Article

The Social Context of Urban Classrooms: Measuring Student Psychological Climate

Journal of Early Adolescence
2015, Vol. 35(5-6) 759–789

DOI: 10.1177/0272431615570056
jea.sagepub.com
SAGE

Stacy L. Frazier[1], Tara G. Mehta[2], Marc S. Atkins[2], Charles Glisson[3], Philip D. Green[3], Robert D. Gibbons[4], Jong Bae Kim[4], Jason E. Chapman[5], Sonja K. Schoenwald[5], Grace Cua[2], Robert R. Ogle[1], and the LINKS (Linking Neighborhoods, Kids, and Schools) Center

Abstract

Classrooms are unique and complex work settings in which teachers and students both participate in and contribute to classroom processes. This article describes the measurement phase of a study that examined the social ecology of urban classrooms. Informed by the dimensions and items of an established measure of organizational climate, we designed the *Student Climate Survey* (*n* = 53 items) to assess student psychological climate in third through eighth grades. We administered the survey to 621 students at three time points within one school year in 69 classrooms within eight urban schools. A multidimensional item response theory (IRT) analysis based on a full-information item bifactor model revealed 18 items that loaded on a primary factor and demonstrated good criterion and predictive validity.

[1]Florida International University, Miami, USA
[2]University of Illinois at Chicago, USA
[3]The University of Tennessee, Knoxville, USA
[4]The University of Chicago, IL, USA
[5]Medical University of South Carolina, Charleston, USA

Corresponding Author:
Stacy L. Frazier, Ph.D., Center for Children and Families, Department of Psychology, Florida International University, 11200 SW 8th Street, Miami, FL 33199, USA.
Email: slfrazi@fiu.edu

Opportunities for the *Student Climate Survey* to advance our contextual understanding of urban classrooms and inform intervention are discussed.

Keywords
urban classrooms, psychological climate, classroom measurement

Classrooms are unique organizational work units in that both teachers and students influence classroom process, and students are both contributors to and recipients of the educational services provided. There has been substantive effort during the last two decades to improve the methods by which we observe and examine classrooms, responding to data that demonstrate the important contribution of classroom context to children's learning and accumulating evidence that variables at multiple levels (student, teacher, classroom) influence outcomes. Indeed, the tremendous complexity and dynamic nature of classrooms warrant equally complex measurement sufficient to capture the delicate interplay of instruction, relationships, and behaviors that comprise the classroom environment. At the same time, competing priorities and increasing demands on teachers require that measures be brief and unobtrusive, minimizing interference with instructional time. The present study introduces a measure of *student psychological climate*.

What Is Psychological Climate?

There is widespread and long-standing agreement that every classroom has a unique "climate," and efforts to understand and examine its impact on learning are rooted in a century of research. Earliest measures of climate emphasized teachers' verbal behavior (e.g., Anderson & Brewer, 1945, 1946; Withall, 1949, 1951) but over time expanded to include both structural (e.g., rules, organizational, authority) and affective (i.e., interpersonal relationships among teachers and students) features of the social environment (e.g., Trickett & Moos, 1973). Alongside this conceptual evolution, dialogue ensued regarding methods and measurement, and observations that once were considered the gold standard declined in favor of aggregated student reports that prioritized subjective experience over objective interpretations (Trickett & Quinlan, 1979).

Current definitions of climate encompass a multiplicity of factors that include rules and norms, policies and procedures, safety and supports, relationships and interactions, instruction, discipline, and autonomy (Matsumura, Slater, & Crosson, 2008). Positive climates describe classrooms with clear

rules and high achievement expectations that promote caring, respect, cooperation, and emotional safety. Negative climates describe classrooms characterized by inconsistent rules, punitive discipline, and low expectations that minimize collaboration, reduce motivation, and allow disrespect, negativity, and bullying. Measures include a combination of teacher-report, student-report, and independent observations. Climate is ascribed to the level of classroom (e.g., Trickett & Quinlan, 1979) and often to the level of school (e.g., Brand, Felner, Shim, Seitsinger, & Dumas, 2003) reflecting a higher-order dimension that captures an overall appraisal of the educational environment.

However, this view of climate lies in contrast to the construct of psychological climate as defined and measured in the organizational literature, where it represents an *individual-level*, rather than a *setting-level*, characteristic. Specifically, in the organizational literature, psychological climate is an individual-level construct that reflects the extent to which an individual believes he or she is affected positively or negatively by his or her work environment (James & James, 1989). The construct is best illustrated with a simple example. *Temperature* is an objective measure of a room's warmth or coldness with reference to a standard value (e.g., the room's temperature is 78 degrees). *Climate* is a subjective measure of a person's comfort (e.g., I feel chilly), and thus, multiple people in the same room may report different experiences (e.g., warm, cold, or comfortable). Climate as defined this way is assessed by self-report, as the construct itself represents an individual experience (rather than description) of an environment.

Why Is Psychological Climate Important?

If several people report comparable experiences (e.g., everyone feels cold), then their individual data can be aggregated to represent an organizational climate variable (setting-level characteristic). Extending this to a workplace example, when individuals in the same work environment agree on their perceptions of that environment as stressful, their shared perceptions can be aggregated, and the organizational climate can be described as stressful (Jones & James, 1979; Joyce & Slocum, 1984). Positive organizational climates in youth mental health and child welfare systems (e.g., engaged, functional climates) have been associated with lower staff turnover, higher service quality, and better youth outcomes in numerous studies, including both nationwide surveys and randomized controlled trials (e.g., Glisson & Green, 2011; Glisson, Hemmelgarn, Green, & Williams, 2013; Glisson, Schoenwald, et al., 2008; Olin et al., 2014).

Extending this to schools, a rich literature on teacher stress and burnout (Byrne, 1999; Kyriacou, 1987, 2001; Watts & Robertson, 2011) illustrates

one measure and outcome of teacher psychological climate. Stress is the result of negative emotions (e.g., anger, anxiety, frustration) triggered by the workplace (Kyriacou, 2001), and burnout, in turn, is a response to chronic job stress, characterized by exhaustion, cynicism, and lack of felt efficacy (Maslach, Schaufeli, & Leiter, 2001). Teacher stress reflects classroom-level (e.g., limited resources, heterogeneous learners, extensive behavior problems, role overload) and school-level (e.g., disorganization, excessive workload, accountability policies) factors that influence teaching quality and effectiveness via harsher punishment styles, decreased effort, and lower quality of instruction and affect teachers' personal relationships, physical health, work performance, and emotional well-being (Shernoff, Mehta, Atkins, Torf, & Spencer, 2011). Hence, "stressful" may represent a teacher's perception of the psychological impact of their work environment, offering an example of how the construct of psychological climate may be applied to schools.

Just as for teachers, student perceptions of the psychological impact of their educational environment may vary and reflect the extent to which they believe they are affected positively or negatively by going to school. We propose that a student's overall appraisal of how his or her school experience affects him or her psychologically has important implications for outcomes and intervention. Merging educational and organizational literatures, we expect student appraisals to reflect many aspects of their experience including but not limited to their individual instructional level (is the work too hard or too easy), position in the social network (connectedness to other students), relationship to the teacher (ratio of praise to reprimands), and opportunities for accomplishment (performance, grades). Whereas two students in the same classroom may offer similar *descriptions* of their teacher and classmates, their psychological *experiences* may vary as a function of their individual histories, resources, and relationships. Specifically, we are interested in determining the extent to which a student's perceptions regarding the personal impact of his or her educational workplace—is it a healthy, safe, and productive space for me to learn—contributes to a better understanding of linkages between classroom ecology, student engagement, and academic progress.

What Is Unique About Student Psychological Climate?

To our knowledge, there are no measures to assess student psychological climate as defined in the organizational literature (i.e., perceptions of the extent to which a student's daily experience in school affects his or her own mental health and well-being), limiting what we know about its contribution to understanding associations between classroom context and learning outcomes, and

its malleability as a potential lever for change. Widely utilized student-report measures of classroom environment are composed largely of items that describe features of the classroom environment (e.g., Trickett & Quinlan, 1979): teacher practice and personality (e.g., the teacher goes out of his way to help students; this teacher "talks down" to students), student behaviors and interactions (e.g., "Students are expected to follow set rules in doing their work"; "Students enjoy working together on projects in this class"), and classroom expectations (e.g., "Grades are not very important in this class"; "This class is more a social hour than a place to learn something"). They produce a score that represents the rules, norms, organization, and relationships common to the classroom—the way things are done—reflecting what would be defined in much of the organizational literature as "culture" instead of "climate" (Glisson, 2002).

