors. This may suggest that the task dimensions are not as cohesive as expected; future work is needed to further establish this domain.

Table 3

Pilot Study: Fall inCLASS Confirmatory Factor Analyses

Items	Teacher Interactions ($\alpha = 0.80$)	Peer Interactions ($\alpha = 0.92$)	Task Orientation ($\alpha = 0.72$)	Conflict Interactions ($\alpha = 0.71$)
Positive Engagement with Teacher	.78	-.15	.15	-.05
Teacher Communication	.90	.16	.25	.19
Peer Sociability	-.08	.90	.44	.12
Peer Assertiveness	.05	.88	.50	.33
Peer Communication	.00	.94	.40	.29
Engagement within Tasks	.13	.24	.68	-.25
Self-Reliance	.23	.53	.88	.05
Teacher Conflict	.17	.05	-.29	.82
Peer Conflict	-.04	.31	-.03	.75
Eigenvalue	1.85	3.35	.78	1.74
% of variance	20.53	37.18	8.69	19.31

Table 4

Field Study: Fall inCLASS Confirmatory Factor Analyses

Items	Teacher Interactions ($\alpha = 0.79$)	Peer Interactions ($\alpha = 0.87$)	Task Orientation ($\alpha = 0.61$)	Conflict Interactions ($\alpha = 0.71$)
Positive Engagement with Teacher	.89	.26	.13	.00
Teacher Communication	.79	.30	.30	-.04
Peer Sociability	.31	.83	.35	.04
Peer Communication	.18	.78	.51	-.11
Peer Assertiveness	.30	.92	.48	-.12
Engagement within Tasks	.33	.34	.69	.26
Self-Reliance	.12	.42	.65	.02
Teacher Conflict	.02	-.01	-.13	-.64
Peer Conflict	-.06	.04	-.05	-.65
Behavior Control	-.10	-.19	.22	.93
Eigenvalue	1.41	3.27	.91	2.20
% of variance	14.08	32.73	9.13	22.02

Note: Behavior Control dimension is reversed coded to calculate reliability.

Table 5

Hands on Science Study: Fall inCLASS Confirmatory Factor Analyses

Items	Teacher Interactions ($\alpha = 0.79$)	Peer Interactions ($\alpha = 0.90$)	Task Orientation ($\alpha = 0.46$)	Conflict Interactions ($\alpha = 0.82$)
Positive Engagement with Teacher	.83	.08	.22	-.13
Teacher Communication	.72	.19	.16	.00
Peer Sociability	.15	.80	.32	-.07
Peer Communication	.08	.84	.34	.03
Peer Assertiveness	.18	.86	.35	.04
Engagement within Tasks	.25	.35	.78	-.36
Self-Reliance	.29	.51	.53	-.14
Teacher Conflict	.02	-.09	-.35	.60
Peer Conflict	-.11	.05	-.15	.71
Behavior Control	.06	.07	.55	-.77
Eigenvalue	1.52	3.20	.78	2.05
% of variance	15.22	31.96	7.78	20.50

Note: Behavior Control dimension is reversed coded to calculate reliability.

RELIABILITY

How do people become reliable users of the inCLASS?

All inCLASS observers must attend an intensive training and successfully code video clips before observing live in the field in order to establish inter-rater reliability. At the end of inCLASS training, all observers are tested on five reliability clips; to be deemed reliable, they must score within one point from the mastercode on 80% of the dimensions. All trainees are required to demonstrate reliability before going out into the field. For the fall Pilot Study, the group of coders scored within one of the mastercode for 86% of the dimensions across all five training videos (a range of 74 to 92% across the 9 dimensions). For the fall Field Study, the group of coders scored within one point of the mastercode for 90% of the dimensions across all five training videos (a range of 84 to 96% across the 10 dimensions). For the Hands on Science Study, the group of coders passed the reliability standard across all five training videos through the initial test or with follow-up consultation.

How much consistency is there across users of the inCLASS?

Inter-rater reliability for live observations was calculated across a small portion of all live classroom observations, where two coders independently observed and rated the same children. All studies and time points had 20% of observations double coded except during the spring of the field study (12%).

For the Pilot Study, observers scored within one point of each other 87% of the time in the fall and 90% in the spring. An intraclass correlation was calculated across all dimensions and was within the excellent range according to standards in the field (Cicchetti & Sparrow, 1981). Additionally, intraclass correlations remained relatively consistent for each of the domains and dimensions across fall and spring observation periods (with a 4-month break in between) and in the majority of cases actually improved over time.

