The Meriden School Climate Survey-Student Version: Reliability and Validity

Nicholas A. Gage

University of Florida

Alvin Larson

Meriden Public School

Sandra Chafouleas

University of Connecticut

Abstract

School climate has been linked with myriad positive student outcomes and the measurement of school climate is widely advocated at the national and state level. However, districts have little guidance about how to define and measure school climate. This study examines the psychometric properties of a district-developed school climate measure that was created in response to state policy pressure and an interest in student perceptions of school climate. The Meriden School Climate Survey-Student Version, a 38-item scale, was found to be valid and reliable based on exploratory and confirmatory factor analysis, test-retest reliability, and comprehensive assessment of internal consistency. Overall, results suggest that the MSCS-SV is a sound measure of student perceptions of a broad school climate construct.

The Meriden School Climate Survey–Student Version: Reliability and Validity

Schools are tasked with creating positive school climates to increase the social and academic performance of all students (Cohen, McCabe, Michelli, & Pickeral, 2009). Although there is no consensus definition of school climate (Thapa, Cohen, Guffey, & Higgins-D'Alessandro, in press), it has generally been described as the social and environmental conditions of a school. Positive school climate refers to schools that are safe (emotionally and physically), engaged, collaborative (between teachers, students and parents), and respectful (National School Council, 2007). Student perceptions of positive school climate have been linked to a number of positive academic, social, and behavioral outcomes, including academic achievement (Brand, Felner, Shim, Seitsinger, & Dumas, 2003; Griffith, 1999); attitudes and motives in school (Battistich, Solomon, Kim, Watson, & Schaps, 1995); and increased attendance (Brand et al., 2003; Welsh, 2000). Further, student perception of positive school climate have been linked to decreased negative outcomes such as student delinquency (Gottfredson, Gottfredson, Payne, & Gottfredson, 2005; Welsh, 2000), use of illegal substances (Brand et al., 2003), bullying (Nansel et al., 2001), victimization (Gottfredson et al., 2005; Welsh, 2000), depression (Brand et al., 2003; Way, Reddy, & Rhodes, 2007), and general behavior problems (Battistich & Horn, 1997; Kuperminc, Leadbeater, & Blatt, 2001; Welsh, 2000). Taken together, these positive associations suggest that evaluations of school climate may provide useful to school leaders in decision making regarding interventions and supports.

Understanding and examining school climate is imperative given the significant amount of research identifying the relationship between positive school climate and positive student outcomes, and has been widely advocated by the federal department of

education and many state departments of education (Piscatelli & Lee, 2011). Unfortunately, there is no consensus on the definition or the measurement of school climate, particularly at the student-level, to guide school districts (Cohen, Pickeral, & McCloskey, 2009). As a result, school districts must rely on state recommendations, consultant recommendations, or develop their own measures of school climate. The purpose of this paper is to review work undertaken to evaluate the technical adequacy of a school climate measure developed by a district in response to the lack of an appropriate measure yet increased pressures within state policy to assess school climate. In particular, district-level leadership requested an assessment of student perceptions of school climate as part of their district improvement plan, and thus, a decision was made to invest in developing a new measure. Leaders were interested in student perceptions in order to collect actionable data that teachers and school administrators could use to improve school climate. This study is the result of a partnership between the district and a local university research center to evaluate the psychometric properties of the Meriden School Climate Survey-Student Version (MSCS-SV).

Defining School Climate

School climate is a multi-dimensional construct with myriad definitions and critical features described across the literature. In an early review of school climate research, Anderson (1982) found that most researchers developed their own definitions, which they "verified intuitively rather than empirically" (p. 369). For example, Halpin and Coft (1963) described climate as the personality of an organization, whereas Tagiuri (1968) described dimensions of an environment, including ecology, milieu, social systems, and culture. Recent reviews have updated and refined the definition and

conceptualization of school climate as the quality and character of school life (Cohen et al., 2009), involving the patterns of people's experiences of school life, reflecting norms, goals, values, interpersonal relationships, teaching and learning practices, and organizational structures (Thapa et al., in press). Researchers have identified a number of domains or dimensions of school climate, including (a) safety (e.g., rules and norms, perceived safety, physical safety), (b) social relationships (e.g., teacher-student relationships, peer relationships, social support), (c) teaching and learning (e.g., social/emotional learning, support for academic achievement), (d) institutional environment (e.g., physical surroundings, resources), and (e) school connectedness (e.g., students' and parents' feelings about school, parent support of student academic success, student enthusiasm) (Cohen et al., 2009; Thapa et al., in press; Zullig, Koopman, Patton, & Ubbes, 2010).

Further complicating clear definition of school climate is the unit of analysis and interpretation. Although school climate is a school-level construct, the experience and interpretation of that "climate" may vary by unit of measurement (i.e., student, teacher, parent, etc.) and may be moderated by any number of variables, including student ethnicity, teacher experience, or geographical location. Students, teachers, administrators, and parents all experience school life differently. All four stakeholders have an active role in school life. Students are the primary focus and consumers of education; teachers are the agents delivering instruction and contents, and employees of the school. Administrators are building leaders setting the tone and enforcing rule compliance, wheras parents are collaborators and supporters of educational outcomes produced by the schools. Overall, all four play a critical role; therefore, definitions should be either

flexible enough to include all four stakeholders, regardless of unit of analysis, or specific to only one stakeholder group.

The MSCS-SV was informed by the definitions developed by Cohen et al. (2009) and Thapa et al. (in press), and incorporates the district interest in home-school connectedness. The MSCS-SV was developed as a measure of the quality and character of school life as experienced by students, including school connectedness, school safety, social and emotional well-being, including social interactions with adults and peers.