A large body of literature presents several other student-report measures that assess, for example, motivation (e.g., Ginsburg-Block & Fantuzzo, 1998), school attachment (e.g., Libbey, 2004), and school liking or avoidance (Ladd & Price, 1987). While these may be associated with the construct of psychological climate that derives from the organizational literature, psychological climate is theoretically and methodologically distinct, characterized by several specific underlying domains (e.g., role clarity, emotional exhaustion, opportunities for growth) reflecting student perceptions of their school experience on personal mental health and well-being (i.e., "This is a healthy place for me") rather than individual differences on constructs describing work, friendships, or overall satisfaction (i.e., "I like it here").

Current Study

This article introduces the construct of student psychological climate and its unique contribution to the comprehensive measurement of classroom ecology. We offer a developmentally appropriate measure of psychological climate for elementary school students, informed by the items used to measure climate from a well-validated instrument, the Organizational Social Context (OSC) measure (Glisson, Green, & Williams, 2012; Glisson, Landsverk, et al., 2008). After rewriting selected climate items to be appropriate for students in a classroom context, we administered the survey to 621 students (Grades 3 to 8) at three time points within one school year in 69 classrooms within eight urban elementary schools. Items were calibrated using a multidimensional item response theory (IRT) analysis based on a full-information item bifactor model (Gibbons et al., 2007; Gibbons & Hedeker, 1992). Mixed-effects regression models were used to assess (a) criterion validity by examining longitudinal associations among the estimated student climate

scores with student-reported academic motivation and school liking/avoidance, and (b) predictive validity by examining longitudinal associations among student climate scores with teacher-reported classroom functioning (problem behavior, social skills, academic competence), observed student engagement, and curriculum-based measures of reading performance (fluency and comprehension).

Method

Participants

Schools. Eight K-8 schools in a large Midwestern city participated in this study. Schools were located in four low-income communities (including two high-poverty communities), as determined by 87% or more students eligible for subsidized meals. Schools showed substantial variability on student mobility (range = 3.5%-47.6%, $\overline{X}$ = 29.1%, *SD* = 14.9) and student enrollment (range = 405-1,045, $\overline{X}$ = 673.0, *SD* = 225.8). Seven schools served predominantly African American students (range = 88.1%-100%, $\overline{X}$ = 96.96%, *SD* = 4.10), and one school served predominantly Hispanic or Latino students (91.6%). The proportion of students meeting or exceeding national averages on standardized tests in reading ranged from 45.8% to 84.6% ($\overline{X}$ = 60.8%, *SD* = 13.3) and in math ranged from 52.9% to 89.3% ($\overline{X}$ = 68.5%, *SD* = 11.9). The number of students from Grades 3 to 8 ranged from 237 to 617 ($\overline{X}$ = 412.3, *SD* = 133.4).

Classrooms. Classrooms (*n* = 69) were distributed across the eight participating schools, with a range of 4 to 17 ($\overline{X}$ = 8.9, *SD* = 4.1) classrooms per school. Classrooms also were distributed across Grades 3 to 8, with 80.2% representing Grades 3 to 6. The number of students participating in each classroom ranged from 1 to 19 ($\overline{X}$ = 9.0, *SD* = 3.9). The majority of classrooms in Grades 3 to 5 were self-contained (with some fifth-grade students changing classrooms for particular subjects). The majority of students in Grades 6 to 8 changed classrooms for different subjects per middle school norms. Class size data were obtained from observations (*n* = 61) and teacher reports (*n* = 3); five class sizes were unreported. The total number of students in each classroom ranged from 12 to 33 ($\overline{X}$ = 24.3, *SD* = 4.4).

Teachers. Teachers (*n* = 69) were primarily female (81.2%), African American (47.8%), Caucasian (27.5%), Hispanic (8.7%), Other (5.8%), and unreported (10.2%) and ranged in age from 23 to 62 years ($\overline{X}$ = 35.4 years, *SD* = 10.26). The number of teachers from each school ranged from 4 to 17 ($\overline{X}$ = 8.9, *SD* = 4.1).

Students. The total sample included 621 students in Grades 3 through 8 ($\overline{X}$ = 4.6, *SD* = 1.5), ages 8 to 15 ($\overline{X}$ = 10.22, *SD* = 1.63), 68.8% African American, 22.4% Latino, 1.0% White, 5.3% multi-ethnic, 0.6% American Indian/Alaskan Native, 1.9% unreported, and 50% female. Students in Grades 6 through 8 comprised 25% of the sample. Students represented 11 classrooms across two schools in Year 1 (*n* = 102), 36 classrooms across four schools in Year 2 (*n* = 308), and 22 classrooms across two schools in Year 3 (*n* = 211). Recruitment at each school occurred in two waves (fall and spring, described in further detail below), mobility in the schools was high, and students changed classrooms for many reasons throughout the year, resulting in variable sample sizes across time points: Time 1 (*n* = 263), Time 2 (*n* = 391), and Time 3 (*n* = 565). The bifactor model was calibrated on one time point with the most data (Time 3) from six schools once sufficient item-level data became available following Year 2 of the study, for a subsample of *n* = 335 students. Mixed-effects linear regression models were estimated utilizing longitudinal data from the full sample of *n* = 621 students, including all 3 years of the study, all eight schools, and all three time points.

Measures

Student Climate Survey. The development of the Student Climate Survey began with a close examination of items from the widely used OSC measure (Glisson et al., 2012; Glisson, Landsverk, et al., 2008) that was developed over three decades to assess climate, culture, and work attitudes in organizations that provide mental health services, social services, and other human services. The scales comprising the OSC have been tested in thousands of these organizations throughout the United States and in other countries. The factor structure of the OSC and the items comprising climate have been confirmed in two nationwide studies and shown to predict key service quality and outcome criteria in scores of additional studies (e.g., Glisson, 1978; Glisson & Durick, 1988; Glisson et al., 2012; Glisson & Hemmelgarn, 1998; Glisson et al., 2013; Glisson, Landsverk, et al., 2008; Glisson, Schoenwald, et al., 2008).

Items deemed relevant to classrooms as work units were selected and rewritten for Grades 3 to 8 in accordance with the following considerations. First, items (*n* = 66) reflecting personal impact of work (e.g., amount, difficulty, resources, frustration, attitudes); relationships (e.g., number and nature of connections, frequency/intensity of positive/negative interactions); opportunities for autonomy, growth, learning, and achievement (e.g., achievement expectations, accomplishment, reward); and classroom organization (e.g., rules, routines, behavior expectations, equity) were retained. Items (*n* = 49)

reflecting professional development (e.g., advancement), organizational commitment (e.g., employment options), structure (e.g., operating procedures, management), and personality (e.g., temperament) were eliminated. Second, item content was reworded to be developmentally appropriate and accurately reflect the classroom environment while also retaining the underlying meaning and nuance of each item. For example, "I feel used up at the end of the work day" became "I feel worn out at the end of the school day," and "I deal very effectively with the problems of the clients I serve" became "I deal very well with problems in my class."

Third, Microsoft Works Readability statistics were used to assess the Flesh-Kincaid readability level of each reworded item. Flesh-Kincaid scores account for the number and difficulty of words and multi-syllabic words in a sentence. Items that scored higher than a third-grade reading level were revised again, and the readability analysis was repeated, toward the goal of having all items achieve a Flesh-Kincaid score at or below Grade 3.0. Following this, items were reviewed to eliminate redundancy or items that held little relevance for students and classrooms, resulting in the final measure with 53 items used in the present study. To make the response options more suitable for administration to young students, the rating scale was revised from a 5-point to a 4-point format with response options "NO," "no," "yes," and "YES." Following procedures previously utilized by the Metropolitan Area Child Study Research Group (2007), these response options were verbally presented to students to reflect very strong endorsement of "NO," moderate endorsement of "no," moderate endorsement of "yes," and very strong endorsement of "YES."