For the Field Study, observers scored within one point of each other 95% of the time in the fall, 97% in the spring, and 95% in fall of year two. An intraclass correlations were calculated for each time point and was within the excellent range according to standards in the field (Cicchetti & Sparrow, 1981). Additionally, intraclass correlations remained relatively consistent for each of the domains and dimensions across fall and spring observation periods (with approximately a 4-month break in between) and follow-up year two after the summer break.

For the Hands on Science Study, observers scored within one point of each other 85% of the time in the fall and 90% in the spring. An intraclass correlation was calculated across all dimensions and was within the excellent range according to standards in the field (Cicchetti & Sparrow, 1981). Additionally, intraclass correlations remained relatively consistent for each of the domains and dimensions across fall and spring observation periods (with a 3-month break in between) and in the majority of cases actually improved over time.

Table 6

Inter-Rater Reliability for Live Observations

Domain	Dimension	Time	Pilot Study		Field Study		Hands on Science	
			Within 1	Intraclass Correlation	Within 1	Intraclass Correlation	Within 1	Intraclass Correlation
Teacher Interactions		Fall	.93	.92	.96	.84	.84	.73
		Spring	.95	.85	.99	.90	.91	.76
		Fall Yr 2			.94	.78		
	Positive Engagement	Fall	.93	.86	.92	.79	.84	.69
		Spring	.90	.74	.99	.86	.90	.71
		Fall Yr 2			.96	.77		
	Teacher Communication	Fall	.87	.90	.97	.80	.87	.64
		Spring	.94	.89	.99	.86	.95	.70
		Fall Yr 2			.96	.73		
Peer Interactions		Fall	.88	.84	.96	.85	.81	.77
		Spring	.91	.86	.97	.84	.89	.80
		Fall Yr 2			.93	.82		
	Peer Sociability	Fall	.87	.80	.96	.75	.87	.74
		Spring	.89	.76	.99	.79	.89	.68
		Fall Yr 2			.92	.72		
	Peer Communication	Fall	.87	.84	.98	.86	.82	.74
		Spring	.92	.88	.99	.83	.89	.74
		Fall Yr 2			.96	.84		
	Peer Assertiveness	Fall	.83	.71	.97	.78	.79	.64
		Spring	.82	.76	.96	.67	.90	.68
		Fall Yr 2			.93	.68		
Task Orientation		Fall	.77	.63	.91	.73	.77	.63
		Spring	.79	.57	.93	.56	.84	.63
		Fall Yr 2			.90	.69		
	Engagement	Fall	.83	.56	.92	.69	.88	.68
		Spring	.81	.62	.93	.65	.90	.65
		Fall Yr 2			.92	.65		
	Self-Reliance	Fall	.71	.56	.88	.60	.75	.56
		Spring	.77	.59	.93	.37	.78	.54
		Fall Yr 2			.90	.56		
Conflict Interactions		Fall	.99	.55	.99	.78	.93	.76
		Spring	.99	.68	.99	.82	.97	.77
		Fall Yr 2			.99	.68		
	Teacher Conflict	Fall	.99	.44	.99	.69	.96	.73
		Spring	.99	.68	.99	.81	.97	.66
		Fall Yr 2			.99	.63		
	Peer Conflict	Fall	.95	.58	.99	.77	.92	.71
		Spring	.99	.71	.99	.75	.92	.65
		Fall Yr 2			.99	.73		
	Behavior Control	Fall			.95	.63	.84	.66
		Spring			.98	.71	.87	.66
		Fall Yr 2			.94	.54		

Note: The Behavior Control dimension is not included in the Conflict Domain averages.

How stable are scores on the inCLASS?

In this section, test-retest reliability, stability across observation cycles, and stability across times of data collection are presented. Users of the inCLASS may be interested in minimizing the number of observation cycles required to assess each child, due to limited time and resources available for assessment. However, preschool-aged children can be highly variable in their behavior; thus, it is important to maintain multiple observation cycles and time points of data collection.

Test-Retest Reliability. During the pilot study, two visits (or time points) of observation were collected across a two-week period in both the fall and spring. This design allowed for basic test-retest analyses to be completed, and presented in Table 7. In the Teacher and Peer Interaction domains, there were significant moderate correlations; whereas for the Task Orientation and Conflict domains, there were less consistent relations across two-week periods during the fall and spring (i.e., smaller magnitude but still significant).

