Measuring School Climate: Validating the Education Department School Climate Survey in a Sample of Urban Middle and High School Students

Renee Ryberg
Sarah Her
Deborah Temkin (iD)
Rebecca Madill
Claire Kelley
Joy Thompson
Alexander Gabriel

Child Trends

The U.S. Department of Education's School Climate Survey (EDSCLS) is a free, open-source school climate survey available for any local or state education agency to use to measure three domains of school climate: engagement, safety, and environment. The present study leverages EDSCLS data from 3,416 students from 26 middle and high schools in Washington, DC to confirm the factor structure of the survey using both single-level and multilevel confirmatory factor analyses. At the individual level, our findings paralleled those from the original validation study conducted by the U.S. Department of Education. At the school level, our findings suggested a simpler factor structure for the engagement and environment domains, and could not identify a reasonable well-fitting model for the safety domain. Particularly, as more states are considering school climate measurement for accountability systems, these findings suggest that simply using the individual-level structure may not yield valid measurement.

Keywords: *school climate, multilevel confirmatory factor analysis*

SCHOOLS, districts, and states are increasingly recognizing the need to improve school climates to promote the academic, social, and emotional well-being of students (Jordan & Hamilton, 2020; Temkin & Harper, 2018; Thapa et al., 2013; Wang & Degol, 2016). Many states, for instance, included references to school climate within their implementation plans for the Every Student Succeeds Act, and at least six states include school climate measurement as an accountability metric (Jordan & Hamilton, 2020; Temkin & Harper, 2018). Improving school climate, let alone using school climate for accountability purposes, requires the ability to validly and reliably measure school climate. Although there are a multitude of available surveys, most are proprietary and require schools to incur costs to use them (National Center on Safe Supportive Learning Environments [NCSSLE], n.d.). In 2016, the U.S. Department of Education (ED) responded to the need for a freely available school climate survey by developing the "Education Department School Climate Survey" (EDSCLS), a freely available tool any state or local education agency (LEA) can adopt to measure school climate. The tool was initially pilot tested with a diverse set of schools across the United States to refine and

validate the measure (National Center for Education Statistics [NCES], 2015) using a single-level confirmatory factor analysis (CFA) prior to launch. However, to date, no external validation studies have been conducted and no analyses have used a multilevel confirmatory factor analysis procedure that may be more appropriate for measures intended to capture attributes at a school level (Cornell & Huang, 2019; Schweig, 2014). This article leverages baseline data from an ongoing evaluation of a school climate improvement framework in Washington, DC to confirm and extend the findings of these initial pilot tests and demonstrate the validity of the EDSCLS.

Defining School Climate

School climate broadly refers to the quality of multiple facets of a school's environment to support student learning (Thapa et al., 2013). Definitions of school climate vary considerably among researchers and among measurement tools (Wang & Degol, 2016). ED considers school climate across three domains: engagement, or the quality of opportunities for students to connect with the school community through

relationships and activities; safety, or the ability of schools to keep students emotionally and physically safe; and environment, or the quality of the structures and supports that surround the physical plant of a school, academic rigor, and discipline (Bradshaw et al., 2014). Other definitions augment these domains; for instance, separating academic climate into its own domain separate from institutional environment (Wang & Degol, 2016; Zullig et al., 2010) or further splitting engagement into a relationships domain and a school connectedness domain (Zullig et al., 2010). Regardless, most available definitions agree that school climate is multidimensional and reliant on the perceptions of those within the community. That is, school climate is largely the subjective reality of those within a school rather than something that can be objectively identified and perceptions of a school's climate can also vary considerably by different members of that community (Bradshaw et al., 2014; Thapa et al., 2013; Wang & Degol, 2016; Zullig et al., 2010).

Still, the aggregated perceptions of school climate within a school building are significantly associated with academic achievement, engagement in problematic and deviant behaviors, and emotional well-being, although certain aspects of school climate (e.g., engagement) have a broader base of research evidence than other domains (e.g., environment; Haynes et al., 1997; Steffgen et al., 2013; Thapa et al., 2013; Wang & Degol, 2016). It is no surprise then, that education stakeholders are increasingly advocating for a focus on school climate improvement in school reform efforts (see, for instance, Berger et al., 2019; Temkin et al., 2019). Such a focus, however, requires schools to be able to precisely measure and monitor school climate over time.

School Climate Measurement

As attention to the importance of school climate has grown, so too has the proliferation of school climate measurement tools. ED maintains a compendium of the small fraction of school climate measures that have at least some established psychometrics (NCSSLE, n.d.). As of the development of this article, 23 tools that include a student survey component are included on this list. Although each of the measurement tools listed in the compendium have been psychometrically tested, in nearly all cases, such testing was conducted by the survey developer. Among those listed, 16 tools are either copyright protected or not easily publicly accessible and eight are available as fee-for-use. Only six surveys are freely available for public use: The Authoritative School Climate Survey (Cornell, 2014); The Consortium on Chicago School Research Survey of Chicago Public Schools (Consortium on Chicago School Research, n.d.), Delaware Bullying Victimization Student Scale (Bear et al., 2014), Delaware School Climate Survey (Bear et al., 2014), Flourishing Children Survey Social Competence Adolescent Scale (Lippman et al., 2012), and the EDSCLS (NCES,

2016). Although each of these surveys covers an array of topics, only EDSCLS covers the full range of school climate topics stressed in the literature (Thapa et al., 2013; Wang & Degol, 2016; Zullig et al., 2010) and it is the only tool that provides a readily downloadable computer-based platform that allows for administration and analysis without extensive preparation work. ED also provides benchmarks against which schools can compare their aggregate school climate scores and be categorized into "least favorable," "favorable," or "most favorable" performance levels (Ye & Wang, 2018).