Academic Motivation Inventory (AMI). The AMI (Ginsburg-Block & Fantuzzo, 1998) is a 13-item, 3-point scale that assesses student motivation for learning, adapted for a population of primarily African American, low-achieving third to fourth graders in urban elementary school. Internal consistency was moderately strong ($\alpha = .77$).

School Liking and Avoidance Questionnaire (SLAQ). The SLAQ (Ladd & Price, 1987) is a 14-item, 3-point scale assessing student satisfaction with school along two domains: School Liking (9 items, $\alpha = .82$) and School Avoidance (5 items, $\alpha = .70$). The subscales were moderately correlated ($r = -.64$). The form for children in Grades 3 to 6 was used.

Behavioral Observation of Students in Schools (BOSS). The BOSS (Shapiro, 2004) uses momentary time sampling to measure engaged and off-task behavior of targeted students in elementary classrooms, along with peer

comparison data every fifth interval. Observations included two 15-minute time samples (sixty 15-second intervals) on two consecutive days. Scores were computed to reflect percentage of time engaged and off-task (DuPaul et al., 2004; Ota & DuPaul, 2002), and these were highly correlated (r = −.91). BOSS observers were trained to a minimum of 80% inter-observer agreement with an expert rater at the beginning of each school year. Due to time and resource constraints, BOSS observations were obtained for a subset of randomly selected students in each classroom during Years 1 and 3 (n = 133).

Curriculum Based Reading Measures (CBM). Academic performance was assessed using standardized oral fluency and reading comprehension probes for Grades 3 to 8 obtained from AIMSweb (http://www.aimsweb.com; Shapiro, 2004). To assess oral fluency, students read aloud for 1 minute from three grade-level reading passages. Scores were computed as the median number of words read correctly. To assess reading comprehension, students read silently for 3 minutes from a passage during which they were required to substitute every seventh word with one of three options that included the semantically correct word and two distractors. Scores were computed as the number of correct selections. Reading fluency and reading comprehension were highly correlated (r = .79). Research assistants were trained at the beginning of each school year on standardized AIMSweb master-coded DVDs to a criterion of 80% agreement.

Social Skills Improvement System (SSIS). Teachers reported problem behaviors (46 items, α = .96) and social skills (30 items, α = .98) along a 4-point scale. Academic competence (seven items, α = .96) was based on each student's performance compared with their peers on a range of subjects (1 = *lowest 10%* to 5 = *highest 10%*). The subscales were moderately correlated (r ranged from −.56 to −.73). The SSIS (Gresham & Elliot, 2008) was not utilized in Year 1 of the study. To minimize burden during Years 2 and 3, teachers completed the SSIS for a subset of randomly selected students in their classroom (n = 193) at all three waves of data collection.

Procedures

This study was conducted with approval from the Institutional Review Board (IRB) and school district Research Review Board for recruitment, informed consent, and data collection procedures.

School recruitment. Eight schools participated in three cohorts: 2 schools in Year 1, 4 schools in Year 2, and 2 schools in Year 3. The two schools that

participated during Year 1 were involved in a larger ongoing study with the investigative team. Two schools that participated during Year 2 had been involved in prior research with the investigative team. These 4 schools were originally selected from a list of 58 elementary schools (of 325 in the district) that met the following criteria: (a) 85% or greater low income, (b) 85% or greater African American students, (c) average reading scores on statewide testing below the 35th percentile ($\overline{X}$ = 27.9, *SD* = 3.8), and (d) school population within a standard deviation of the district mean ($\overline{X}$ = 702, *SD* = 306). Four new schools were selected for participation in Years 2 and 3 based on their large size and close proximity to the university from a list of 35 schools meeting similar criteria: (a) 80% or greater low-income households, (b) 50% or greater underrepresented minority students, (c) average reading scores on statewide testing below the 65th percentile. School recruitment followed IRB- and district-approved procedures used previously by the investigative team, beginning with written correspondence to school principals during the spring prior to their participation. Discussion with principals was followed by a series of additional meetings, through summer and early fall, with school administrators, local school councils (parents and teachers), and school personnel regarding the objectives and procedures of the research.

Teacher recruitment and consent. Recruitment and consent procedures were specific to the current study and identical across all eight schools, including the four schools involved in prior or ongoing research by the investigative team. Research staff introduced and distributed information about study objectives and procedures during a pre-arranged meeting on fall teacher planning days. They scheduled subsequent individual meetings with teachers to provide further detail, answer questions, and obtain written informed consent.

Parent permission and student assent. As for teachers, recruitment and consent procedures were specific to the current study and identical across all eight schools. Recruitment occurred in two waves, during fall and spring report card distribution days, designed for parents to visit their child's classroom and meet briefly with their child's teacher. With permission from each principal and from participating teachers, research staff arranged tables with research materials outside of participating classrooms. Teachers encouraged parents to spend a few minutes after their conference learning about the objectives and procedures of the study. Research staff provided detailed information, answered questions, and invited parents to enroll their child. Parents who attended the fall report card day and agreed for their children to participate provided written informed consent at that time. Their children

participated in all three waves of data collection. Parents who requested more time to consider participation provided their name and phone number. Research staff contacted them during the weeks (and sometimes months) that followed and (if appropriate) scheduled meetings at the school to obtain written permission for their children to participate in the study. Several of these students enrolled following completion of Time 1 data collection; therefore, they participated at Time 2 and Time 3. Students for whom parental consent was obtained during spring report card day participated only in Time 3 data collection, accounting for differences in sample size across time points but increasing the power for Time 3 IRT analyses. Following parent permission, research staff assented individual students during the school day.

Data collection. Data collection procedures were consistent across all eight schools. Data were collected at three time points (fall: August-December, winter: January-March, spring: April-June) during one school year. Information was collected from teachers outside of contract hours, and they received a box of classroom supplies to compensate them for their time. All participating students completed the Student Climate Survey during classroom time pre-arranged with each teacher. To minimize interference with instructional time and adhere to school district conditions for research, students within classrooms were randomly assigned to complete either the SLAQ or the AMI in order not to exceed one half hour of classroom data collection at each time point. Measures were group administered, and students completed them independently at their desks, with assistance from research staff as needed. Nonparticipating students worked quietly at their seats during data collection. All students received small prizes as tokens of appreciation (e.g., pencils, erasers). Students in Grades 6 through 8 who changed classrooms throughout the school day completed measures during their homeroom (also their first academic class) and were instructed to use that class as a reference. Data recorded by research staff during classroom observations at each time point indicated that despite students changing classrooms throughout the school day, individual teachers still taught multiple academic subjects (reflecting school district emphasis on integrated curriculum), and homerooms therefore included a combination of instruction in the following areas, ordered from most frequently observed to least frequently observed: English (40%; language arts, literature, reading, writing), social studies (27%), math (21%), and science (12%). For these students, homeroom teachers also completed the SSIS. Research assistants removed students individually from their classrooms to administer CBMs privately and observed students with the BOSS during morning instruction on days pre-arranged with individual teachers and approved by the school administration.

Data Analytic Plan

A multidimensional IRT analysis based on a full-information item bifactor model (Gibbons et al., 2007; Gibbons & Hedeker, 1992) was used to calibrate a Time 3 subsample, following Year 2 of the study (6 schools, $n = 335$ students), once sufficient item-level data became available. The advantage of the bifactor model over unidimensional IRT models is that it extends the assumption of conditional independence to all of the domains from which the items were sampled and not just the primary domain that the test measures. The advantage of the bifactor model over an unrestricted full-information item factor analytic model is that it is computationally tractable for any number of dimensions (i.e., domains) whereas item factor analysis is limited to no more than five or six dimensions (see Gibbons & Hedeker, 1992). The bifactor model requires all items to load on the primary dimension and no more than one subdomain. All bifactor models were fitted to the data using IRTPRO (Cai, du Toit, & Thissen, 2009). We began with the full set of 53 items, which were then rescored in a common direction and classified by their content into one of nine theoretically meaningful domains corresponding to previous research on climate (Glisson & Hemmelgarn, 1998) but postulated to reflect children's experiences in classrooms: Disinterest, Fatigue, Equity, Involvement, Predictability, Discord, Pressure, Peer Support, and Self-Efficacy. Based on the original (53 items) and final (18 items) bifactor models, Bayes estimates of the primary dimension scores were computed for each student ($n = 621$) at each of the three measurement occasions (within one school year) and were then related to longitudinal changes in student outcome measures using a three-level (time nested within student nested within classrooms) linear mixed-effects regression model (Hedeker & Gibbons, 2006).