Table 7

Pilot Study: Fall inCLASS Test-Retest Reliability across Two Weeks

	Fall	Spring
Teacher Interactions	.45***	.35***
Positive Engagement with Teacher	.39***	.23***
Teacher Communication	.55***	.43***
Peer Interactions	.44***	.59***
Peer Sociability	.40***	.52***
Peer Communication	.50***	.56***
Peer Assertiveness	.33***	.49***
Task Orientation	.24**	.05
Engagement within Tasks	.26***	.09
Self-Reliance	.17*	.02
Conflict Interactions	.29***	.17[†]
Teacher Conflict	.18*	.05
Peer Conflict	.22***	.13

[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

Stability across Cycles. Tables 8, 9, and 10 show the correlations between the scores a child received in each cycle of observation, and their total morning average across the other three cycles for each of the inCLASS studies. Note that correlations between the first cycle and the rest of the morning's average are of only medium size, but by the second cycle and on, all correlations are generally high. It should be noted that in some of the inCLASS studies, children were only observed for three cycles due to time constraints.

Table 8

Pilot Study: Fall inCLASS Correlations across Observation Cycles

	Cycle 1 with 2,3,4	Cycle 2 with 1,3,4	Cycle 3 with 1,2,4	Cycle 4 with 1,2,3
Teacher Interactions	.39***	.48***	.26**	.34**
Positive Engagement with Teacher	.34***	.39***	.34***	.28***
Teacher Communication	.38***	.48***	.23**	.30*
Peer Interactions	.40***	.41***	.36***	.27*
Peer Sociability	.29***	.29***	.26**	.19
Peer Communication	.38***	.36***	.30***	.21 [†]
Peer Assertiveness	.40***	.36***	.32***	.23 [†]
Task Orientation	.15[†]	.50***	.45***	.40***
Engagement within Tasks	.18*	.25**	.19*	.40**
Self-Reliance	.26***	.42***	.27***	.30*
Conflict Interactions	.17*	.43***	.37***	.39***
Teacher Conflict	.14 [†]	.37***	.38***	.19
Peer Conflict	.07	.28***	.15 [†]	.42***

[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

Table 9

Field Study: Fall inCLASS Correlations across Observation Cycles

	Cycle 1 with 2,3,4	Cycle 2 with 1,3,4	Cycle 3 with 1,2,4	Cycle 4 with 1,2,3
Teacher Interactions	.43***	.40***	.37***	.25**
Positive Engagement with Teacher	.40***	.38***	.35***	.22*
Teacher Communication	.38***	.35***	.31***	.30***
Peer Interactions	.27***	.30***	.34***	.38***
Peer Sociability	.28***	.29***	.29***	.34***
Peer Communication	.20***	.26***	.30***	.34***
Peer Assertiveness	.27***	.29***	.36***	.33***
Task Orientation	.38***	.38***	.38***	.44***
Engagement within Tasks	.33***	.26***	.29***	.37***
Self-Reliance	.43***	.45***	.47***	.44***
Conflict Interactions	.45***	.42***	.33***	.47***
Teacher Conflict	.30***	.20***	.16**	.34***
Peer Conflict	.21***	.15**	.22***	.16 [†]
Behavior Control	.45***	.43***	.32***	.46***

Note: Behavior Control dimension is reversed coded to calculate Conflict Interactions domain score.

[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

Table 10

Hands on Science Study: Fall inCLASS Correlations across Observation Cycles

	Cycle 1 with 2,3,4	Cycle 2 with 1,3,4	Cycle 3 with 1,2,4	Cycle 4 with 1,2,3
Teacher Interactions	.49***	.38***	.34***	.37***
Positive Engagement with Teacher	.43***	.32***	.35***	.34***
Teacher Communication	.46***	.39***	.30***	.34***
Peer Interactions	.24***	.35***	.34***	.36***
Peer Sociability	.25***	.26***	.30***	.23***
Peer Communication	.18**	.33***	.32***	.31***
Peer Assertiveness	.25***	.36***	.34***	.40***
Task Orientation	.21***	.29***	.32***	.25***
Engagement within Tasks	.22***	.32***	.31***	.24***
Self-Reliance	.33***	.39***	.41***	.37***
Conflict Interactions	.53***	.62***	.60***	.50***
Teacher Conflict	.46***	.54***	.45***	.36***
Peer Conflict	.42***	.49***	.47***	.39***
Behavior Control	.44***	.51***	.56***	.44***

Note: Behavior Control dimension is reversed coded to calculate Conflict Interactions domain score.