Education Department School Climate Survey

EDSLCS was designed to allow any state or LEA to quickly and easily assess climate in its schools. EDSCLS is available for download as a virtual machine that can be embedded on a school system's server for administration and data storage. This allows school systems to use the survey fully free of charge and control access to students' data (i.e., without needing to subscribe to third-party survey providers). The survey tool itself is also available and can be administered through other tools (e.g., online survey tools such as SurveyMonkey or SurveyGizmo) or via paper-and-pencil. EDSCLS consists of three surveys: a student survey, a family survey, and a staff survey. The virtual machine automatically scores the survey data and creates scale scores that can be directly compared with national benchmarks. The student survey, on which this article focuses, measures 12 topics (cultural and linguistic competence, relationships, school participation, emotional safety, physical safety, bullying/cyberbullying, substance abuse, emergency readiness/management, physical environment, instructional environment, mental health, discipline) categorized into the three domains (engagement, safety, environment) of the ED's school climate model. The ED's school climate model, although criticized by some for lacking sufficient theoretical basis for its structure (Wang & Degol, 2016), has shown to be valid for other school climate surveys (Bradshaw et al., 2014).

ED does not maintain a comprehensive list of states and LEAs using EDSCLS. However, in recent years, it has made use of EDSCLS a requirement for certain grant programs (see, e.g., the 2019 School Climate Transformation grant program). As such, the number of schools collecting EDSCLS data is expected to increase as more LEAs and states compete for and receive these federal grants.

The American Institutes for Research, on behalf of NCES, conducted an initial pilot and validation study to refine the EDSCLS survey, which included a convenience sample of 50 schools from across the United States (NCES, 2015). The study used a balanced incomplete block design to minimize burden on students. This meant that no participating student answered all tested items from the EDSCLS. Instead, items

were split into three blocks, with each student responding to two of the three blocks of items. Because the EDSCLS was specifically designed to measure ED's school climate framework, as described previously, the initial validation study used a hierarchical CFA for each of the three primary domains (engagement, safety, and environment) with subtopics under each domain fit as first-order factors. CFAs test whether a given measure aligns with a theoretical model. If the model fits well, this analysis gives validation that the tool is measuring what it is supposed to (Harrington, 2009). After removing items with low loadings, the original study demonstrated good fit for the safety (comparative fit index [CFI] = .91, Tucker–Lewis index [TLI] = .92, root mean square error of approximation [RMSEA] = .09, α = .91) and environment (CFI = .92, TLI = .92, RMSEA = .08, α = .90) domains and marginal fit for the engagement domain (CFI = .87, TLI = .89, RMSEA = .10, α = .90), based on the generous thresholds used in the original study of CFI and TLI greater than .90 and RMSEA less than .10 (Bentler, 1990; Browne & Cudeck, 1993). The study purposefully used relaxed thresholds due to concerns about Type I error (Marsh et al., 2004).

Although this original validation study was foundational in establishing the psychometric properties of the EDSCLS, the present study seeks to address critical limitations of this initial work. Specifically, the original study was focused on narrowing a larger set of items into the final EDSCLS. As such, although the original study demonstrated that the set of selected items fit well together after eliminating nonloading items, the original study could not test whether this remained true in a setting where only those items were presented to students and all students were presented with all items. At present, no study has validated the final EDSCLS student survey tool. Additionally, the initial validation did not take into account the clustered nature of the data. Because EDSCLS is designed to measure school climate at a school level, rather than for each individual student, using a multilevel CFA is necessary to fully validate the measure (Cornell & Huang, 2019; Schweig, 2014). This article leverages baseline data from an ongoing school climate study in the District of Columbia (DC) to test whether the same factor structures identified in the initial validation study of EDSCLS remain when the survey tool is used under typical administration conditions and when accounting for the clustered nature of the data.

Method

Participants

The District of Columbia Public School system (DCPS) served more than 45,000 predominately Black (60%) and economically disadvantaged (77%) students during the 2016–2017 school year (District of Columbia Public Schools, n.d.). Public charter schools in DC served an additional approximately 40,000 students with similar demographic composition (Government of the District of Columbia Office of the State Superintendent of Education, 2017). As part of the baseline data collection for an evaluation of a school climate technical assistance framework, 3,908 students at 26 public and public charter schools in DC completed the EDSCLS during the fall/winter of the 2016–2017 school year. Data were collected in partnership with the DC Office of the State Superintendent of Education (OSSE), which managed the recruitment and consent procedures. For purposes of the evaluation, data were collected from two focal grades at each school, depending on the grade levels served. In 20 schools, data were collected from seventh- and eighth-grade students (N = 2,999) and in six schools, data were collected from ninth- and 10th-grade students (N = 603). Schools additionally had the option to survey students from nonfocal grades (which included Grades 6–12) to provide additional context for data-based decision-making. Schools were asked to survey all students in the given focal grades; however, some schools opted to survey only a sample of students based on time and resource constraints (e.g., availability of computers and tablets for data collection). Because OSSE and the participating schools led the data collection and consenting process, a precise response rate cannot be calculated.