To aid in interpretability, a second model was fitted on a log scale (for the outcome) so that parameters can be interpreted as percentage change in response to a one standard deviation unit change in the primary dimension scale score. For example, if the estimated parameter for the primary dimension score is 0.2, then a person having a primary dimension score of 1.0 has a 20% increase in the outcome relative to a person with primary dimension score of 0.0. Across the underlying primary dimension scale, there is a 5-point range (underlying normal random variable from −2.5 to 2.5); hence, there is a 0.2*5*100 = 100% increase from the low end of the scale to the high end of the scale in terms of change in the outcome measure.

To assess stability of the trait over time, the IRT model was fitted separately at each of the three time points. Estimated factor loadings were compared visually and by computing raw and weighted root mean square error

(RMSE). Raw RMSE is in the original scale of measurement (an underlying unit normal distribution) whereas the weighted is scale invariant and describes agreement in terms of the percentage of a standard error.

Results

Results of fitting the bifactor model with all nine subdomains and all 53 items are displayed in Table 1. The estimated factor loadings in Table 2 reveal that several of the items do not have high loadings on the primary climate dimension. A final bifactor model was derived by iteratively deleting items until all of the items had a loading of 0.4 or greater on the primary climate dimension. Deleted items were distributed evenly across domains, such that at least 50% of items within each domain were dropped (except Disinterest, for which only one of three items was dropped). The ninth domain, Peer Support (four items), was dropped completely.

The final model parameters based on 18 items and eight subdomains are displayed in Table 2. The category thresholds and observed and expected proportions are displayed in Table 3, which indicates that the model did an excellent job of tracking the observed item response proportions. The bifactor model provided significant improvement in fit over a unidimensional IRT model alternative ($\chi^2 = 71.8$, $df = 18$, $p < .0001$). This reduced set of items can be used to derive psychological climate scores, either by a simple summation or preferably by obtaining Bayes estimates as described by Gibbons and colleagues (2007). These 18 items and response categories are presented in Table 4.

Table 5 presents results for criterion and predictive validity based on bifactor models fitted to the original 53-item survey and the better fitting 18-item version. Scores based on the primary dimension from the bifactor model were used as time varying predictors in a three-level mixed-effects linear regression model (Hedeker & Gibbons, 2006) for each academic endpoint. Both 53-item and 18-item scores revealed strong and significant longitudinal associations in the expected direction with student-reported academic motivation and school liking and avoidance (indicators of criterion validity). Both measures also revealed strong and significant longitudinal associations in the expected direction with teacher-reported social skills, problem behaviors, and academic competence, and with curriculum-based measures of reading fluency (trend) and comprehension (indicators of predictive validity). Associations with observed classroom behavior (engaged and off-task) were nonsignificant. The correlations between the 18- and 53-item versions of the measure were .83, .88, and .87 at Time Points 1, 2, and 3, respectively.

Table 1. Bifactor Solution for 53 Items: Factor Loadings and Standard Errors.

		Factor									
Subdomain	Item	Primary factor	2	3	4	5	6	7	8	9	10
Disinterest	9	.53 (.06)	.84 (.04)								
	16	.63 (.08)	.23 (.11)								
	19	.25 (.10)	.21 (.12)								
Fatigue	4	.48 (.10)		.44 (.13)							
	7	.14 (.11)		.57 (.14)							
	11	.41 (.10)		.52 (.14)							
	14	.52 (.09)		.56 (.12)							
	17	.16 (.11)		.29 (.13)							
	20	.48 (.09)		.26 (.13)							
Equity	23	.31 (.10)			.13 (.19)						
	24	.57 (.09)			.09 (.27)						
	25	−.19 (.11)			.57 (.77)						
	37	.63 (.08)			.04 (.23)						
	41	.34 (.11)			.21 (.50)						
Involvement	8	.38 (.10)				.32 (.22)					
	12	.55 (.09)				.26 (.19)					
	18	.59 (.09)				−.01 (.15)					
	21	.56 (.09)				.62 (.32)					
	22	.47 (.09)				.22 (.18)					
Predictability	38	.07 (.11)					.43 (.17)				
	45	.57 (.10)					.33 (.15)				

(continued)

Table 1. (continued)

Subdomain	Item	Factor									
		Primary factor	2	3	4	5	6	7	8	9	10
	47	.52 (.10)					.65 (.21)				
	52	.49 (.10)					.30 (.14)				
Discord	2	.28 (.11)						.18 (.15)			
	10	.63 (.08)						.19 (.12)			
	13	.53 (.09)						.14 (.13)			
	43	.42 (.10)						.69 (.13)			
	48	.50 (.10)						.55 (.12)			
	49	.25 (.11)						.48 (.13)			
Pressure	1	.44 (.09)							.03 (.14)		
	5	.20 (.11)							.34 (.16)		
	39	.08 (.11)							.54 (.18)		
	42	.39 (.10)							.46 (.15)		
	46	.18 (.11)							.28 (.15)		
	50	.27 (.10)							.31 (.15)		
	51	.34 (.10)							.43 (.16)		
Peer Support	3	.15 (.11)								.27 (.17)	
	40	.49 (.09)								.38 (.17)	
	44	.46 (.09)								.26 (.16)	
	53	.29 (.10)								.60 (.24)	

(continued)

Table 1. (continued)

		Factor									
Subdomain	Item	Primary factor	2	3	4	5	6	7	8	9	10
Self-Efficacy	26	.58 (.09)									.14 (.15)
	27	.60 (.12)									−.24 (.21)
	28	.57 (.09)									.17 (.13)
	29	.51 (.09)									.10 (.14)
	30	.59 (.09)									.16 (.14)
	32	.30 (.10)									.57 (.18)
	33	.11 (.11)									.66 (.20)
	34	.65 (.08)									−.07 (.14)
	35	.71 (.08)									−.13 (.15)
	36	.66 (.11)									−.25 (.21)
Primary Factor Only	6	.46 (.13)									
	15	.65 (.09)									
	31	.68 (.09)									

Note. Fitted by IRTPRO (Cai, du Toit, & Thissen, 2009).

Table 2. Eight-Dimensional Bifactor Solution for 18 Items: Factor Loadings.

		Factors							
Subdomain	Item	1	2	3	4	5	6	7	8
Disinterest	9	0.745	0.180						
	16	0.766	−0.046						
Fatigue	4	0.438		0.526					
	11	0.450		0.583					
Equity	41	0.476							
Involvement	18	0.587			−0.058				
	22	0.462			0.243				
Predictability	45	0.524				0.275			
	52	0.415				0.434			
Discord	13	0.639					−0.002		
	43	0.519					0.577		
	48	0.534					0.554		
Pressure	1	0.430						0.116	
	42	0.500						0.210	
	51	0.434						0.706	
Self-efficacy	27	0.480							0.349
	28	0.416							−0.056
	30	0.595							0.144

Table 6 displays estimated primary factor loadings at each of the three time points. In general, the loadings are quite similar across the three time points indicating a stable trait. The raw RMSEs were 0.13 between Times 1 and 2, 0.15 between Times 1 and 3, and 0.09 between Times 2 and 3. These absolute differences translate into 0.65, 0.78, and 0.65 standard error differences, all of which demonstrate trait invariance over time.

Discussion

This study introduced the construct of student psychological climate and its potential to contribute uniquely to the comprehensive measurement of classroom ecology. Results indicated that psychological climate—perceptions of the psychological impact of one's work environment on one's own well-being and mental health—reasonably may extend to children in their classroom work environment. Based on a multidimensional IRT analysis, we extracted theoretically meaningful climate items that loaded on a primary student climate dimension. The derived 18-item measure performed better

Table 3. Estimated Category Thresholds, and Observed and Expected Proportions From the Eight-Dimensional Graded Bifactor Analysis.