Measuring School Climate

A bevy of school climate measures have been developed, going back over half a century ago to Halpin and Croft's (1963) *Organizational Climate Descriptive Questionnaire*. As noted previously, there is no agreement about (a) the definition of school climate, (b) the dimensions of schools climate, and (c) the unit of measurement and analysis (e.g., teacher, student, parent). As such, extant measures were developed with different definitions of school climate and focused on different units of measurement. Since the proliferation of school climate measures began, three comprehensive reviews have been undertaken regarding measures of school climate in order to identify existing measures of school climate and report features and psychometric properties of existing measures of school climate.

Anderson (1982) reviewed existing school climate measures with established psychometric properties, identifying nine. Students were the unit of analysis in five of the nine measures, but all assessed different aspects of school climate and all were based on a different definition of school climate. For example, the *Pupil Control Behavior* examined students' perceptions of teacher's orientation towards students, focusing on a custodial

continuum, whereas the *Quality of School Life Scale* assesses students' attitudes toward school, including relationships with authority figures. Although comprehensive, Anderson's review was conducted over 30 years ago and most of the measures are no longer available.

In a more recent review, Gangi's (2010) dissertation study examined existing school climate measures to identify the most valid and reliable available in the literature. Following a comprehensive literature search, 102 measures of school climate were identified. The study only included measures that assessed (a) teacher perceptions of school climate and (b) available items for review, resulting in three studies. This dissertation study reviewed the *Comprehensive School Climate Inventory* (CSCI), *Tennessee School Climate Inventory-Revised*, and the *Western Alliance for the Assessment of School Climate* (WAASC), finding that, although all three were promising, the WAASC was the most promising based on availability of quantitative evidence of validity, reliability, and the norming sample (e.g., sample size and date of study). However, the WAASC did not demonstrate test-retest reliability, did not utilize a nationally representative sample, and does not have a technical manual.

Most recently, Clifford, Menon, Gangi, Condon, and Hornung (2012) conducted a comprehensive review of school climate measures to identify potentially valid and reliable measures for principal evaluation. Their review identified 125 surveys, but only reviewed measures that reported psychometric properties and were completed by teachers or administrators. Thirteen measures met criteria and were assessed for reliability and validity. Of those, two included student perceptions along with teacher perceptions and three measures included teachers', students' and parents' perceptions. The report does not

rank order the measures, but instead provides a table for school administrators to use as a starting point for identifying a measure of school climate for principal evaluation. However, none of the measures focus exclusively on student perceptions.

Taken together, it is clear from these reviews that (a) there are over 100 measures of school climate in the literature, (b) only a few of the measures report psychometric properties, and (c) even fewer are completed by students and valid and reliable. Although not included in the reviews above, a recently developed measure of student perceptions of school climate necessitates mention. Bear, Gaskin, Blank, and Chen (2011) developed the *Delaware School Climate Survey-Student* (DSCS) from a social-ecological perspective (Bandura, 1986, 1997) to assess a bi-dimensional framework of school climate (Stockard & Mayberry, 1992), which posits that school climate consists of social action (i.e., the presence of caring, understanding, concern, and respect among students, staff, and teachers) and social order (i.e., structure to reduce behavior problems and promote safety). More importantly, the DSCS targets bullying as a central construct, meaning they measure school climate to measure bullying and student aggression in schools. Unlike all previously reviewed measures, the DSCS can be accessed and used by school districts, but it's focus on bullying and aggression do not provide broader insight into the construct of school climate (e.g., home/school connectedness), which was the goal of the MSCS-SV.

Limitations of Existing School Climate Measures

Although a handful of empirically validated school climate measures exist, limitations substantiate the need for further development of new and contemporary measures. First, many measures have not proven to be widely used, most likely because

they do not fulfill a contemporary definition of school climate. For example, the measures identified in the review by Anderson (1982) were developed over thirty years ago and have not been widely used this century. Second, the 13 measures identified by Clifford et al. (2012), including the three by Gangi (2010), are completed primarily by teachers and were identified as potential evaluation tools for principal performance. Teacher perception of school climate is more akin to organizational climate research because it assesses employee perceptions of their working environment, including working conditions (i.e., safety). Although teacher perceptions are important generally, student perceptions are relevant for developing interventions at the school- or student-level to increase student perceptions of school climate, which have been directly linked to positive student outcomes. Taken together, to address the broad and contemporary definition of school climate adopted by the district, either multiple measures would have been required, an inefficient use of resources, or a new measure needed to be developed and evaluated.

Purpose

A large school district in Connecticut developed a school climate measure, the Meriden School Climate Survey-Student Version, in response to (a) state policy pressure to assess school climate, (b) no widely advocated measure of student perceptions school climate, and (c) a district-level interest in student perceptions of school environments. In this study, an evaluation of the psychometric properties of the Meriden School Climate Survey-Student Version is provided. Specific research aims were to confirm construct validity, test-retest validity, and internal consistency of the measure. An additional

exploratory aim was to identify similarities and differences in response patterns by student characteristics.

Methods

Sample

Meriden School District is comprised of eight elementary schools, two middle schools and two high schools, with a total enrollment of ~8,900 students. The district serves a diverse student population, with 62% receiving free or reduced lunch, 11% not fluent in English, and 13.4% receiving special education services. Total minority population in the district is approximately 61%, including 45% White/Hispanic, 14% African-American, 3% Asian, and 0.3% Native American students.

A total of 3,504 students completed the MSCS-SV in June of 2011-2012. Forty-eight percent of the students were in elementary school, 31% were in middle school and 21% were in high school. Across grades, the largest percentage of students was in fourth grade (17%) and the smallest was 12th grade (3%). Fifty percent of students were boys, 39% were White/Non-Hispanic, 43% were White/Hispanic, 15% were Black, and 2% were Asian. The majority of students (64%) received free or reduced lunch. The district classified 9%of students as English Language Learners (ELL) and 11% received special education services, with the largest group receiving services for Learning Disabilities (6%).