The final analytical sample included 3,416 students in Grades 7 to 10. Data from 492 participants were not used because they either did not respond to any of the EDSCLS items (N = 97), provided the same response for at least 90% of the items (N = 10; consistent with the treatment of data in the initial validation study, NCES, 2015), indicated they were in a grade that the school did not offer or did not survey (N = 18), indicated they were not in the focal-grades of the study (N = 318), or did not respond to the race/ethnicity items used to construct survey weights (N = 49). Student demographic information is presented in Table 1. Notably, half of students in the weighted sample were female, the majority of students (67%) were in Grades 7 or 8, two thirds of students were non-Hispanic Black (66%), and almost one fifth were Hispanic (17%).

Instrumentation

School Climate. Students completed the 68-item EDSCLS (66 items for those in middle school), which measures students' perceptions of school climate in the domains of engagement, safety, and environment, with several topic areas within each domain (e.g., the *engagement* domain includes a *relationships* topic area; the *safety* domain includes a *bullying/cyberbullying* topic area; see Table 2 for items in each topic area). Students responded to each item on a scale from 1 to 4, with 1 indicating "strongly disagree" and 4 indicating "strongly agree." Two items were asked of high school students only (SENGREL153 and SSAFBUL77B in

TABLE 1
Demographic Characteristics (N = 3,416)

Variables	Unweighted *N*	Weighted %
Grade level		
7	1,401	33
8	1,416	34
9	300	19
10	299	13
Gender		
Female	1,685	50
Male	1,694	49
Transgender	27	1
Sexual orientation		
Straight	2,276	81
LGBQ	494	19
Race/ethnicity		
Non-Hispanic Black	1,917	66
Hispanic	607	17
Non-Hispanic, Other, or multiple races	420	4
Non-Hispanic White	472	13

Note. Percentages may not add up to 100 due to rounding. Ten students in the analytic sample were missing data on gender and 646 students were missing data on sexual orientation.

Table 2), as is standard within the EDSCLS platform. One large middle school also agreed to ask students these more sensitive items. The present study focused on 64 items from the EDSCLS, following the developers' decision to exclude the two items in the topic area *emergency readiness/management* and two items in the *safety* domain that loaded poorly but were nonetheless kept on the final tool (NCES, 2015).

Student Demographic Information. EDSCLS includes a number of demographic items (grade level, race, ethnicity, and gender) on the core survey. In addition, the research team added measures of sexual orientation and gender identity based on the request of the participating district (see Temkin et al., 2017 for further information).

Procedure

Consent from students' parents/guardians was obtained passively: OSSE coordinated with each participating school to send parents information about the survey and provided instructions for parents to opt students out of completing the survey a week prior to data collection. Assent was collected from each participating student as part of introductory text at the front of the survey; students were provided with information about the survey and prompted as to whether they would like to continue.

All students completed the survey through a web-browser link to the EDSCLS virtual machine. Depending on the resources at each school, participating students were either brought to a computer lab or tablets or laptops were brought to students' classrooms. Students were provided with a unique username to log in to the survey and complete the assent procedure. A research team member served as a proctor during survey administration to assist with any issues with the login procedures and help maintain student privacy.

Data Analysis

Single-Level Confirmatory Factor Analysis. Consistent with the initial validation study, we conducted a series of CFAs to determine whether the theoretical three-domain model proposed by ED (Figure 1; NCES, 2015) fit the EDSCLS data collected in the present study. Our specific aim was to determine whether the current data produced similar findings as the original validation. As such, we aimed to follow the original procedures to the greatest extent possible. Because EDSCLS was specifically designed to measure this framework, we focused on whether the measure was consistent with this theorized model. Specifically, a hierarchical one-factor model was fit for each of the three domains of the EDSCLS. Items loaded on their respective topic area; in turn, topic areas loaded on their respective domain. Items were analyzed as ordinal categorical measures (as opposed to continuous). These models are depicted in Figures 1 to 3. All analyses were conducted in Mplus Version 8.4 with the weighted least squares mean and variance-adjusted estimator (WLSMV; Muthén & Muthén, 1998–2019). WLSMV accounted for the ordinal nature of the Likert-type response options (Flora & Curran, 2004).

Consistent with the original validation study, we used pairwise deletion to address missing data for the replication analyses. This was important to replicate the original validation study and because two items with more mature content were typically asked only of ninth- and 10th-grade students (the exception being one school where seventh- and eighth-grade students responded to these items). Students were dropped from the CFA if they did not respond to any of the domain's items. Nonresponse was primarily due to students not finishing the survey, leading to the engagement domain having more students with data, compared with the safety and environment domains (which had items at the end of the survey).

Alternative Single-Level Confirmatory Factor Analysis. To improve the model fit and reduce the burden on students, items with standardized factor loadings with an absolute value less than .50 were excluded from the model; a new CFA was run without these items.