Subdomain	Item number	Item thresholds			Observed and expected proportions[a]			
		1-2	2-3	3-4	1	2	3	4
Disinterest	9	−1.019	−0.650	−0.226	0.148 (0.154)	0.118 (0.104)	0.160 (0.153)	0.574 (0.589)
	16	−1.119	−0.790	−0.249	0.127 (0.132)	0.090 (0.083)	0.198 (0.187)	0.586 (0.598)
Fatigue	4	−0.871	–0.335	0.181	0.189 (0.192)	0.173 (0.177)	0.208 (0.203)	0.429 (0.428)
	11	−0.415	0.199	0.588	0.341 (0.339)	0.241 (0.240)	0.146 (0.143)	0.272 (0.278)
Equity	41	−1.198	−0.770	−0.315	0.113 (0.116)	0.110 (0.105)	0.159 (0.156)	0.618 (0.624)
Involvement	18	−1.023	−0.768	−0.097	0.150 (0.153)	0.069 (0.068)	0.246 (0.240)	0.536 (0.539)
	22	−1.077	−0.659	0.190	0.138 (0.141)	0.116 (0.114)	0.324 (0.320)	0.422 (0.425)
Predictability	45	−1.670	−1.329	−0.460	0.046 (0.047)	0.043 (0.044)	0.232 (0.231)	0.679 (0.677)
	52	−1.338	−0.921	−0.122	0.088 (0.090)	0.091 (0.088)	0.276 (0.273)	0.545 (0.548)
Discord	13	−0.749	−0.143	0.310	0.225 (0.227)	0.222 (0.216)	0.182 (0.179)	0.370 (0.378)
	43	−0.883	−0.515	−0.007	0.188 (0.189)	0.120 (0.115)	0.197 (0.194)	0.495 (0.503)
	48	−1.118	−0.541	−0.135	0.128 (0.132)	0.171 (0.163)	0.156 (0.152)	0.545 (0.554)
Pressure	1	−1.090	−0.195	0.490	0.137 (0.138)	0.286 (0.285)	0.267 (0.266)	0.310 (0.312)
	42	−0.763	−0.183	0.354	0.221 (0.223)	0.212 (0.205)	0.212 (0.211)	0.355 (0.362)
	51	−0.789	−0.281	0.241	0.214 (0.215)	0.180 (0.174)	0.207 (0.206)	0.399 (0.405)
Self–Efficacy	27	−1.634	−1.006		0.049 (0.051)	0.107 (0.106)	0.845 (0.843)	
	28	−1.202	−0.747	0.306	0.113 (0.115)	0.113 (0.113)	0.393 (0.393)	0.380 (0.380)
	30	−1.198	−0.810	−0.189	0.111 (0.116)	0.096 (0.093)	0.222 (0.216)	0.571 (0.575)

Note. Empty cells reflect that two categories with frequency rate less than 0.03 were collapsed into one.
[a]Expected proportions in parentheses.

Table 4. Eighteen Items by Subdomain, and Response Proportions (n = 335).

Subdomain	Item	Content	YES	yes	no	NO
Disinterest	9	I care less about school since I entered this class	0.148	0.118	0.160	0.574
	16	This class makes me care less about others	0.127	0.090	0.198	0.586
Fatigue	4	I'm tired of school	0.189	0.173	0.208	0.429
	11	When I wake up in the morning I don't want to go to school	0.341	0.241	0.146	0.272
Equity	41	Being liked is more important than doing well in my class	0.113	0.110	0.159	0.618
Involvement	18	I like school	0.150	0.069	0.246	0.536
	22	I stay calm when other kids in my class have problems	0.138	0.116	0.324	0.422
Predictability	45	I know what I am supposed to do in my class	0.046	0.043	0.232	0.679
	52	Goals in my classroom are clear	0.088	0.091	0.276	0.545
Discord	13	School gets in the way of fun time	0.225	0.222	0.182	0.370
	43	Rules get in the way of getting my work done	0.188	0.120	0.197	0.495
	48	I have to do things in my class that I don't think are right	0.128	0.171	0.156	0.545
Pressure	1	My classmates look stressed out	0.137	0.286	0.267	0.310
	42	The amount of work I have makes it hard to do a good job	0.221	0.212	0.212	0.355
	51	I feel a lot of pressure in my class	0.214	0.180	0.207	0.399
Self-efficacy	27	I feel good when I do my work the right way	0.049	0.107	0.845	
	28	I help other kids in my class	0.113	0.113	0.393	0.380
	30	Rules are fair in my class	0.111	0.096	0.222	0.571

Note. Two categories with frequency rate less than 0.03 were collapsed into one.

Table 5. Estimated Coefficients (and Percentage Change in Outcome for 1 *SD* Unit Change in Primary Factor Score) of Score on Primary Factor in Random Effects Model.

	Time 1		Time 2		Time 3		Coefficient of primary factor score	
Outcome	*N*	*M* (*SD*)	*N*	*M* (*SD*)	*N*	*M* (*SD*)	53 items	18 items
Academic Motivation	122	18.80 (4.33)	189	18.16 (4.72)	269	18.19 (4.83)	2.383 (14.7%)***	2.678 (16.1%)***
School Liking	110	2.45 (0.41)	179	2.47 (0.43)	257	2.32 (0.45)	0.240 (7.9%)***	0.288 (8.9%)***
School Avoidance	109	1.87 (0.51)	179	1.90 (0.54)	256	2.03 (0.52)	−0.230 (−8.2%)***	−0.322 (−11.0%)***
Social Skills	51	85.63 (27.63)	43	88.12 (21.61)	83	88.67 (27.55)	10.615 (14.0%)***	9.319 (11.5%)***
Problem Behavior	52	17.90 (17.63)	47	15.64 (13.40)	94	18.00 (16.31)	−5.425 (−46.0%)***	−5.345 (−43.6%)***
Academic Competence	47	22.68 (7.80)	45	22.73 (8.31)	86	22.91 (8.26)	2.015 (7.9%)*	2.244 (10.4%)**
Observed Engagement	54	55.95 (17.84)	20	64.39 (15.63)	97	61.30 (19.12)	2.112 (4.2%)	1.129 (2.7%)
Observed Off-Task	54	56.32 (18.77)	22	42.39 (20.54)	98	43.46 (18.68)	−2.252 (−5.8%)	−1.294 (−3.1%)
Reading Fluency	219	90.17 (42.21)	308	103.81 (42.63)	443	111.14 (43.23)	1.073 (1.3%)****	1.422 (1.6%)****
Reading Comprehension, Correct Responses	214	9.79 (6.41)	310	11.94 (6.90)	439	14.02 (7.88)	0.547 (3.6%)****	0.664 (4.9%)†
Reading Comprehension, Errors	214	3.10 (2.80)	310	2.81 (3.25)	439	3.15 (3.81)	−0.338(−5.8%)†	−0.506 (−7.4%)*

† $p < .05$. * $p < .01$. ** $p < .001$. *** $p < .0001$. **** $p < .10$.

Table 6. Stability of the Primary Factor Loadings Over Time for the 18-Item Solution.

		Factor		
Subdomain	Item	Time 1 (*SE*)	Time 2 (*SE*)	Time 3 (*SE*)
Disinterest	9	.44 (.17)	.64 (.10)	.74 (.08)
	16	.44 (.17)	.57 (.10)	.65 (.09)
Fatigue	4	.68 (.11)	.72 (.07)	.57 (.08)
	11	.53 (.13)	.62 (.08)	.52 (.08)
Equity	41	.43 (.18)	.28 (.13)	.42 (.09)
Involvement	18	.80 (.11)	.69 (.09)	.68 (.07)
	22	.23 (.17)	.42 (.11)	.38 (.11)
Predictability	45	.37 (.18)	.45 (.13)	.55 (.09)
	52	.53 (.14)	.49 (.10)	.48 (.10)
Discord	13	.58 (.14)	.67 (.08)	.62 (.09)
	43	.57 (.13)	.61 (.10)	.47 (.31)
	48	.45 (.16)	.48 (.11)	.49 (.25)
Pressure	1	.11 (.16)	.21 (.11)	.36 (.08)
	42	.30 (.16)	.54 (.08)	.47 (.08)
	51	.38 (.15)	.51 (.10)	.43 (.08)
Self-efficacy	27	.23 (.25)	.43 (.16)	.49 (.10)
	28	.48 (.14)	.49 (.11)	.45 (.09)
	30	.66 (.13)	.58 (.10)	.58 (.08)

Note. SE = standard error.

than the original 53-item version, demonstrating good criterion and predictive validity, thus meeting criteria as an instrument with properties of both effectiveness (reliable and valid) and efficiency (brief and feasible to administer) as recommended by Schoenwald and colleagues (2011).