Instrument Development

The district's Research and Evaluation Specialist developed the MSCS-SV from a review of existing school climate measures the National School Climate Standards (National School Climate Council, 2007). All items were vetted by a team of district

staff, including teachers and administrators. The original pilot measure consisted of 47 items broadly based on three core features of school climate: (a) norms, values, and expectations that support people feeling socially, emotionally and physically safe; (b) students, families and educators working together to develop, live and contribute to a shared school vision; and (c) educators model and nurture attitudes that emphasize the benefits and satisfaction gained from learning (National School Climate Council, 2007). All items use a 5-point Likert scale ranging from "strongly disagree" to "strongly agree". Internal consistency of the pilot measure with all 47-items was very high ($\alpha = .93$).

The MSCS-SV is delivered online to students in grade 3 to 12 and is completed in the fall (September 1 to October 11) and spring (April 22 to June 15) of each school year. The survey takes approximately 15 minutes to complete. Data for this study were completed in October and June of the 2011-2012 school year. All primary analyses were conducted using the MSCS-SV collected in June. We focus on the end-of-the-year (EOY) collection period because perceptions of school climate at EOY reflect student experiences throughout the year.

Data Analysis

Missing data. Across all 47 items, missing values accounted for 1.6%. The item "*During the past few months, I have hit, pushed or spread mean rumors at the bus stop or on the bus*" had the most missing data (3.1%), whereas "*I know the school rules*" had the least amount of missing data (0.3%). To avoid deletion of students, we used the Expectation-Maximization (EM) multiple imputation procedure to impute missing values for each variable. The EM algorithm attempts to find a value for theta which maximizes g(y | theta) given an observed y, and does so by making use of the associated family f(x,

y | theta) (Dempster, Laird, & Rubin, 1977). We treated all 47-items as a single family because the measure was designed to measure a single construct. All EM imputation was conducted using SPSS 19.0 Missing Values Add-in software.

Exploratory and confirmatory factor analysis. A goal of this study was to examine and refine the factor structure of the MSCS-SV. First, we randomly split the participant sample into two equivalent groups of 1,752 students using SPSS 19.0 random sample generator. The first sample was used for exploratory factor analysis (EFA) and the second sample was used for confirmatory factor analysis (CFA). The sample size was sufficient to meet the "sample size to number of items" recommendations in factor analysis (i.e., sample size should be approximately 10 times the number of items in the survey) (Netemeyer, Bearden, & Sharma, 2003). The EFA was conducted in SPSS 19.0 using Principal Component Analysis with Varimax rotation. Items were retained if the factor loading was .40 or greater and did not cross-load on another factor above .35. Items that (a) did not sufficiently load on a factor or (b) cross-loaded were excluded from subsequent analyses.

The CFA analysis was conducted in AMOS 19 using Maximum Likelihood estimation and based on the factor structure resulting from the EFA. Model fit was assessed based on sufficient comparative fit index (CFI) and Tucker-Lewis Index (TLI) values defined as > .90, standardized root mean squared residual (SRMR) ≤ .09, and root mean squared error of approximation (RMSEA) ≤ .06. Hu and Bentler's (1999) joint criteria specify that CFI and TLI values should be ≥ .96, however, research suggests that CFI and TLI model fit decreases as the number of variables and parameters increase (Kenny & McCoach, 2003). Therefore, we used a more liberal criterion (i.e., CFI and TLI

≤ .90), which aligns with other school climate research (see Furlong et al., 2005). To improve model fit, we used model modification indices and added covariance parameters between item-level error terms within factors (i.e., no cross-factor covariance parameters at the item level) (Kline, 2005). Once the final model was identified, the parameter estimates and model fit was reassessed using the full October sample of 3,868 students to confirm parameter stability. Sixty-eight percent of the sample (2,630 students) completed the MSCS-SV in both October and June.

Reliability. Following the EFA and CFA confirmation of the MSCS-SV factor structure, we examined item-level test-rest reliability from October to June using results from the 2,630 students that completed the MSCS-SV on both occasions. Test-retest reliability was calculated using Pearson correlation coefficients in SPSS 19.0. We also calculated reliability statistics for the full survey (all retained items) and for each factor. Typically, applied researchers report Cronbach's alpha (α) as a single point estimate of reliability. Recent research supports the calculation of a standard error of alpha and reporting a 95% confidence interval (Duhachek & Iacobuci, 2004), which we calculated using Microsoft Excel. In addition to alpha, we calculated three other reliability statistics. Sijtsma (2009) argues that alpha is a very limited and over-used statistic that only estimates a lower bound of reliability and, in many cases, is a gross underestimate. As a result, Revelle and Zinbarg (2009) recommend reporting McDonald's ω, Guttman's λ_4, and Revelle's β.

McDonald's (1999) ω is an estimate of the general factor saturation of a test and is calculated using a CFA model, summing the factor loadings and dividing by the variance in the scale scores obtained by summing the *k* indicators comprising the scale. The larger

this variance ratio (ω) is, the more accurately we can predict an individual's relative standing on the latent variable common to all the scale's indicators based on his or her observed scale score (Revelle, 2008). Essentially, ω is the proportion of common variance explained by the latent factor. Guttman (1945) developed a series of lower bound reliabilities (λ_1 to λ_6), with λ_4 representing the maximum lower-bound estimate of reliability. Revelle's (1979) β is the lowest possible split-half coefficient possible. Zibarg, Revelle, Yovel, & Li (2005) empirically demonstrated that ω is a better estimate of reliability, particularly when the researcher is interested in the proportion of scale variance due to a general factor. Therefore, following the recommendations of Revelle and Zinbarg (2009), we reported the four reliability statistics for the full scale and each identified factor. Reliability was calculated using the *psych* package (Revelle, 2008) for R (R Development Core Team, 2008). Reliability was judged based on the following criteria: > .9 – Excellent, > .8 – Good, > .7 – Acceptable, > .6 – Questionable, > .5 – Poor, and, < .5 – Unacceptable (George & Mallery, 2003, p. 231) with the caveat that those values are for alpha and not the other statistics.