TABLE 2
Descriptive Statistics for the EDSCLS Student Survey Items

Variable	Item text	N	Weighted mean	M	SD
Domain: Engagement; Topic area: Cultural and linguistic competence					
SENGCLC1	All students are treated the same, regardless of whether their parents are rich or poor.	3,385	2.86	2.84	0.89
SENGCLC2	Boys and girls are treated equally well.	3,362	2.84	2.79	0.87
SENGCLC3	This school provides instructional materials (e.g., textbooks, handouts) that reflect my cultural background, ethnicity, and identity.	3,276	3.09	3.00	0.76
SENGCLC4	Adults working at this school treat all students respectfully.	3,381	2.80	2.67	0.86
SENGCLC7	People of different cultural backgrounds, races, or ethnicities get along well at this school.	3,337	3.03	3.02	0.75
Domain: Engagement; Topic area: Relationships					
SENGREL9	Teachers understand my problems.	3,324	2.65	2.56	0.87
SENGREL11	Teachers are available when I need to talk with them.	3,331	2.99	2.87	0.76
SENGREL12	It is easy to talk with teachers at this school.	3,331	2.89	2.76	0.81
SENGREL14	My teachers care about me.	3,306	3.07	3.03	0.76
SENGREL153	At this school, there is a teacher or some other adult who students can go to if they need help because of sexual assault or dating violence.	1,252	3.02	3.02	0.81
SENGREL17	My teachers make me feel good about myself.	3,285	2.93	2.83	0.76
SENGREL20	Students respect one another.	3,326	2.36	2.26	0.86
SENGREL21	Students like one another.	3,306	2.60	2.57	0.80
SENGREL29	If I am absent, there is a teacher or some other adult at school that will notice my absence.	3,307	3.24	3.20	0.76
Domain: Engagement; Topic area: Participation					
SENGPAR44	I regularly attend school-sponsored events, such as school dances, sporting events, student performances, or other school activities.	3,315	2.64	2.75	0.91
SENGPAR45	I regularly participate in extracurricular activities offered through this school, such as, school clubs or organizations, musical groups, sports teams, student government, or any other extracurricular activities.	3,285	2.79	2.81	0.90
SENGPAR46	At this school, students have lots of chances to help decide things like class activities and rules.	3,273	2.65	2.55	0.90
SENGPAR47	There are lots of chances for students at this school to get involved in sports, clubs, and other school activities outside of class.	3,271	3.35	3.24	0.76
SENGPAR48	I have lots of chances to be part of class discussions or activities.	3,275	3.16	3.12	0.70
Domain: Safety; Topic area: Emotional safety					
SSAFEMO49	Students at this school get along well with each other.	3,272	2.56	2.48	0.82
SSAFEMO52	At this school, students talk about the importance of understanding their own feelings and the feelings of others.	3,205	2.42	2.32	0.89
SSAFEMO53	At this school, students work on listening to others to understand what they are trying to say.	3,189	2.48	2.46	0.85
SSAFEMO54	I am happy to be at this school.	3,210	2.90	2.84	0.90
SSAFEMO56	I feel like I am part of this school.	3,184	2.92	2.91	0.81
SSAFEMO57	I feel socially accepted.	3,168	2.98	2.99	0.79
Domain: Safety; Topic area: Physical safety					
SSAFPSAF60	I feel safe going to and from this school.	3,204	3.08	3.03	0.77
SSAFPSAF63	I sometimes stay home because I don't feel safe at this school.	3,215	1.53	1.56	0.77
SSAFPSAF65	Students at this school carry guns or knives to school.	3,141	1.70	1.71	0.84
SSAFPSAF67	Students at this school threaten to hurt other students.	3,140	2.43	2.53	0.95
SSAFPSAF68	Students at this school steal money, electronics, or other valuable things while at school.	3,137	2.41	2.43	0.98
SSAFPSAF69	Students at this school damage or destroy other students' property.	3,127	2.41	2.54	0.92
SSAFPSAF71	Students at this school fight a lot.	3,138	2.66	2.65	0.91

(continued)

5

TABLE 2. (CONTINUED)

Variable	Item text	N	Weighted mean	M	SD
Domain: Safety; Topic area: Bullying					
SSAFBUL74	Students at this school are teased or picked on about their race or ethnicity.	3,103	2.14	2.19	0.93
SSAFBUL75	Students at this school are teased or picked on about their cultural background or religion.	3,072	2.04	2.05	0.90
SSAFBUL76	Students at this school are teased or picked on about their physical or mental disability.	3,073	2.25	2.29	0.97
SSAFBUL77B	Students at this school are teased or picked on about their real or perceived sexual orientation.	1,197	2.25	2.33	0.95
SSAFBUL73	Students at this school are often bullied.	3,084	2.31	2.37	0.89
SSAFBUL80	Students at this school try to stop bullying.	3,078	2.59	2.55	0.89
SSAFBUL83	Students often spread mean rumors or lies about others at this school on the internet (i.e., Facebook, e-mail, and instant message).	3,030	2.61	2.62	0.95
Domain: Safety; Topic area: Substance abuse					
SSAFSUB88	Students use/try alcohol or drugs while at school or school-sponsored events.	3,033	2.05	1.86	0.87
SSAFSUB91	It is easy for students to use/try alcohol or drugs at school or school-sponsored events without getting caught.	3,029	2.11	1.99	0.95
SSAFSUB92	Students at this school think it is okay to smoke one or more packs of cigarettes a day.	2,954	1.92	1.76	0.85
SSAFSUB93	Students at this school think it is okay to get drunk.	2,933	2.17	1.96	0.90
SSAFSUB94	Students at this school think it is okay to try drugs.	2,933	2.42	2.18	0.99
Domain: Environment; Topic area: Physical environment					
SENVPENV100	The bathrooms in this school are clean.	3,017	2.53	2.21	0.97
SENVPENV102	The temperature in this school is comfortable all year round.	3,015	2.41	2.25	0.88
SENVPENV105	The school grounds are kept clean.	3,022	2.81	2.62	0.89
SENVPENV106	I think that students are proud of how this school looks on the outside.	2,913	2.98	2.79	0.88
SENVPENV107	Broken things at this school get fixed quickly.	2,933	2.60	2.41	0.91
Domain: Environment; Topic area: Instructional environment					
SENVINS111	My teachers praise me when I work hard in school.	2,949	3.00	2.99	0.82
SENVINS113	My teachers give me individual attention when I need it.	2,917	2.95	2.83	0.80
SENVINS114	My teachers often connect what I am learning to life outside the classroom.	2,856	2.83	2.72	0.85
SENVINS115	The things I'm learning in school are important to me.	2,910	2.97	3.05	0.83
SENVINS121	My teachers expect me to do my best all the time.	2,910	3.42	3.39	0.70
Domain: Environment; Topic area: Mental health					
SENVMEN130	My teachers really care about me.	2,844	3.02	2.95	0.82
SENVMEN132	I can talk to my teachers about problems I am having in class.	2,851	3.03	2.90	0.85
SENVMEN133	I can talk to a teacher or other adult at this school about something that is bothering me.	2,812	3.00	2.92	0.85
SENVMEN134	Students at this school stop and think before doing anything when they get angry.	2,791	2.16	2.04	0.87
SENVMEN137	Students at this school try to work out their disagreements with other students by talking to them.	2,782	2.36	2.23	0.91
Domain: Environment; Topic area: Discipline					
SENVDIS142	My teachers make it clear to me when I have misbehaved in class.	2,801	3.15	3.14	0.75
SENVDIS143	Adults working at this school reward students for positive behavior.	2,794	2.92	2.90	0.85
SENVDIS146	Adults working at this school help students develop strategies to understand and control their feelings and actions.	2,722	2.81	2.82	0.83
SENVDIS147	School rules are applied equally to all students.	2,809	2.84	2.83	0.95
SENVDIS147C	Discipline is fair.	2,736	2.79	2.66	0.92