More specifically, 18 of 53 items survived to comprise the *Student Climate Survey* presented here. All 18 items were good discriminators and represented eight of the nine proposed underlying theoretical dimensions, suggesting that student psychological climate, like the psychological climate of work units in other contexts, is itself multifaceted. The 18-item scale, like the initial 53-item version, longitudinally predicted proximal student-reported outcomes of academic motivation and school liking and avoidance and distal teacher-reported student problem behaviors, social skills, and academic competence as well as curriculum-based measures of reading fluency and comprehension.

Contrary to expectations, neither the 18-item nor the 53-item scale predicted observed on- or off-task student behaviors, despite an extensive

literature relating student engagement with academic performance (DiPerna, Volpe, & Elliott, 2002). Although this may be attributed to relatively low power (given that only slightly more than one quarter of the sample was observed), it may also reflect the sensitivity of student engagement to teacher instruction and behavior management (Stringfield, 1994). Nevertheless, given that the *Student Climate Survey* related to teacher reports of academic competence and to curriculum-based measures of reading performance, this finding merits replication with a larger number of students.

The original 53 items were classified into theoretically meaningful domains that included Disinterest, Fatigue, Equity, Involvement, Predictability, Discord, Pressure, Peer Support, and Self-Efficacy. We employed a fairly stringent threshold of 0.4 in fitting the bifactor model in order to produce a very clear pattern with a reduced set of items to decrease participant burden. Deleted items were distributed in roughly equivalent proportions across eight domains (50% items dropped). A ninth domain, Peer Support (four items), was dropped completely. Close examination of the items offers a possible explanation. The majority of items on the original 53-item survey described a student's personal experience (e.g., "I feel good when I do my work the right way" or "I stay calm when other kids in my class have problems"). On the other hand, the four items representing Peer Support appeared to characterize the classroom environment more than its impact on individual well-being (e.g., "Kids in my class argue" or "There is a feeling of teamwork in my class"), reflecting culture (i.e., group norms) more than climate (i.e., personal experience) as defined by the organizational literature, and perhaps explaining why these items failed to survive.

Unique Contribution of Student Psychological Climate

Consistent with an extensive literature on organizational social context, and closely aligned with findings of teacher stress and burnout, results from the present study suggest that student psychological climate may add another dimension to our comprehensive understanding of classroom ecology. Several well-established and widely utilized measures (including teacher-report, student-report, and independent observations) already yield rich descriptions of the educational environment with regard to instruction, organization, relationships, productivity, and management. Reflecting a conceptualization of "climate" that characterizes the education literature, these measures ascribe climate to the level of school or classroom (e.g., Koth, Bradshaw, & Leaf, 2008; Wang, 2009; Zullig, Huebner, & Patton, 2011). The *Student Climate Survey* extends our understanding of classrooms—not by characterizing the environment—but by assessing the extent to which

individual students perceive it to be a healthy or unhealthy workspace. Findings indicated that students reporting a positive psychological climate enjoy school more, are more motivated, exhibit more teacher-reported social skills, and perform well academically, whereas students reporting a negative psychological climate dislike school, are less motivated to achieve, exhibit more teacher-reported behavior problems, and struggle academically. These findings mirror reports by teachers, especially early in their careers, when chronic stress (reflecting, for example, inadequate administrative support, isolation, and high rates of student behavior problems) predicts harsher discipline, reduced effort, and poor quality instruction (e.g., Shernoff et al., 2011), disengagement (Hakanen, Bakker, & Schaufeli, 2006), and attrition (Ingersoll & Strong, 2011).

Reflecting an extensive organizational literature that links psychological climate to work attitudes, behaviors, and outcomes (Glisson, 2002), we propose that student psychological climate may help to identify early students at risk for disengagement, academic failure, and school dropout. Increasing attention to school dropout as a public health issue reflects ongoing disparities in health and education and unacceptably low promotion and school completion rates, in particular among African American, Latino, and American Indian students, and especially in our nation's largest urban centers (Freudenberg & Ruglis, 2007). While contributing factors predictably reflect a convergence of individual, family, school, and neighborhood characteristics, the eight dimensions underlying student psychological climate together reflect several individual factors associated with dropout, including not liking school (Involvement), not belonging (Disinterest), perceptions of unfair or harsh discipline (Discord), feeling unsafe (Pressure), and not engaged (Fatigue) as summarized by Freudenberg and Ruglis (2007). Therefore, while other student-report measures assess classroom-level characteristics (e.g., instruction, academic press, warmth, organization) that may warrant classroom-level intervention, or one particular dimension of the school experience (e.g., school liking, motivation, or attachment), the *Student Climate Survey* provides a multifaceted but economical assessment of individual-level experiences that may reveal early risk and facilitate student-level intervention. This may be especially important in the context of a literature that suggests school performance and adaptation as early as first-grade initiate stable trajectories and predict long-term academic success or failure (Ensminger & Slusarcick, 1992).

Related, persistently high rates of mental health need in middle and high school—reflected by increased violence, substance abuse, and suicide attempts—are fueling efforts nationwide to equip school personnel to identify and respond to youth in distress via school-wide screenings (e.g., Levitt,

Saka, Romanelli, & Hoagwood, 2007) and gatekeeper programs (e.g., Wyman et al., 2008). While attention is largely directed to early identification of specific behavioral (e.g., impulsivity, conduct problems) and emotional (e.g., depressed mood) symptoms, we propose the *Student Climate Survey* may offer a brief and easy-to-administer tool that yields unique and important information related to children's perceptions of their daily school experience, its impact on their mental health, and possible levers for change.

Limitations and Future Directions

As noted above in procedures, school district regulations limited the instructional time allocated to administer student measures. Therefore, whereas all participating students completed the *Student Climate Survey*, they were randomly assigned to receive measures associated with criterion validity, thereby reducing the sample size for those analyses. Strong longitudinal associations resulted, nonetheless, between student climate and academic motivation and school liking and avoidance. Although method variance may have contributed to this association, its strength over time, coupled with findings of predictive validity, help attenuate this concern. Relatedly, resource constraints minimized capacity of the investigative team to conduct student time-on-task observations, resulting in a very small sample, which, as noted, may partially explain the nonsignificant association between student climate and observed engagement. Second, the sample includes a broad age range of students, from third to eighth grades. Although there are important developmental and classroom differences between elementary and middle school, the theoretical dimensions underlying the construct of psychological climate were determined to have broad relevance across grades. Furthermore, although older students changed classrooms, their homeroom was consistently used across student and teacher measures as the reference, and ultimately their climate score represents an individual-level psychological experience of their overall educational environment.

This study was conducted in urban schools with high proportions of economically disadvantaged students and characterized by low standardized test scores in reading. Previous research documents the robust disparity in educational and mental health needs that affects students and teachers in high-poverty urban schools, evidenced by higher prevalence of cognitive deficits, disruptive behavior problems, and emotional disturbance (Cappella, Frazier, Atkins, Schoenwald, & Glisson, 2008; Murali & Oyebode, 2004). We believe these disproportionate needs place underperforming urban schools at greatest risk for students to experience a negative psychological climate and hence most warranted a close examination of student classroom experiences. Hence,

the impetus for development of this survey was to provide a standardized assessment of factors affecting teaching and learning in urban elementary schools. Future work will need to examine the extent to which the contribution of the *Student Climate Survey* for understanding classrooms extends to schools situated in other communities (e.g., rural) or characterized by other populations (e.g., less economic disadvantaged, other racial or ethnic groups).