Descriptive Characteristics of Students and Schools by Factor and Scale

Once the final factor structure was confirmed and the measure met reliability criteria, student- and school-level descriptive characteristics were examined by factor. Standardized z-scores ($M = 0.0$, $SD = 1.0$) for each factor were calculated. We used standard scores instead of summed total scores because item influence on the factor scores is weighted by factor loadings while summed scores treat all items as equal. In addition to factor scores for each of the identified factors, we calculated a factor score for the full scale using a principal component factor analysis with Varimax rotation, but

constrained the number of possible factors to one. The factor scores were retained, resulting in a full-scale factor score for each student. Differences between groups were calculated using Hedges' *g* effect sizes and assessed using What Works Clearinghouse equivalence standards of .25 standard deviations. Differences between groups greater than .25 standard deviations were considered "substantively important".

Results

Comparing EFA and CFA Samples

Two samples of students were artificially created by randomly splitting the full June 2012 MSCS-SV sample of 3,504 into two equal groups. In order to establish equivalence the two random samples were compared across available demographic characteristics. The percentage of students by demographic characteristic and EFA and CFA sample are presented in Table 1. To statistically assess equivalence, we calculated Hedges' *g* effect sizes for dichotomous variables for each characteristic following conventions outlined by What Works Clearinghouse (CITE). Equivalence is defined as differences less than 0.25 of a standard deviation. Equivalence was established for all group comparisons except for Asian students ($g = 0.26$). Based on the small number of Asian students in the sample ($n = 70$), it was determined that nonequivalence on this characteristic is not threatening to the instrument development process.

In addition to calculating effect size comparisons between all characteristics, we calculated chi-square statistics to examine overall group equivalence. The EFA and CFA samples did not differ on grade-level, χ^2 ($df = 9$) = 8.25, $p = .51$; gender, χ^2 ($df = 2$) = 0.37, $p = .83$; ethnicity, χ^2 ($df = 4$) = 6.52, $p = .16$; lunch status, χ^2 ($df = 3$) = 0.42, $p = .94$; EL status, χ^2 ($df = 2$) = 2.06, $p = .36$; and disability status, χ^2 ($df = 9$) = 14.18, $p = .12$.

Exploratory Analysis

The EFA analysis was conducted in two iterations, with adjustments made based on recommended factor identification procedures (eigenvalues greater than 1.0 and visual inspection of the Scree plot; CITE) and the a priori analysis criteria established for this study (i.e., factor loading was .40 or greater and no cross-loadings above .35). First, we examined assumptions. The Kaiser-Meyer-Olkin measure of sampling adequacy was .92, above the recommended value of .6, and Bartlett's test of sphericity was significant (χ^2 (666) = 24090.6, $p < .00$). Using all 47 items, the first EFA with principal component analysis resulted in seven factors meeting recommended criteria but 10 items did not meet the a priori criteria (i.e., factor loading .40 or greater and no cross-loading above .35). Nine of the ten items were then removed from the analysis. Item 2, "*I feel safe at school*" cross-loaded on both factor 1 and factor 2, but because the item was deemed central to the construct "school climate", assigned decision was made to retain it and assign the item to the factor with the highest loading (Factor 1). The following items were removed based on a lack of model fit: 1. *My teachers want me to work hard and do well*, 4. *I know the school rules*, 6. *This school wants all students to do their very best*, 13. *At my school, there is a teacher or other adult who always wants me to do my best*, 15. *I try to understand how other students feel*, 16. *At home, I have a parent or other adult who talks with me about my problems*, 17. *Other students in this school are polite and listen to what I say*, 20. *In class, I try to understand other students who disagree with me*, and 29. *I do my homework on time*.

The new 38-item MSCS-SV was re-analyzed with the same sample to re-establish the seven-factor structure and accurate statistics. The final factor loadings,

communalities, eigenvalues, and percentage of variance by factor are presented in Table 2. Each factor was defined by the second author and confirmed for face validity by the first and third author. We defined factor 1 as Adult Support at School (11 items), factor 2 as School Safety (7 items), factor 3 as Respect for Differences (5 items), factor 4 as Adult Support at Home (4 items), factor 5 as Academic Support at Home (4 items), factor 6 as Peer Support (4 items) and factor 7 as Aggression Towards Others (3 items). Factor loadings ranged from .42 to .78 suggesting that the items were highly saturated to each construct. Overall, the seven-factor structure explained 55.3% of the variance.

Confirmatory Analysis

Using structural equation modeling methods, a fully correlated 7-factor model based on the EFA results was then fit using the CFA sample of students. The first model fit statistics did not meet all a priori criteria for model fit: $\chi^2 = 3948.3$ ($df = 608$, $p < .000$), CFI = .862, TLI = .849, RMSEA = .056, and SRMR = .067. The RMSEA and SRMR met criteria, but the CFI and TLI did not. The χ^2 value as significant, but expected based on the number of parameters and sample size (Hu & Bentler, 1999). In order to increase model fit, we followed the model fit modification recommendations from AMOS and used the procedures outlined by Kline (2005). Once all within factor covariances were included (contact first author for a complete list of covariance parameters), our 7-factor model met our model fit criteria: $\chi^2 = 2734.3$ ($df = 585$, $p < .000$), CFI = .911, TLI = .900, RMSEA = .046, and SRMR = .062. All factor loadings and covariance estimates were statistically significant. Correlations between the seven factors ranged from .17 between factors 5 (Academic Support at Home) and 6 to .60 between factors 2 (School Safety) and 3 (Respect for Differences). To confirm model fit, we conducted a series of model

reductions, removing a single factor at a time in a systematic fashion to determine whether model fit could be improved by removing one or more factors. Model fit was consistently worse than the 7-factor model, with no CFI and TLI statistics above .900 across all reduced factor models.