** $p \leq .01$, *** $p \leq .001$.

Stability across Times of Data Collection. Tables 11, 12, and 13 shows the correlations between the average inCLASS domain score across the different times of data collected for each of the inCLASS studies. There are some correlations within time points, as discussed by Downer et al. (2010), but many of the inCLASS domains were also correlated across times of data collections. For example, the Conflict Interactions domain is consistently correlated across time points in all three studies, suggesting that the inCLASS Conflict domain is capturing a characteristic that is stable in children over time. Other domain scores appear to be stable across fall and spring time points, albeit sometimes with lower magnitudes. In the Field Study, which followed children into another school year, correlations were generally higher between fall and spring ratings within the first year of data collection and lower when looking at correlations between first year and second year ratings. This may be due to many children transitioning from preschool into kindergarten classrooms and therefore experiencing a new teacher, a new peer group, and new behavioral expectations.

Table 11

Pilot Study: inCLASS Correlations across Times of Data Collection

	1	2	3	4	5	6	7
1. Fall Teacher Interactions							
2. Fall Peer Interactions	.07						
3. Fall Task Interactions	.33***	.42***					
4. Fall Conflict Interactions	-.06	.11	-.09				
5. Spring Teacher Interactions	.48***	.13	.17*	.14 [†]			
6. Spring Peer Interactions	.05	.57***	.32***	.21**	.19*		
7. Spring Task Interactions	.23**	.34***	.30***	.12	.21**	.55***	
8. Spring Conflict Interactions	-.01	.19*	-.03	.54***	-.14 [†]	.13	.02

* $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

Table 12

Field Study: inCLASS Correlations across Times of Data Collection

	1	2	3	4	5	6	7	8	9	10	11
1. Fall Teacher Interactions											
2. Fall Peer Interactions	.30***										
3. Fall Task Interactions	.25***	.46***									
4. Fall Conflict Interactions	.07	.11 [†]	-.16**								
5. Spring Teacher Interactions	.26***	-.01	.04	.05							
6. Spring Peer Interactions	.05	.24***	.19***	.02	.13*						
7. Spring Task Interactions	-.05	.10 [†]	.22***	-.01	.18**	.43***					
8. Spring Conflict Interactions	.05	-.00	-.12*	.34***	.11*	.03	-.20***				
9. Fall Yr 2 Teacher Interactions	.10	.04	.15*	.02	.23**	.02	-.04	.22**			
10. Fall Yr 2 Peer Interactions	-.01	.16*	.14 [†]	.00	-.04	.18*	-.08	.12	.28***		
11. Fall Yr 2 Task Interactions	-.06	.11	.09	-.06	-.10	.05	-.02	-.04	.13 [†]	.58***	
12. Fall Yr 2 Conflict Interactions	.03	.04	-.07	.25***	.04	.06	-.18*	.44***	.21**	.33***	-.07

Note: Behavior Control dimension is reversed coded to calculate Conflict Interactions domain score. $N = 188 - 341$.[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

Table 13

Hands on Science Study: inCLASS Correlations across Times of Data Collection

	1	2	3	4	5	6	7
1. Fall Teacher Interactions							
2. Fall Peer Interactions	.32***						
3. Fall Task Interactions	.39***	.55***					
4. Fall Conflict Interactions	.00	-.03	-.29***				
5. Spring Teacher Interactions	.24***	.16*	.15*	.19**			
6. Spring Peer Interactions	.18**	.26***	.17*	.00	.10		
7. Spring Task Interactions	.09	.13*	.24***	-.10	.24***	.40***	
8. Spring Conflict Interactions	.01	.00	-.09	.33***	.10	-.01	-.33***

Note: Behavior Control dimension is reversed coded to calculate Conflict Interactions domain score.

* $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

VALIDITY

Does the inCLASS measure constructs of importance in children’s classroom behavior (face validity)?

The inCLASS was developed based on an extensive literature review of the important cognitive and socioemotional skills developing during the preschool period, which predict children’s later social and academic performance in school. The choice of dimensions was additionally informed by a review of constructs assessed in other observational, teacher report, and direct assessment instruments currently used in child care and research. Finally, the operational definitions of the dimensions were specified through extensive piloting and revision after sampling with more diverse populations of children. Consultation with early childhood practitioners, as well as researchers with expertise in child development and school readiness, confirm that the inCLASS measures aspects of children’s classroom behavior that impact their school performance and socioemotional competency, suggesting considerable face validity.