Concluding Remarks

This study extends school mental health services research in communities of concentrated urban poverty to consider student reports of psychological climate, informed to a significant extent by the earlier work of Glisson and colleagues on the importance of organizational social context across a wide array of service settings. Ongoing research will examine the extent to which student reports of psychological climate within a classroom can be aggregated to comprise a classroom-level variable, its degree of correspondence with teacher psychological climate, and the concurrent correspondence of each to a measure of school-wide culture and climate. Specifically, we plan to examine associations among student climate, teacher climate, observed classroom processes (e.g., instruction, behavior management), and the extent to which climate may help to explain associations between classroom functioning and student outcomes. These data are beginning to inform a vision for classroom-level intervention supported by the literature on organizational social context interventions and aligned with a problem-solving model by which students participate in identifying and solving the barriers to learning in their classrooms.

In conclusion, there is long-standing interest among researchers, school personnel, and policy stakeholders to understand the extent to which classroom environment and processes influence teacher practice and student outcomes. Results from this study indicate that classrooms are unique work units in which both teachers and students contribute to classroom processes, and in which students are both contributors to and recipients of the services provided by the teachers and schools. Student descriptions of their classrooms, therefore, are unlikely to capture adequately the extent to which each individual student perceives his or her environment to influence—positively or negatively—his or her own mental health. The current findings suggest that student psychological climate is strongly and significantly related to academic and behavioral outcomes. Future research is warranted to examine the extent to which student perceptions contribute to our comprehensive understanding of classroom functioning and the linkages between teacher practices and student experiences.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Primary support for this article was provided by the National Institute of Mental Health (NIMH) Research Grant P20MH078458 (Principal Investigator [PI]: Atkins). The authors are grateful for the time and contribution of participating schools, teachers, and students.

References

Anderson, H. H., & Brewer, H. M. (1945). Dominative and socially integrative behavior of kindergarten teachers [Monograph]. *Applied Psychology Monographs, No. 6*.

Anderson, H. H., & Brewer, J. E. (1946). Effects of teachers' dominative and integrative contacts on children's classroom behavior [Monograph]. *Applied Psychology Monographs*, *No. 8.*

Brand, S., Felner, R., Shim, M., Seitsinger, A., & Dumas, T. (2003). Middle school improvement and reform: Development and validation of a school-level assessment of climate, cultural pluralism, and school safety. *Journal of Educational Psychology*, *95*, 570-588. doi:10.1037/0022-0663.95.3.570

Byrne, B. M. (1999). The nomological network of teacher burnout: A literature review and empirically validated model. In R. Vandengerghe & A. M. Huberman (Eds.), *Understanding and preventing teacher burnout: A sourcebook of international research and practice* (pp. 15-37). Cambridge, UK: Cambridge University Press.

Cai, L., du Toit, S. H. C., & Thissen, D. (2009). IRTPRO: Flexible, Multidimensional, Multiple Categorical IRT Modeling [Computer software]. Chicago, IL: SSI International.

Cappella, E., Frazier, S. L., Atkins, M. S., Schoenwald, S. K., & Glisson, C. (2008). Enhancing schools' capacity to support children in poverty: An ecological model of school-based mental health services. *Administration and Policy in Mental Health*, *35*, 395-409. doi:10.1007/s10488-008-0182-y

DiPerna, J. C., Volpe, R. J., & Elliott, S. N. (2002). A model of academic enablers and elementary reading/language arts achievement. *School Psychology Review*, *31*, 298-312.

DuPaul, G. J., Volpe, R. J., Jitendra, A. K., Lutz, J. G., Lorah, K. S., & Gruber, R. (2004). Elementary school students with AD/HD: Predictors of academic achievement. *Journal of School Psychology*, *42*, 285-301. doi:10.1016/j.jsp.2004.05.001

Ensminger, M. E., & Slusarcick, A. L. (1992). Paths to high school graduation or dropout: A longitudinal study of a first-grade cohort. *Sociology of Education*, *65*, 95-113. doi:10.2307/2112677

Freudenberg, N., & Ruglis, J. (2007). Reframing school dropout as a public health issue. *Preventing Chronic Disease*, *4*(4), 1-11.

Gibbons, R. D., Bock, R. D., Hedeker, D., Weiss, D. J., Segawa, E., Bhaumik, D. K., …Stover, A. (2007). Full-information item bifactor analysis of graded response data. *Applied Psychological Measurement*, *31*, 4-19. doi:10.1177/0146621606289485

Gibbons, R. D., & Hedeker, D. (1992). Full-information item bi-factor analysis. *Psychometrika*, *57*, 423-436. doi:10.1007/BF02295430

Ginsburg-Block, M. D., & Fantuzzo, J. W. (1998). An evaluation of the relative effectiveness of NCTM standards-based interventions for low-achieving urban elementary students. *Journal of Educational Psychology*, *90*, 560-569. doi:10.1037/0022-0663.90.3.560

Glisson, C. (2002). The organizational context of children's mental health services. *Clinical Child and Family Psychology Review*, *5*, 233-253. doi:10.1023/A:1020972906177

Glisson, C., & Durick, M. (1988). Predictors of job satisfaction and organizational commitment in human service organizations. *Administrative Science Quarterly*, *33*(1), 61-81. doi:10.2307/2392855

Glisson, C., & Green, P. (2011). Organizational climate, services, and outcomes in child welfare systems. *Child Abuse & Neglect*, *35*, 582-591. doi:10.1016/j.chiabu.2011.04.009

Glisson, C., Green, P., & Williams, N. J. (2012). Assessing the organizational social context (OSC) of child welfare systems: Implications for research and practice. *Child Abuse & Neglect*, *36,* 621-632. doi:10.1016/j.chiabu.2012.06.002

Glisson, C., & Hemmelgarn, A. (1998). The effects of organizational climate and interorganizational coordination on the quality and outcomes of children's service systems. *Child Abuse & Neglect*, *22*, 401-421. doi:10.1016/S0145-2134(98)00005-2

Glisson, C., Hemmelgarn, A., Green, P., & Williams, N. J. (2013). Randomized trial of the availability, responsiveness, and continuity (ARC) organizational intervention for improving youth outcomes in community mental health programs. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*, 493-500. doi:10.1016/j.jaac.2013.02.005

Glisson, C., Landsverk, J., Schoenwald, S., Kelleher, K., Hoagwood, K. E., Mayberg, S., . . . The Research Network on Youth Mental Health. (2008). Assessing the organizational social context (OSC) of mental health services: Implications for research and practice. *Administration and Policy in Mental Health*, *35,* 98-113. doi:10.1007/s10488-007-0148-5

Glisson, C., Schoenwald, S. K., Kelleher, K., Landsverk, J., Hoagwood, K. E., Mayberg, S., . . . The Research Network on Youth Mental Health. (2008). Therapist turnover and new program sustainability in mental health clinics as a function of organizational culture, climate, and service structure. *Administration and Policy in Mental Health*, *35*, 124-133. doi:10.1007/s10488-007-0152-9

Glisson, C. A. (1978). Dependence of technological routinization on structural variables in human service organizations. *Administrative Science Quarterly*, *23*, 383-395. doi:10.2307/2392416

Gresham, F. M., & Elliott, S. N. (2008). *Social skills improvement system: Rating scales manual*. Minneapolis, MN: Pearson

Hakanen, J. J., Bakker, A. B., & Schaufeli, W. B. (2006). Burnout and work engagement among teachers. *Journal of School Psychology*, *43*, 495-513. doi:10.1016/j.jsp.2005.11.001

Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis*. New York, NY: Wiley.