To further confirm the validity of the 7-factor structure, the CFA model was replicated with full October 2011 sample of students. Of the 3,868 students that completed the October MSCS-SV, 68% also completed the June 2012 MSCS-SV. The replicated model met most of the a priori model fit criteria: $\chi^2 = 5862.4$ ($df = 585$, $p < .000$), CFI = .900, TLI = .880, RMSEA = .048, and SRMR = .069.

Reliability

Following confirmation of the 38-items and 7-factors, reliability statistics were examined. First, we examined the test-retest reliability at the item level. As noted, 2,630 students completed both the October and the June MSCS-SV. The range of correlation coefficients was .39 to .45 in Adult Support at School, .35 to .45 in School Safety, .36 to .46 in Respect Differences, .32 to .35 in Academic Support at Home, .31 to .35 in Adult Support at Home, .23 to .38 in Peer Support, and .43 to .45 in Aggression Towards Others. All correlation coefficients were statistically significant at the $p < .000$. The smallest correlation (.23) was for the item "*During the past few months, I have hit, pushed or spread mean rumors at the bus stop or on the bus*", which was also the only item with a correlation below .30.

Reliability of each the full scale and each factor (latent construct) was examined across four statistics. Reliability results are reported in Table 3. All but one alpha value was above .70, indicating acceptable reliability. The factor with an $\alpha < .70$ (Aggression

Towards Others) only had 3 items (27, 28, and 46) and those items were the most skewed compared to all other items (-2.18, -3.51, and -3.34 respectively). Upon review of the 95% confidence interval and standard errors of alpha, the Aggression Towards Others factor had the largest standard error and the upper limit of the confidence interval was >1. McDonald's ω values were all >.70 and indicated acceptable reliability. Revelle's Beta results were also acceptable, with lowest possible split-half correlation found in the full scale, which is due to the large number of items in the scale. Guttman's Lamba4 was acceptable across all items, but again, the Aggression Towards Others factor was <.70, indicating questionable reliability.

Characteristic Differences

Standard scores were calculated for each student based for the full scale and the 7-factor model. Results are presented in Table 4. Based on the full-scale results, students in elementary school report increasing positive school climate whereas students in middle and high school reported decreasing positive school climate. No substantive differences were found on the full scale for gender, race/ethnicity, Free and reduced lunch status, ELL status, and special education status. Students receiving special education services reported more negative school climate, but the difference was not >.25.

Discussion

In response to policy pressure and interest by school leaders in student perceptions of school climate, a school district developed a measure of school climate. A decision was made to base the school climate survey on a broad and contemporary definition of school climate, defined as the quality and character of school life as experienced by students, including school connectedness, school safety, social and

emotional well-being, and social interactions with adults and peers. The district partnered with a university-based research center to evaluate the psychometric properties of the developed measure. Based on the analysis, the MSCS-SV was modified from 42-items to 38-items, with seven factors, or domains, including (a) Adult Support at School, (b) School Safety, (c) Respect for Differences, (d) Adult Support at Home, (e) Academic Support at Home, (f) Peer Support, and (f) Aggression Towards Others.

The EFA analysis identified a seven-factor structure, but ten items did not meet a priori criteria of factor loadings .40 or greater and no cross-loadings 35. We removed nine of the ten items, retaining *I feel safe at school*. We retained the item because the item was directly related to a primary domain of interest, school safety. Keeping the item may have influenced our CFA results because of the cross loading. However, a decision was made to accept the trade-off because although the central focus of this study was validation of the survey measure, the district has a vested interest in both factor scores and individual item responses. One alternative would be to remove the item and report it as a single survey item, not an item representing a latent construct. For now, we have chosen to retain the item, with the caveat that future analysis may find that the item does not load and, in fact, significantly impacts model fit with future samples.

To confirm the seven-factor structure of the MSCS-SV, a CFA was conducted. Due to inconsistent model fit, modifications in the form of correlated errors for within factor items were added. We chose to modify the model using Kline's (2005) recommendations due to the match between the proposed factors and the theoretical model based on the school climate definition used to develop the survey. Once all modifications (i.e., within factor correlated errors) were made, the model fit indices met

all a priori defined fit criteria, confirming the validity of the seven-factor structure. To further confirm construct validity, the exact same model was recalculated, but with the Fall 2011 school climate data. The modification indices indicated consistent fit across the two models. Overall, confirmation of the seven-factor structure confirmed that the broader definition of school climate could be empirically validated. Unlike other contemporary measures with established psychometric properties (e.g., California School Climate and Safety Survey, Delaware School Climate Survey), the MSCS-SV measures a broader definition of school climate.

Once construct validity was established, reliability was examined using test-retest reliability for each item and internal consistency reliability of (a) the full-scale and (b) each of the seven individual factors. Retest correlations at the item-level were all statistically significant and moderately large (Cohen, 1992). We calculated multiple internal consistency reliability statistics because of the known problems with Cronbach's alpha (Sijtsma, 2009) and a desire for a comprehensive assessment of reliability because the measure will be used for intervention development. Reliability was above acceptable for all domains, with the exception of the *Aggression Towards Others* factor. We retained the factor because acceptable results were dependent upon which reliability statistic was used and because of the district's interest in the factor. Based on the review of existing measures, the MSCS-SV is the only school climate measure that asks students about their own behavior towards others. Interestingly, although aggression towards others might not be a construct typically considered in the broader school climate construct, inclusion of the construct was empirically supported and items may provide important contributions

to school information of students reporting aggression towards others in decisions about targets for intervention.