How does the inCLASS relate to other measures of children’s behaviors (criterion and concurrent validity)?

To establish criterion and concurrent validity, the inCLASS observations were compared to teacher ratings of similar behaviors. Specifically, bivariate correlations were conducted between inCLASS teacher, peer, task, and conflict domain scores and teacher ratings from several established measures collected during the same time point of data collection. Teacher ratings included scales from the Student-Teacher Relationship Scale (STRS; Pianta, 2001), the Academic Rating Scale (ARS; Rock & Pollack, 2002), the California Preschool Social Competency Scales (CPSCS; Levine et al., 1970), the Teacher-Child Rating Scale (TCRS; Hightower et al., 1986), the Emotion Regulation Checklist (ERC; Shields & Cicchetti, 1997), the Devereux Early Childhood Assessment (DECA; LeBuffe & Naglieri, 1999), and the Preschool Learning Behaviors Scale (PLBS; McDermott et al., 2002). Refer to Tables 14, 15, and 16 for separate correlations between observations and teacher ratings for each of the inCLASS studies.

Across studies there were associations in the expected direction with small to moderate magnitudes. For example, observations of children’s interactions with teachers were positively correlated with teacher ratings of closeness with that child. As expected, children who were observed as having higher quality interactions with their peers had higher teacher ratings of behaviors that likely facilitate peer interactions, including assertiveness, social communication, and language and literacy skills. Children who were observed as having higher quality interactions with classroom tasks and activities were also rated more highly by their teachers on related skills, including task orientation and language and literacy skills. Finally, children with higher observed conflict were also rated by the teacher as showing more problem behaviors in the classroom.

Table 14

Pilot Study: Fall inCLASS Bivariate Pearson Correlations with Teacher Ratings

Items	Teacher Interactions	Peer Interactions	Task Orientation	Conflict Interactions
Closeness (STRS)	.28***	.08	.23**	-.08
Conflict (STRS)	.04	.17*	-.03	.54***
Language and Literacy (ARS)	-.03	.27***	.33***	-.08
Social Communication (CPSCS)	.05	.18*	.35***	-.25***
Frustration Tolerance (TCRS)	-.10	-.24**	-.02	-.45***
Assertiveness (TCRS)	.29***	.38***	.35***	.11
Task Orientation (TCRS)	.03	.02	.24**	-.24**
Social Skills (TCRS)	.13	.21**	.28***	-.21**
Problem Behaviors (TCRS)	.02	.05	-.24**	.39***
Emotion Regulation (ERC)	.18*	.23**	.32***	-.05

Note: Conflict Interactions domain score does not include Behavior Control. $N = 135-164$.

* $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

Table 15

Field Study: Fall inCLASS Bivariate Pearson Correlations with Teacher Ratings

Items	Teacher Interactions	Peer Interactions	Task Orientation	Conflict Interactions
Closeness (STRS)	.16**	.11	.17**	-.02
Conflict (STRS)	-.05	.02	-.12*	.25***
Language and Literacy (ARS)	.20***	.13*	.25*	-.06
Social Communication (CPSCS)	.07	.15*	.29***	-.16**
Peer Assertiveness (TCRS)	.18**	.20***	.31***	-.09
Task Orientation (TCRS)	.05	.05	.27***	-.20***
Social Skills (TCRS)	.01	.18**	.28***	-.17**
Problem Behaviors (TCRS)	-.00	-.06	-.23***	.30***
Emotion Regulation (ERC)	.16**	.11 [†]	.24***	-.13*

Note: Behavior Control dimension is reversed coded to calculate Conflict Interactions domain score. $N = 303 - 341$.

[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

Table 16

Hands on Science Study: Fall inCLASS Bivariate Pearson Correlations with Teacher Ratings

Items	Teacher Interactions	Peer Interactions	Task Orientation	Conflict Interactions
Initiative (DECA)	.23***	.13*	.21**	-.15*
Self-Control (DECA)	.04	.10	.20**	-.31***
Attachment (DECA)	.18**	.18**	.18**	-.13 [†]
Behavioral Concerns (DECA)	-.07	-.07	-.19**	.40***
Competence Motivation (PLBS)	.11	.09	.17**	-.16*
Attention/Persistence (PLBS)	.17*	.11	.22***	-.28***
Attitude Toward Learning (PLBS)	.08	.13*	.15*	-.27***

Note: Behavior Control dimension is reversed coded to calculate Conflict Interactions domain score. $N = 242$.