Note. All items were rated on a scale from 1 to 4. EDSCLS = U.S. Department of Education's School Climate Survey.
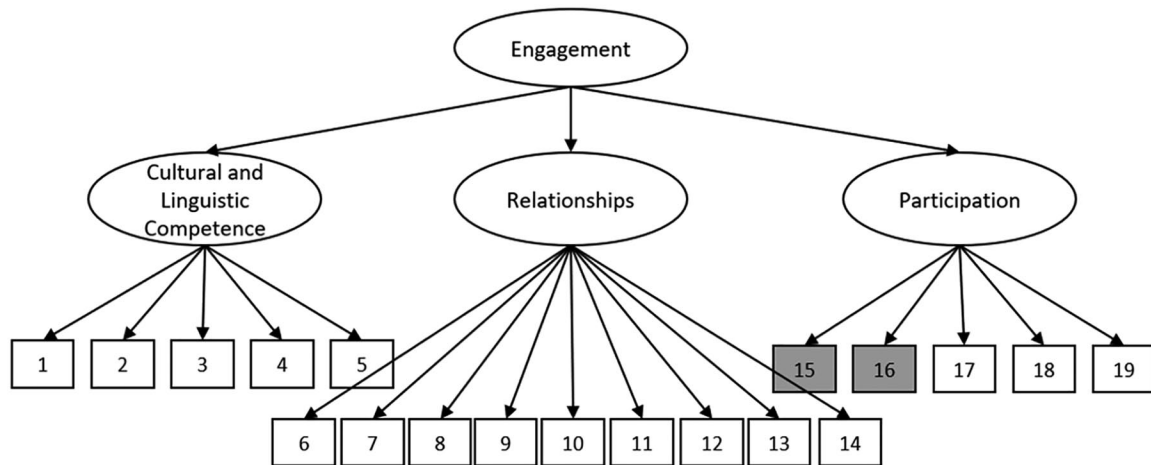
FIGURE 1. *Hierarchical factor structure for the U.S. Department of Education's School Climate Survey (EDSCLS) Engagement domain.*
Note. Indicators in gray boxes are behavioral items that were dropped from the alternative models. Additional items in the engagement and safety domains were dropped once multiple imputation was introduced.

FIGURE 2. *Hierarchical factor structure for the U.S. Department of Education's School Climate Survey (EDSCLS) Safety domain.*
Note. Indicators in gray boxes are behavioral items that were dropped from the alternative models. Additional items in the engagement and safety domains were dropped once multiple imputation was introduced.

Alternative Single-Level Confirmatory Factor Analysis With Multiple Imputation. Moving beyond replication, we tested additional models treating missing data with multiple imputation.

As noted above, nonresponse was primarily due to students not finishing the survey. Item-level missingness ranged from 0.91% at the beginning of the survey to 20.32% at the end of the survey (see Table 2 for the number of respondents for each item). There were also times of technical difficulties where students had trouble obtaining charged laptops, connecting to the server, and other school events which reduced the time the students had to complete the survey.

Imputation was performed in Mplus, running the procedure separately for each of the three domains, as we did not hypothesize a relationship between the three domains of

school climate in our models. We imputed five data sets (Rubin, 1987) with a two-level structure (students nested within schools) using all survey items in that domain, student grade level, and race and gender (Pedersen et al., 2017). All indicators were imputed as categorical. The two items that were only asked to high school students were excluded from the analysis since data for these two items were not missing at random, violating the assumptions of multiple imputation (Jakobsen et al., 2017).