To extend the results beyond psychometric validation to descriptive information regarding findings, student and school characteristics were examined using full-scale and factor-based standard scores. Overall, student perceptions of school climate appear to decrease into middle and high school. This finding has not been fully explored empirically before, but may be due to environmental changes (i.e., moving from a single classroom, to switching among many classrooms) or developmental reasons. However, these are merely hypotheses that necessitate further investigation. Small differences were found between some student-level characteristics. For example, Black students were more likely to report less *Respect for Differences* than White/Non-Hispanic students, and students classified as ELL and those receiving special education services significantly reported lower levels of *Adult Support at Home*.

Limitations and Future Directions

A number of limitations necessitate highlighting. First, students were nested within schools and estimates did not account for this structure. Multi-level factor analysis was not conducted because students were clustered within only 10 schools, well below the minimum recommendation of approximately100 level 2 cluster units (Preacher, Zyphur, & Zhang, 2010). Discussions are currently underway to expand the use of the MSCS-SV in additional districts, providing opportunity for additional analysis using multi-level SEM approach, similar to that used by Bear et al. (2011). Relatedly, students in this study were all from the same New England school district, limiting generalization. Future studies should be conducted with students in a variety of locations

with even more diverse samples of students. Additionally, the model fit statistics met our a priori criteria, but did not meet Hu and Bentler's (1999) joint criteria. Additional models, based on either a reduction or expansion of the school climate definition, should be tested to identify a model with CFI and TLI statistics closer to 1.0. Models should also be extended beyond instrument validation to empirical models examining the relationship between the full-scale and factor scores and school and student outcomes. Descriptive results were presented in this study, but additional studies should employ SEM and other modeling procedures to identify interrelationships among factors and dependent variables. Future studies should also examine school-level intervention efforts on the full-scale and factor scores to determine whether or not intervention increases positive school climate.

Conclusions

Increasing positive school climate can have a significant impact on students' academic and behavioral outcomes. To assess school climate, districts must utilize valid and reliable measures. To date, a number of measures have been developed, but few contain a broad and contemporary definition of school climate. The results of this study indicate that the MSCS-SV is not only broad, but also valid and reliable. As such, the results of the MSCS-SV can be used for (a) measuring current student perceptions of school climate, (b) identifying areas to target for interventions and supports, and (c) re-assess using the MSCS-SV to identify change. More research is needed, but the results of this study are promising.

References

Anderson, C. S. (1982). The search for school climate: A review of the research. *Review of Educational Research*, *52*, 368-420. doi.10.3102/00346543052003368

Battistich, V., Solomon, D., Kim, D., Watson, M., & Schaps, E. (1995). Schools as communities, poverty levels of student populations, and students' attitudes, motives, and performance: A multilevel analysis. *American Educational Research Journal*, *32*, 627−658. doi:10.2307/1163326

Battistich, V., & Horn, A. (1997). The relationship between students' sense of their school as a community and their involvement in problem behaviors. *American Journal of Public Health*, *87*, 1997−2001. doi:10.2105/AJPH.87.12.1997

Bear, G. G., Gaskin, C., Blank, J., & Chen, F. F. (2011). Delaware School Climate Survey-Student: Its factor structure, concurrent validity, and reliability. *Journal of School Psychology*, *49*, 157-174. doi. 10.1016/j.jsp.2011.01.001

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

Clifford, M., Menon, R., Gangi, T., Condon, C., & Hornung (2012). *Measuring school climate for gauging principal performance: A review of the validity and reliability of publicly accessible measures*. Retrieved from http://www.air.org/files/school_climate2.pdf

Cohen, J. (1982). A power primer. *Psychological Bulletin*, *112*, 155-159.

Cohen, J., McCabe, E. M., Michelli, N. S., & Pickeral, T. (2009). School climate:

Research, policy, practice, and teacher education. *Teachers College Record*, *111*, 180-213.

Cohen, J., Pickeral, T., & McCloskey, M. (2009). Assessing school climate. *The Education Digest*, *74*, 45-48.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, *39*, 1-38.

Duhachek, A., & Iacobucci, D. (2004). Alpha's standard error (ASE): An accurate and precise confidence interval estimate. *Journal of Applied Psychology*, *89*, 792-808. doi: 10.1037/0021-9010.89.5.792

Furlong, M. J., Greif, J. L., Bates, M. P., Whipple, A. D., Jimenez, T. C., & Morrison, R. (2005). Development of the California School Climate and Safety Survey-Short Form. *Psychology in the Schools*, *42*, 137-149. doi: 10.1002/pits.20053

Gottfredson, G. D., Gottfredson, D. C., Payne, A. A., & Gottfredson, N. C. (2005). School climate predictor of school disorder: Results from a national study of delinquency in schools. *Journal of Crime and Delinquency*, *42*, 412-444. doi:10.1177/0022427804271931

Griffith, J. (1999). School climate as "social order" and "social action": A multi-level analysis of public elementary school student perceptions. *Social Psychology of Education*, *2*, 339−369.

Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, *10*, 255–282. doi:10.1007/BF02288892

Halpin, A. W., & Croft, D. B. (1963). *The organizational climate of schools*. Chicago, IL: University of Chicago Press.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55. doi.10.1080/10705519909540118

Kenny, D. A., & McCoach, D. B. (2003): Effect of the number of variables on measures of fit in structural equation modeling, *Structural Equation Modeling: A Multidisciplinary Journal*, *10*, 333-351. doi. 10.1207/S15328007SEM1003_1

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

Kuperminc, G. P., Leadbeater, B. J., & Blatt, S. J. (2001). School social climate and individual differences in vulnerability to psychopathology among middle school students. *Journal of School Psychology, 39*, 141-159. doi:10.1016/S0022-4405(01)00059-0

Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among US youth: Prevalence and associations with psychosocial adjustment. *Journal of the American Medical Association*, *285*, 2094−2100. doi:10.1001/jama.285.16.2094

National School Climate Council. (2007). *The School Climate Challenge: Narrowing the gap between school climate research and school climate policy, practice guide-lines and teacher education policy*. Retrieved from http://www.schoolclimate.org/climate/advocacy.php

Netwmeyer, R. G., Bearden, W. O., Sharma, S. (2003). *Scaling procedures: Issues and*

applications. Thousand Oaks, CA: Sage.

Piscatelli, J & Lee, C. (2011). State Policies on School Climate and Bully Prevention Efforts: Challenges and Opportunities for Deepening State Policy Support for Safe and Civil Schools. National School Climate Center (July 2011) Available at http://schoolclimate.org/climate/policy.php

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, *15*, 209-233. doi: 0.1037/a0020141

R Development Core Team. (2008). *R: A language and environment for statistical computing*. Vienna, Austria (ISBN 3-900051-07-0).

Revelle, W. (1979). Hierarchical cluster-analysis and the internal structure of tests. *Multivariate Behavioral Research*, *14*(1), 57–74. doi:10.1207/s15327906mbr1401_4

Revelle, W. (2008). *psych: Procedures for personality and psychological research* (R package version 1.0-51).

Revelle, W., & Zinbarg, R. E. (2009). Cofficient alpha, beta, omega, and the GLB: Comments on Sijtsma. *Psychometrika*, *74*, 145-154. doi: 10.1007/S11336-008-9102-Z

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alphs. *Psychometrika*, *74*, 107-120. doi. 10.1007/S11336-008-9101-0

Tagiuri, R. (1968). The concept of organizational climate. In R. Tagiuri & G. H. Litwin (Eds). *Organizational climate: Exploration of a concept*. Boston, MA: Harvard University, Division of Research, Graduate School of Business Administration.

Way, N., Reddy, R., & Rhodes, J. (2007). Students' perceptions of school climate during the middle school years: Associations with trajectories of psychological and behavioral adjustment. *American Journal of Community Psychology*, *40*, 194–213. doi:10.1007/s10464-007-9143-y

Welsh, W. N. (2000). The effects of school climate on school disorder. *Annals of the American Academy of Political and Social Science*, *567*, 88−107. doi:10.1177/0002716200567001007

Zinbarg, R.E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and McDonald's ωH : Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, *70*(1), 123–133. doi:10.1007/s11336-003-0974-7

Zullig, K. J., Koopman, T. M., Patton, J. M., & Ubbes, V. A. (2010). School climate: Historical review, instrument development, and school assessment. *Journal of Psychoeducational Assessment*, *28*, 139-152. doi. 10.1177/0734282909344205

Table 1
June 2012 EFA and CFA Sample Characteristics

Characteristic		EFA Sample %	CFA Sample %	*g*
Grade				
	3	15.5	15.5	0.0
	4	17.5	16.5	0.00
	5	15.6	14.6	0.00
	6	10.6	12.4	0.01
	7	11.3	11.6	0.00
	8	8.6	8.1	0.00
	9	6.8	6.2	-0.01
	10	6.4	7.0	0.01
	11	5.0	4.6	-0.01
	12	2.6	3.6	0.10
Gender				
	Female	50.1	49.1	0.00
	Male	49.4	50.4	0.00
Ethnicity				
	American Indian	0.3	0.3	
	Asian	1.7	2.7	0.26*
	Black	14.5	15.6	0.00
	White	38.7	39.3	0.00
	Hispanic	44.8	42.1	0.00
Lunch Status				
	Free	54.5	55.0	0.00
	Reduced	9.1	8.8	0.00
	Full Pay	36.0	35.8	0.00
EL Status				
	No	90.2	91.4	0.00
	Yes	9.4	8.1	-0.01
Disability Status				
	Yes	11.8	9.9	-0.01
	No	88.2	90.1	0.00

Note. *g values greater than .25 are considered non-equivalent. *EFA* is exploratory factor analysis, CFA is confirmatory factor analysis, disability status included students receiving IEP services for autism, emotional disturbance, hearing impairment, intellectual disability, learning disabilities, multiple disabilities, OHI, and speech or language impairment.

Table 2
Factor Loadings, Communality, Eigenvalues, and Percentage of Variance Explained From Exploratory Principal Components Factor Analysis with Varimax Rotation

	Factors							
Items	1	2	3	4	5	6	7	Communality
30. There are teachers in my school that help me to really want to learn	.73							.60
25. The teachers in my school make learning fun	.73							.56
42. My school handles student behavior problems fairly	.68							.58
40. The adults in my school treat students with respect	.68							.59
37. The adults in my school treat all students fairly	.68							.59
26. I am happy to be at this school	.67							.55
43. At my school, there is a teacher or other adult who listens to me when I have something to say	.66							.60
36. At my school, there is a teacher or other adult who tells me when I do a good job	.66							.53
7. At my school, there is a teacher or other adult whom I can trust	.61							.53
3. There are teachers at my school who care about me	.59							.55