[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

How does the inCLASS relate to later teacher reports of similar behavior (predictive validity)?

To establish predictive validity, the inCLASS fall observations were compared to spring child outcomes from teacher ratings and direct assessments. Due to the nested nature of these data, hierarchical linear models were conducted using fall observations to predict spring child outcomes. Teacher ratings included scales from the Student-Teacher Relationship Scale (STRS; Pianta, 2001), the Academic Rating Scale (ARS; Rock & Pollack, 2002), the California Preschool Social Competency Scales (CPSCS; Levine et al., 1970), the Teacher-Child Rating Scale (TCRS; Hightower et al., 1986), the Emotion Regulation Checklist (ERC; Shields & Cicchetti, 1997), the Devereux Early Childhood Assessment (DECA; LeBuffe & Naglieri, 1999), the Preschool Learning Behaviors Scale (PLBS; McDermott et al., 2002), and the Learning Express (McDermott et al., 2009). Associations between inCLASS dimensions and children's outcomes were assessed after adjusting for a variety of covariates, including child gender, age and maternal education, through conditional hierarchical linear models. Refer to Tables 17, 18, 19, and 20 for the predictors and outcomes from each inCLASS study.

Observations of children's classroom interactions were related to teacher ratings of similar constructs. For example, observed conflict consistently predicted teacher-rated conflict at the following time point. Additionally, task observations predicted teacher-rated academic skills including language and literacy skills and task orientation. Similarly, observations of peer interactions predicted teacher-rated peer assertiveness.

Some cross-domain relationships were also observed. For example, task observations predicted teacher-rated closeness, social communication, social skills, assertiveness, possibly because the capacity for engagement that children demonstrate in relation to tasks can affect their ability to manage interpersonal relationships. Observed conflict was negatively related to teacher-rated social-emotional competencies (e.g. social skills and emotion regulation).

The inCLASS observations also predicted children's academic outcomes; for example, one study suggested that more positive teacher interactions were associated with increased alphabet knowledge. Additionally, more conflict was associated with lower listening comprehension. More positive peer interactions were related to greater math knowledge and listening comprehension.

Table 17

Pilot Study: Hierarchical Linear Models (HLM) of Fall inCLASS Observations Predicting Spring Teacher Ratings

	Closeness		Conflict		Language & Literacy		Social Communication		Assertiveness		Task Orientation		Social Skills		Emotion Regulation	
	B	SE	B	SE	B	SE	B	SE	B	SE	B	SE	B	SE	B	SE
Intercept	4.51***	.07	1.74***	.10	3.36***	.10	3.02***	.06	3.49***	.09	3.82***	.11	4.03***	.11	3.35***	.07
Teacher Interactions	.12	.09	.19	.12	-.25 [†]	.13	.07	.08	.34**	.13	-.15	.14	-.06	.14	.08	.08
Peer Interactions	.04	.08	.18	.12	.12	.13	-.10	.08	.21 [†]	.12	-.08	.13	.08	.13	-.06	.07
Task Orientation	.13	.08	-.06	.11	.42**	.12	.19*	.07	.38**	.12	.37**	.13	.31*	.13	.17*	.07
Conflict Interactions	.32 [†]	.08	.82***	.22	.17	.24	-.06	.15	1.05***	.23	-.35	.25	-.04	.25	.19	.14

Note: The predictors are centered around the aggregated mean, and control for child age, child gender, and maternal education. Conflict Interactions domain score does not include Behavior Control.

[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

Table 18

Field Study: Hierarchical Linear Models (HLM) of inCLASS Observations Predicting Later Teacher Ratings