Multilevel Confirmatory Factor Analysis. To further improve the model fit and take into account the clustered nature of the data, a series of multilevel confirmatory factor analyses (MCFAs) were run, based on the alternative single-level CFA that excluded the items with standardized factor
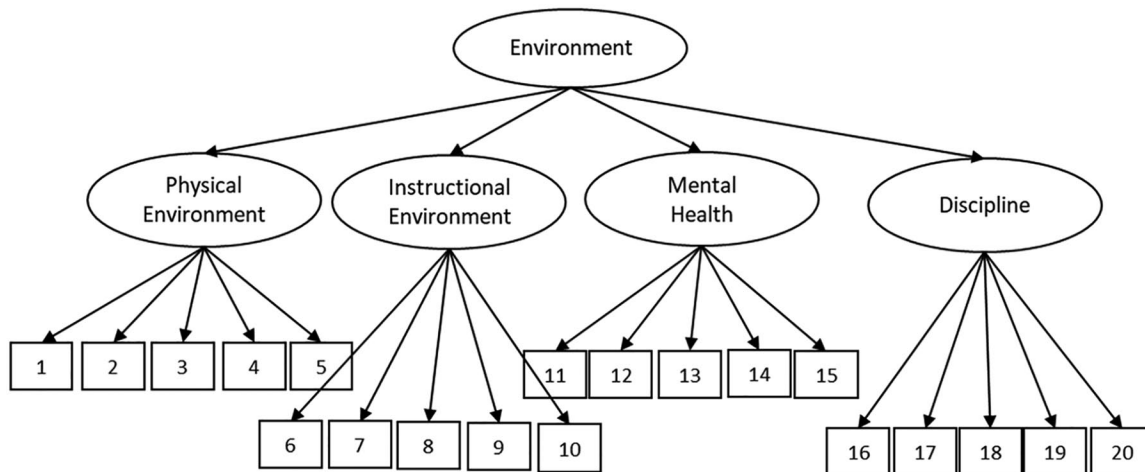
7

FIGURE 3.   *Hierarchical factor structure for the U.S. Department of Education's School Climate Survey (EDSCLS) Environment domain.*

loadings less than .50 in the original CFA. All MCFA analyses used multiple imputation to address missing data.

Weights. All analyses were weighted to adjust for potential bias in the sample due to differential student nonresponse. We constructed poststratification weights based on the inverse probability that a student responded to the survey based on their race/ethnicity and the size of their grade-level according to publicly available aggregate data for each school. With the weights, results are generalizable to the schools and grades surveyed. The weights were scaled in Mplus using the wtscale command in the multilevel confirmatory factor analyses (Asparouhov, 2008; Carle, 2009).

Results

Single-Level Confirmatory Factor Analysis

Standardized parameter estimates from each CFA are shown in Tables 3 to 7. Table 8 shows the fit indices from the original study in comparison to the current study. We use two sets of thresholds for fit indices to determine if models were an acceptable description of the underlying data—those used in the original validation study (NCES, 2015): >.90 for the CFI and the TLI (Bentler, 1990) and <.10 for the RMSEA (Browne & Cudeck, 1993); and stricter, more conventional thresholds: >.95 for the CFI and TLI and <.06 for the RMSEA (Hu & Bentler, 1998).

Overall, results suggest that the three measurement models had a similar fit in the DC data as in the pilot data.

Engagement Domain. The engagement domain met the original study's criteria for the CFI and RMSEA but, as in the original validation study, the TLI (.89) was slightly outside the recommended threshold. The CFI (.91) and TLI (.89) both fall short of the stricter, more conventional thresholds,

and the RMSEA (.06) is right at the threshold. Standardized factor loadings were at least .50 with the exception of two items with factor loadings between .40 and .50. These two items were the only items that asked about students' actual behavior, rather than students' perceptions about the school climate:

> I regularly attend school-sponsored events, such as school dances, sporting events, student performances, or other school activities. (Topic area: Participation)

> I regularly participate in extra-curricular activities offered through this school, such as, school clubs or organizations, musical groups, sports teams, student government, or any other extra-curricular activities. (Topic area: Participation)

These items had similarly low, but marginally acceptable standardized factor loadings (both .53) in the original validation study. For both the current and the original studies, these items were the lowest loading across all items in the engagement domain.

Safety Domain. The safety domain met criteria for each of the model fit indices according to the original study's criteria, but the CFI (.93) and TLI (.92) fall below conventional thresholds, and the RMSEA (.06) is right at the threshold. Standardized factor loadings were at least .50 with the exception of one item with a factor loading of .44. Similar to the engagement domain, the item with a poor loading was the only item that asked about students' actual behavior:

> I sometimes stay home because I don't feel safe at this school.

This item had a similarly low standardized factor loading (.49) during the original validation study, which was the lowest loading across all items in the safety domain and fell below the .50 threshold.