Ingersoll, R. M., & Strong, M. (2011). The impact of induction and mentoring programs for beginning teachers: A critical review of the research. *Review of Educational Research*, *81*, 201-233. doi:10.3102/0034654311403323

James, L. A., & James, L. R. (1989). Integrating work environment perceptions: Explorations into the measurement of meaning. *Journal of Applied Psychology*, *74*, 739-751. doi:10.1037/0021-9010.74.5.739

Jones, A., & James, L. (1979). Psychological climate—Dimensions and relationships of individual and aggregated work-environment perceptions. *Organizational Behavior and Human Performance*, *23*, 201-250. doi:10.1016/0030-5073(79)90056-4

Joyce, W. F., & Slocum, J. W. (1984). Collective climate: Agreement as a basis for defining aggregate climates in organizations. *Academy of Management Journal*, *27*, 721-742. doi:10.2307/255875

Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2008). A multilevel study of predictors of student perceptions of school climate: The effect of classroom-level factors. *Journal of Educational Psychology*, *100*, 96-104. doi:10.1037/0022-0663.100.1.96

Kyriacou, C. (1987). Teacher stress and burnout: An international review. *Educational Research*, *29*, 146-152. doi:10.1080/0013188870290207

Kyriacou, C. (2001). Teacher stress: Directions for future research. *Educational Review*, *53*, 27-35. doi:10.1080/00131910120033628

Ladd, G. W., & Price, J. M. (1987). Predicting children's social and school adjustment following the transition from preschool to kindergarten. *Child Development*, *58*, 1168-1189. doi:10.2307/1130613

Levitt, J. M., Saka, N., Romanelli, L. H., & Hoagwood, K. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology*, *45*, 163-191. doi:10.1016/j.jsp.2006.11.005

Libbey, H. P. (2004). Measuring student relationships to school: Attachment, bonding, connectedness, and engagement. *Journal of School Health*, *74*, 274-283. doi:10.1111/j.1746-1561.2004.tb08284.x

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, *52*, 397-422. doi:10.1146/annurev.psych.52.1.397

Matsumura, L. C., Slater, S. C., & Crosson, A. (2008). Classroom climate, rigorous instruction and curriculum, and students' interactions in urban middle schools. *The Elementary School Journal*, *108*, 293-312. doi:10.1086/528973

The Metropolitan Area Child Study Research Group. (2007). Changing the way children "think" about aggression: Social-cognitive effects of a preventive intervention. *Journal of Consulting and Clinical Psychology*, *75*, 160-167. doi:10.1037/0022-006X.75.1.160

Murali, V., & Oyebode, F. (2004). Poverty, social inequality, and mental health. *Advances in Psychiatric Treatment*, *10*, 216-224. doi:10.1192/apt.10.3.216

Olin, S. S., Williams, N., Pollock, M., Armusewicz, K., Kutash, K., Glisson, C., & Hoagwood, K. E. (2014). Quality indicators for family support services and their relationship to organizational social context. *Administration and Policy in Mental Health and Mental Health Services Research*, *41*, 43-54. doi:10.1007/s10488-013-0499-z

Ota, K. R., & DuPaul, G. J. (2002). Task engagement and mathematics performance in children with attention-deficit hyperactivity disorder: Effects of supplemental computer instruction. *School Psychology Quarterly*, *17*, 242-257. doi:10.1521/scpq.17.3.242.20881

Schoenwald, S. K., Garland, A. F., Chapman, J. E., Frazier, S. L., Sheidow, A. J., & Southam-Gerow, M. A. (2011). Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*, 32-43. doi:10.1007/s10488-010-0321-0

Shapiro, E. S. (2004). *Academic skills problems: Direct assessment and intervention* (3rd ed.). New York, NY: Guildford Press.

Shernoff, E., Mehta, T., Atkins, M., Torf, R., & Spencer, J. (2011). A qualitative study of the sources and impact of stress among urban teachers. *School Mental Health*, *3*, 59-69. doi:10.1007/s12310-011-9051-z

Stringfield, S. (1994). A model of elementary school effects. In D. Reynolds, B. P. M. Creemers, P. S. Nesselrodt, E. C. Schaffer, S. Stringfield, & C. Teddlie (Eds.), *Advances in school effectiveness research and practice* (pp. 153-188). Tarrytown, NY: Elsevier.

Trickett, E. J., & Moos, R. H. (1973). Social environment of junior high and high school classrooms. *Journal of Educational Psychology*, *65*, 93-102. doi:10.1037/h0034823

Trickett, E. J., & Quinlan, D. M. (1979). Three domains of classroom environment: Factor analysis of the Classroom Environment Scale. *American Journal of Community Psychology*, *7*, 279-291. doi:10.1007/BF00890692

Wang, M. (2009). School climate support for behavioral and psychological adjustment: Testing the mediating effect of social competence. *School Psychology Quarterly*, *24*, 240-251. doi:10.1037/a0017999

Watts, J., & Robertson, N. (2011). Burnout in university teaching staff: A systematic literature review. *Educational Research*, *53*, 33-50. doi:10.1080/00131881.2011.552235

Withall, J. G. (1949). The development of a technique for the measurement of social-emotional climate in the classroom. *Journal of Experimental Education*, *17*, 347-361.

Withall, J. G. (1951). The development of the climate index. *The Journal of Educational Research*, *45*, 93-100.

Wyman, P. A., Brown, C. H., Inman, J., Cross, W., Schmeelk-Cone, K., Guo, J., & Pena, J. B. (2008). Randomized trial of a gatekeeper program for suicide prevention: 1-year impact on secondary school staff. *Journal of Consulting and Clinical Psychology*, *76*, 104-115. doi:10.1037/0022-006X.76.1.104

Zullig, K. J., Huebner, E. S., & Patton, J. M. (2011). Relationships among school climate domains and school satisfaction. *Psychology in the Schools*, *48*, 133-145. doi:10.1002/pits.20532

Author Biographies

Stacy L. Frazier is an associate professor in the Center for Children and Families at Florida International University and the director of Clinical Training for the Clinical Science Program in Child and Adolescent Psychology. She examines mental health promotion for families living in urban poverty through capacity building in the natural neighborhood settings that support them.

Tara G. Mehta is an assistant professor at the Institute of Juvenile Research in the Department of Psychiatry at the University of Illinois at Chicago. Her research includes the development and study of effective and efficient systems to support evidence-based practices in community mental health service settings for families living in urban poverty.

Marc S. Atkins is a professor of psychiatry and psychology and the director of the Institute for Juvenile Research at the University of Illinois at Chicago. He has a long-standing interest in the development of effective mental health services for children and families living in high-poverty urban communities, and is an active researcher in the areas of childhood attention deficit hyperactivity disorder (ADHD) and aggression, and school-based mental health services for children and families.

Charles Glisson is Chancellor's professor and University Distinguished professor in the Children's Mental Health Services Research Center at the University of Tennessee, Knoxville. His work examines organizational strategies for assessing and improving mental health, child welfare, and other youth and family serving organizations.

Philip D. Green is a research scientist in the Children's Mental Health Services Research Center in the Department of Social Work at the University of Tennessee in Knoxville. He oversees research design and data collection methods in addition to scale design and analysis. He examines organizational level climate and culture dimensions as predictors of staff and client outcomes.

Robert D. Gibbons is a professor of biostatistics, in the Departments of Health Studies, Medicine, and Psychiatry, and he directs the Center for Health Statistics at the University of Chicago. He works in a variety of statistical areas including analysis of longitudinal data and multidimensional item response theory (IRT). Results of his statistical work are widely applied in mental health research.

Jong Bae Kim is a statistical analyst at the Center for Health Statistics at the University of Chicago. He works on a variety of research projects in the center, and he is currently involved in the development of a computerized adaptive test for depression.

Jason E. Chapman is an associate professor in the Family Services Research Center in the Department of Psychiatry and Behavioral Sciences at the Medical University of

South Carolina. His work focuses on the application of mixed-effects regression models and IRT-based measurement models in the context on research for adolescents with substance abuse and delinquency.

Sonja K. Schoenwald is a professor of psychiatry and behavioral sciences at the Medical University of South Carolina. In collaboration with investigators in federally and foundation funded research centers and projects, she conducts research on the development and implementation of effective community-based psychosocial treatments for youth and families who come to the attention of juvenile justice, child welfare, and education sector authorities.

Grace Cua is a doctoral student in educational psychology in the College of Education at the University of Illinois at Chicago. Her research examines factors related to promoting teacher effectiveness, especially for those working in high poverty schools.

Robert R. Ogle is a doctoral student in clinical psychology at Florida International University. His research focuses on identifying broad-spectrum, cost-efficient, effective, and disseminable intervention and prevention tools for children and adolescents, with a particular focus on educational achievement in economically disadvantaged communities.