		Closeness		Conflict		Language & Literacy		Social Communication		Assertiveness		Task Orientation		Social Skills		Emotion Regulation	
		B	SE	B	SE	B	SE	B	SE	B	SE	B	SE	B	SE	B	SE
Fall to Spring	Intercept	4.45***	.06	1.52***	.07	3.39***	.09	3.12***	.05	3.74***	.08	3.89***	.06	4.04***	.07	3.26***	.05
	Teacher Interactions	.07	.05	.01	.06	.11	.07	.04	.04	.12	.08	.06	.05	-.04	.07	.02	.04
	Peer Interactions	.00	.06	.00	.07	-.01	.07	-.06	.05	.04	.09	-.03	.06	.04	.08	-.01	.05
	Task Orientation	.16**	.06	-.09	.07	.07	.08	.19***	.05	.29***	.08	.09	.06	.24**	.08	.18***	.04
	Conflict Interactions	.12	.08	.38***	.09	.07	.10	.03	.07	.12	.11	.10	.09	-.18	.11	.02	.06
Spring to Fall Yr 2	Intercept	4.41***	.07	1.39***	.07	3.42***	.10	3.10***	.05	3.59***	.10	3.92***	.09	4.09	.08	3.21***	.05
	Teacher Interactions	-.05	.09	.07	.10	.06	.14	-.01	.07	.06	.15	.06	.13	-.03	.12	.01	.07
	Peer Interactions	.17 [†]	.10	-.16	.10	.12	.14	.11	.08	.26 [†]	.15	.21	.14	.15	.12	.16*	.07
	Task Orientation	-.03	.09	-.04	.09	.02	.13	.02	.07	.11	.14	.08	.13	.05	.11	-.07	.07
	Conflict Interactions	-.16	.13	.58***	.14	-.15	.20	-.25*	.11	-.09	.22	-.45*	.19	-.60**	.17	-.22*	.10

Note: The predictors are centered around the aggregated mean, and control for child age, child gender, and maternal education. Behavior Control dimension is reversed coded to calculate Conflict Interactions domain score.

[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

Table 19

Field Study: Hierarchical Linear Models (HLM) of inCLASS Observations Predicting Later Direct Assessments

		PPVT (Picture Vocabulary)		WJ (Picture Vocabulary)		Alphabet Knowledge		TVIP (Spanish PPVT)		Bateria (Spanish WJ)	
		B	SE	B	SE	B	SE	B	SE	B	SE
Fall to Spring	Intercept	47.82***	2.11	454.54***	1.97	12.60***	.94	18.05***	1.56	413.54***	3.20
	Teacher Interactions	.28	2.06	-1.55	1.94	1.15	.91	1.38	1.33	1.00	2.76
	Peer Interactions	2.40	2.35	1.86	2.25	.20	1.04	-1.24	1.82	-1.98	3.67
	Task Orientation	2.77	2.34	1.88	2.16	.60	1.04	-.14	1.43	-5.05 [†]	2.98
	Conflict Interactions	-2.58	3.15	.05	3.03	-.10	1.40	-.90	2.68	-4.00	5.38
Spring to Fall Yr 2	Intercept	60.56***	2.64	462.99***	2.09	19.11***	.93	21.49***	2.26	419.17	4.52
	Teacher Interactions	5.49	5.09	8.41*	4.24	1.36	1.87	3.88	4.42	6.16	8.85
	Peer Interactions	-1.71	4.52	-6.31 [†]	3.71	-.35	1.64	1.14	3.71	-1.43	7.37
	Task Orientation	-1.71	2.85	1.49	2.34	-.86	1.04	2.94	2.14	3.39	4.27
	Conflict Interactions	29.70*	12.09	21.30	9.95	4.10	4.40	-9.80	9.54	-8.81	19.04

Note: The predictors are centered around the aggregated mean, and control for child age, child gender, and maternal education. Behavior Control dimension is reversed coded to calculate Conflict Interactions domain score. Only children who spoke any Spanish were assessed in the Spanish versions.

[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

Table 20

Hands on Science Study: Hierarchical Linear Models (HLM) of Fall inCLASS Observations Predicting Spring Learning Express Direct Assessment

	Vocabulary		Math		Listening Comp.		Alphabet Knowledge	
	B	SE	B	SE	B	SE	B	SE
Intercept	207.90***	4.08	212.30***	3.88	218.22***	3.61	216.81***	4.02
Teacher Interactions	5.29	3.48	2.33	3.56	2.20	3.28	12.58**	3.82
Peer Interactions	3.53	3.61	7.95*	3.74	6.93*	3.43	-0.24	4.05
Task Orientation	-2.28	4.44	1.70	4.58	-2.33	4.20	-3.28	4.92
Conflict Interactions	-8.49*	3.77	-4.98	3.90	-7.84*	3.58	-1.47	4.22

Note: The predictors are centered around the aggregated mean, and control for child age, child gender, and maternal education. Behavior Control dimension is reversed coded to calculate Conflict Interactions domain score.

* $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

Does the inCLASS measure constructs that show important differences in children (construct validity)?

The inCLASS dimensions were derived from an extensive literature review of the social, emotional, and academic behaviors which develop during the preschool years. It is therefore important to look at relations between the inCLASS dimensions and children's age, socioeconomic status, and gender. To establish construct validity, characteristics of the child were compared to the inCLASS fall observations. Due to the nested nature of these data, hierarchical linear models were conducted using family demographic variables (i.e., child age and gender, and reported maternal education) and fall inCLASS observations. Refer to Tables 21, 22, and 23 for the child characteristic predicting inCLASS observations from each inCLASS study.