8

TABLE 3

Engagement Domain: Standardized Item Loadings From a One-Factor Hierarchical CFA

Topic Area and Item	CFA		Alternative CFA		CFA, multiple imputation	
	β	SE	β	SE	β	SE
Topic area: Cultural and linguistic competence						
All students are treated the same, regardless of whether their parents are rich or poor.	.63	.02	.63	.02	.63	.02
Boys and girls are treated equally well.	.67	.01	.67	.01	.67	.01
This school provides instructional materials (e.g., textbooks, handouts) that reflect my cultural background, ethnicity, and identity.	.53	.02	.53	.02	.52	.02
Adults working at this school treat all students respectfully.	.75	.01	.76	.01	.75	.01
People of different cultural backgrounds, races, or ethnicities get along well at this school.	.57	.02	.56	.02	.56	.02
Topic area: Relationships						
Teachers understand my problems.	.75	.01	.75	.01	.75	.01
Teachers are available when I need to talk with them.	.72	.01	.72	.01	.71	.01
It is easy to talk with teachers at this school.	.75	.01	.75	.01	.75	.01
My teachers care about me.	.75	.01	.74	.01	.75	.01
At this school, there is a teacher or some other adult who students can go to if they need help because of sexual assault or dating violence.	.63	.03	.63	.03	—	—
My teachers make me feel good about myself.	.76	.01	.76	.01	.76	.01
Students respect one another.	.59	.02	.59	.02	.59	.02
Students like one another.	.52	.02	.52	.02	.52	.02
If I am absent, there is a teacher or some other adult at school that will notice my absence.	.54	.02	.53	.02	.53	.02
Topic area: Participation						
I regularly attend school-sponsored events, such as school dances, sporting events, student performances, or other school activities.	.49	.02	—	—	—	—
I regularly participate in extracurricular activities offered through this school, such as, school clubs or organizations, musical groups, sports teams, student government, or any other extracurricular activities.	.48	.02	—	—	—	—
At this school, students have lots of chances to help decide things like class activities and rules.	.65	.02	.64	.02	.64	.02
There are lots of chances for students at this school to get involved in sports, clubs, and other school activities outside of class.	.61	.02	.60	.02	.59	.02
I have lots of chances to be part of class discussions or activities.	.74	.02	.73	.02	.73	.02
Loadings of topic areas on the general engagement factor						
Cultural and linguistic competence	.87	.02	.88	.02	.89	.01
Relationships	.99	.01	.98	.01	.97	.01
Participation	.79	.02	.83	.02	.93	.02

Note. CFA results are from a one-factor hierarchical confirmatory factor analysis estimated using weighted least squares means and variance. $N = 3,416$ for all models. β = standardized factor loading; *SE* = standard error; CFA = confirmatory factor analysis.

Environment Domain. The environment domain met all criteria for fit indices according to the thresholds used in the original validation study and are close to meeting conventional thresholds. All standardized factor loadings were at least .50.

Alternative Confirmatory Factor Analysis

Alternative models excluded the aforementioned items that had standardized factor loadings less than .50. Fit indices for the alternative models are presented in Table 8.

9

TABLE 4

Safety Domain: Standardized Item Loadings From a One-Factor Hierarchical CFA

Item text	CFA		Alternative CFA		CFA, multiple imputation	
	β	SE	β	SE	β	SE
Topic area: Emotional safety						
Students at this school get along well with each other.	.75	.02	.75	.02	.75	.01
At this school, students talk about the importance of understanding their own feelings and the feelings of others.	.61	.02	.61	.02	.60	.02
At this school, students work on listening to others to understand what they are trying to say.	.73	.01	.73	.01	.73	.01
I am happy to be at this school.	.78	.01	.78	.01	.78	.01
I feel like I am part of this school.	.80	.01	.80	.01	.79	.01
I feel socially accepted.	.69	.02	.69	.02	.68	.02
Topic area: Physical safety						
I feel safe going to and from this school.	−.54	.02	−.53	.02	−.53	.02
I sometimes stay home because I don't feel safe at this school.	.44	.02	—	—	—	—
Students at this school carry guns or knives to school.	.67	.02	.67	.02	.66	.02
Students at this school threaten to hurt other students.	.79	.01	.79	.01	.79	.01
Students at this school steal money, electronics, or other valuable things while at school.	.78	.01	.78	.01	.78	.01
Students at this school damage or destroy other students' property.	.82	.01	.82	.01	.82	.01
Students at this school fight a lot.	.72	.01	.72	.01	.72	.01
Topic area: Bullying						
Students at this school are teased or picked on about their race or ethnicity.	.84	.01	.84	.01	.84	.01
Students at this school are teased or picked on about their cultural background or religion.	.84	.01	.84	.01	.84	.01
Students at this school are teased or picked on about their physical or mental disability.	.77	.01	.77	.01	.77	.01
Students at this school are teased or picked on about their real or perceived sexual orientation.	.80	.02	.80	.02	—	—
Students at this school are often bullied.	.80	.01	.80	.01	.80	.01
Students at this school try to stop bullying.	−.51	.02	−.51	.02	−.51	.02
Students often spread mean rumors or lies about others at this school on the internet (i.e., Facebook, email, and instant message).	.66	.02	.66	.02	.66	.02
Topic area: Substance abuse						
Students use/try alcohol or drugs while at school or school-sponsored events.	.82	.01	.81	.01	.81	.01
It is easy for students to use/try alcohol or drugs at school or school-sponsored events without getting caught.	.76	.01	.76	.01	.75	.01
Students at this school think it is okay to smoke one or more packs of cigarettes a day.	.77	.01	.77	.01	.76	.01
Students at this school think it is okay to get drunk.	.84	.01	.84	.01	.85	.01
Students at this school think it is okay to try drugs.	.87	.01	.87	.01	.87	.01
Loadings of topic areas on the general safety factor						
Emotional safety	−.60	.02	−.60	.02	−.60	.02
Physical safety	.95	.01	.94	.01	.94	.01
Bullying	.88	.01	.89	.01	.89	.01
Substance abuse	.69	.01	.69	.01	.68	.02

Note. $N = 3,336$ for the CFA and alternative CFA models. $N = 3,416$ for the model with multiple imputation. CFA results are from a one-factor hierarchical confirmatory factor analysis estimated using weighted least squares means and variance. β = standardized factor loading; *SE* = standard error; CFA = confirmatory factor analysis.