21. I try to do my best at school	.49			.42
33. Other students in my school hurt my feelings (R)		.74		.65
24. I feel sad in school (R)		.73		.58
35. Other students at school have spread mean rumors or lies about me (R)		.63		.56
34. I get hit or threatened by other students (R)		.60		.55
22. I worry about many things (R)		.54		.35
2. I feel safe at school		.43		.49
23. I feel safe on my way to and from school		.42		.36
45. A person's skin color can cause problems at my school (R)			.71	.62
39. Students being mean to other students (harassment) is a problem in my school (R)			.59	.474
32. There is physical fighting between students at my school (R)			.59	.53
31. At school, the color of my skin can get me in trouble (R)			.57	.50
10. Students in my school respect differences in other students			.56	.52

18. At home, I have a parent or other adult who always wants me to do my best	.78			.67
11. At home, I have a parent or other adult who cares about my school work	.78			.67
8. At home, I have a parent or other adult who expects me to follow school rules	.74			.61
12. In the future, I feel I will be successful in life	.45			.42
44. At home, if I need help with homework, a parent or adult will help me		.75		.68
38. At home, I have a quiet place to do my homework		.68		.51
47. At home, I have a parent or other adult who listens to me when I have something to say		.67		.64
41. At home, I have time to do my homework		.60		.50
19. I have a friend about my own age who talks with me about my problems			.78	.63
14. I have a friend about my own age that really cares about me			.78	.65
5. At my school, I have a friend who I can really trust			.69	.55
9. When I have a problem, I find someone to talk with			.52	.44

27. During the past few months, I have hit or pushed other students at school in anger (R)							.73	.60
46. During the past few months, I have hit, pushed or spread mean rumors at the bus stop or on the bus (R)							.72	.56
28. During the past few months, I have spread mean rumors or lies about other students (R)							.72	.57
Eigenvalue	5.94	2.93	2.59	2.52	2.43	2.07	1.98	
% of Variance	16.06	7.92	7.01	6.81	6.56	5.58	5.36	

Note. (R) indicates that the item was reverse coded.

Table 3
Reliability Statistics for the Meriden School Climate Survey-Student Version

				95% CI of α				
Latent Constructs	# of Items	α	SE α	Lower	Upper	ω	β	λ_4
Full-Scale	38	0.91	0.03	0.86	0.97	0.93	0.58	0.93
Adult Support at School	11	0.90	0.01	0.87	0.93	0.92	0.71	0.91
School Safety	7	0.76	0.03	0.71	0.81	0.84	0.59	0.82
Respect for Differences	5	0.76	0.03	0.71	0.81	0.82	0.67	0.81
Adult Support at Home	4	0.75	0.03	0.68	0.81	0.82	0.57	0.75
Academic Support at Home	4	0.72	0.03	0.65	0.79	0.82	0.67	0.78
Aggression Towards Others	3	0.69	0.04	0.60	0.76	0.77	0.67	0.63
Peer Support	4	0.75	0.03	0.68	0.82	0.79	0.69	0.79

Table 4
Factor Score Differences by Student- and School-Level Characteristics

Characteristics		Full Scale	Adult Support at School	School Safety	Respect for Differences	Adult Support at Home	Academic Support at Home	Aggression Towards Others	Peer Support
Grade									
	3	0.40	0.54	-0.09	0.42	-0.09	-0.01	-0.15	-0.12
	4	0.49	0.47	-0.03	0.39	0.08	0.14	-0.08	-0.07
	5	0.44	0.35	0.01	0.38	0.07	0.17	-0.05	0.02
	6	-0.08	0.00	-0.13	-0.22	0.21	0.03	-0.04	-0.14
	7	-0.23	-0.27	0.02	-0.23	0.07	0.03	0.10	-0.07
	8	-0.42	-0.66	0.04	-0.12	0.06	0.06	0.06	0.11
	9	-0.72	-0.65	0.02	-0.54	-0.19	-0.11	0.02	-0.02
	10	-0.63	-0.60	0.20	-0.57	-0.19	-0.38	0.17	0.27
	11	-0.69	-0.66	0.10	-0.61	-0.32	-0.26	0.22	0.34
	12	-0.61	-0.47	0.23	-0.74	-0.27	-0.46	0.28	0.26
School Type									
	Elementary	0.45	0.45	-0.04	0.39	0.02	0.10	-0.09	-0.06
	Middle	-0.22	-0.28	-0.03	-0.20	0.12	0.04	0.04	-0.05
	High	-0.67	-0.61	0.12	-0.59	-0.23	-0.28	0.15	0.19
Gender									
	Female	0.03	0.02	-0.19	-0.02	-0.03	0.05	0.16	0.16
	Male	-0.03	-0.02	0.19	0.01	0.03	-0.05	-0.16	-0.16
Ethnicity									
	Am Indian	0.03	0.00	-0.15	0.52	-0.34	0.15	-0.05	-0.15
	Asian	0.30	0.42	0.10	0.11	0.03	-0.13	-0.42	0.30
	Black	-0.13	-0.12	0.04	-0.21	0.04	0.11	0.02	-0.17
	White	0.01	0.01	-0.04	-0.10	0.02	0.01	0.05	0.12

	Hispanic	0.02	0.01	0.02	0.15	-0.03	-0.04	-0.03	-0.07
SES									
	Free	0.00	0.02	-0.01	0.08	-0.06	0.00	-0.05	-0.08
	Full Pay	0.01	-0.04	0.00	-0.13	0.11	0.00	0.08	0.13
	Reduced	-0.02	-0.01	0.01	0.01	-0.03	-0.02	-0.03	0.00
ELL Status									
	N	0.00	-0.01	0.01	-0.02	0.03	0.00	0.02	0.03
	Y	-0.06	0.14	-0.10	0.17	-0.30	0.03	-0.16	-0.28
Special Education Status									
	N	0.02	-0.02	0.01	-0.01	0.05	0.00	0.02	0.06
	Y	-0.16	0.14	-0.08	0.07	-0.37	-0.01	-0.20	-0.47

Note: Standard Deviation ~1 for all scores (range .78 to 1.22)