Across studies, results generally indicated that inCLASS observations are somewhat sensitive to age differences, providing initial evidence of construct validity such that older children scored higher. There were some gender differences evident, specifically in the Conflict Interactions domain that follows previous research trends of boys being more conflictual than girls. The indicator of socioeconomic differences (i.e., maternal education) did not consistently predict to the inCLASS observations, but occasional significant differences were in expected directions. In the Field Study, children whose parents reported speaking English at home were rated as engaging in more positive peer and task interactions during the fall preschool observation, but no effect of home language was found at later time points.

Table 21

Pilot Study: Hierarchical Linear Models Predicting the Fall inCLASS Dimensions

	Teacher Interactions		Peer Interactions		Task Interactions		Conflict Interactions	
	B	SE	B	SE	B	SE	B	SE
Intercept	4.49***	.07	4.18***	.07	4.66***	.08	1.24***	.04
Age	.00	.01	.04***	.01	.02*	.01	-.00	.00
Gender	-.20*	.09	-.02	.10	-.15	.11	.10 [†]	.05
Maternal Education	-.00	.02	.00	.02	.01	.02	-.04***	.01

Note: The predictors are centered around the aggregated mean; for gender, girl = 0.

[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

Table 22

Field Study: Hierarchical Linear Models Predicting the inCLASS Dimensions

		Teacher Interactions		Peer Interactions		Task Interactions		Conflict Interactions	
		B	SE	B	SE	B	SE	B	SE
Predicting Fall inCLASS	Intercept	2.09***	.09	2.21***	.08	3.41***	.09	1.64***	.05
	Age	.01 [†]	.01	.02***	.01	.02***	.01	-.00	.00
	Gender	.02	.06	.23***	.07	.05	.07	.08	.05
	Maternal Education	-.02	.01	.03*	.01	-.00	.01	-.00	.01
	Home Language	.15 [†]	.08	.25**	.08	.35***	.09	-.02	.06
Predicting Spring inCLASS	Intercept	2.06***	.06	2.37***	.06	3.33***	.07	1.55***	.04
	Age	.00	.00	.01*	.00	.01	.01	-.01	.00
	Gender	-.01	.05	.04	.06	-.03	.06	.11**	.04
	Maternal Education	-.02	.01	.01	.01	.01	.01	-.01	.01
	Home Language	.11 [†]	.06	.12 [†]	.07	.01	.07	.00	.04
Predicting Fall Yr 2 inCLASS	Intercept	2.07***	.08	2.22***	.11	3.74***	.10	1.48***	.05
	Age	-.00	.01	-.01	.01	-.01	.01	-.00	.00
	Gender	.17*	.08	.14	.11	.10	.09	.13*	.05
	Maternal Education	-.00	.01	.04*	.02	.02	.02	.00	.01
	Home Language	-.05	.09	.02	.12	.03	.11	-.12 [†]	.06

Note: The predictors are centered around the aggregated mean; for gender, girl = 0; for home language, English = 1, Spanish = 0.

[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

Table 23

Hands on Science Study: Hierarchical Linear Models Predicting the inCLASS Dimensions

		Teacher Interactions		Peer Interactions		Task Interactions		Conflict Interactions	
		B	SE	B	SE	B	SE	B	SE
Predicting Fall inCLASS	Intercept	2.64***	.08	3.03***	.08	4.09***	.08	1.88***	.07
	Age	.02*	.01	.03**	.01	.03***	.01	-.00	.01
	Gender	-.04	.10	-.06	.11	-.22*	.09	.28**	.09
Predicting Spring inCLASS	Intercept	2.55***	.08	3.02***	.08	4.23***	.10	1.75***	.05
	Age	.01*	.01	.03***	.01	.04***	.01	-.00	.01
	Gender	-.16 [†]	.06	-.08	.09	-.21*	.09	.19**	.07

Note: The predictors are centered around the aggregated mean; for gender, girl = 0.

[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

SUMMARY

The inCLASS provides important descriptive information about children's behaviors and competencies in early childhood classrooms. The inCLASS is associated with and has predictive value for academic and social outcomes for children. As such, the inCLASS may be a useful tool for researchers and teachers seeking a standardized measure of children's behavior in the classroom environment.