TABLE 5

Environment Domain: Standardized Item Loadings From a One-Factor Hierarchical CFA

Item text	CFA		CFA, multiple imputation	
	β	*SE*	β	*SE*
Topic area: Physical environment				
The bathrooms in this school are clean.	.72	.02	.72	.02
The temperature in this school is comfortable all year round.	.63	.02	.64	.02
The school grounds are kept clean.	.78	.01	.78	.02
I think that students are proud of how this school looks on the outside.	.66	.02	.66	.02
Broken things at this school get fixed quickly.	.77	.02	.76	.02
Topic area: Instructional environment				
My teachers praise me when I work hard in school.	.72	.01	.72	.01
My teachers give me individual attention when I need it.	.72	.01	.72	.02
My teachers often connect what I am learning to life outside the classroom.	.67	.01	.67	.01
The things I'm learning in school are important to me.	.65	.02	.65	.02
My teachers expect me to do my best all the time.	.66	.02	.65	.02
Topic area: Mental health				
My teachers really care about me.	.83	.01	.82	.01
I can talk to my teachers about problems I am having in class.	.82	.01	.82	.01
I can talk to a teacher or other adult at this school about something that is bothering me.	.77	.01	.76	.01
Students at this school stop and think before doing anything when they get angry.	.58	.02	.58	.02
Students at this school try to work out their disagreements with other students by talking to them.	.63	.02	.62	.02
Topic area: Discipline				
My teachers make it clear to me when I have misbehaved in class.	.56	.02	.57	.02
Adults working at this school reward students for positive behavior.	.67	.02	.66	.02
Adults working at this school help students develop strategies to understand and control their feelings and actions.	.75	.01	.76	.01
School rules are applied equally to all students.	.71	.01	.71	.02
Discipline is fair.	.67	.02	.67	.02
Loadings of topic areas on the general environment factor				
Physical environment	.66	.02	.66	.02
Instructional environment	.94	.01	.93	.01
Mental health	.93	.01	.93	.01
Discipline	.94	.01	.94	.01

Note. $N = 3,120$ for the CFA. $N = 3,416$ for the model with multiple imputation. CFA results are from a one-factor hierarchical confirmatory factor analysis estimated using weighted least squares means and variance. β = standardized factor loading. *SE* = standard error; CFA = confirmatory factor analysis.

For the engagement domain, the alternative model fit the data slightly better than the original model, and all three fit statistics met the fit criteria of the original validation study, but the CFI and TLI continue to be less than conventional thresholds, and RMSEAs are right at conventional thresholds.

When rerun with multiple imputation, the alternative models had very similar fit statistics across domains.

Multilevel Confirmatory Factor Analysis

MCFA models were consistent with the alternative model in that they dropped the items with standardized factor loadings less than .50 (i.e., items that asked about students' behaviors) in the single-level CFA model. Across the three domains, MCFA models following the factor structure of the single-level CFA models, with hierarchical factor models at both the within and between levels, did not fit the data at the between level. To improve model fit, we tested a progressive series of models to explore whether simplifying or otherwise altering the between-level structure (including correlated factor models) would yield a better-fitting model (Chen et al., 2001; see Appendix Table A1 for detailed descriptions of each model and their fit statistics). In selecting the recommended models, we relied on fit statistics, factor loadings, and other model parameters, and selected the best-fitting

11

TABLE 6

Engagement Domain: Standardized Item Loadings From an MCFA With Hierarchical Structure at the Within Level and a Single Factor at the Between Level (N = 3,416 Students)

Topic area and item	Within level		Between level	
	β	SE	β	SE
Topic area: Cultural and linguistic competence				
All students are treated the same, regardless of whether their parents are rich or poor.	.63	.02	.72	.10
Boys and girls are treated equally well.	.77	.04	.69	.13
This school provides instructional materials (e.g., textbooks, handouts) that reflect my cultural background, ethnicity, and identity.	.52	.02	.78	.11
Adults working at this school treat all students respectfully.	.78	.02	.98	.08
People of different cultural backgrounds, races, or ethnicities get along well at this school.	—	—	—	—
Topic area: Relationships				
Teachers understand my problems.	.78	.02	.87	.08
Teachers are available when I need to talk with them.	.66	.02	.97	.07
It is easy to talk with teachers at this school.	.69	.01	.97	.07
My teachers care about me.	.81	.01	.70	.12
At this school, there is a teacher or some other adult who students can go to if they need help because of sexual assault or dating violence.	—	—	—	—
My teachers make me feel good about myself.	.79	.01	.76	.10
Students respect one another.	.59	.02	.51	.14
Students like one another.	.61	.05	.47	.16
If I am absent, there is a teacher or some other adult at school that will notice my absence.	.51	.02	.70	.15
Topic area: Participation				
I regularly attend school-sponsored events, such as school dances, sporting events, student performances, or other school activities.	—	—	—	—
I regularly participate in extra-curricular activities offered through this school, such as, school clubs or organizations, musical groups, sports teams, student government, or any other extra-curricular activities.	—	—	—	—
At this school, students have lots of chances to help decide things like class activities and rules.	.62	.02	.75	.12
There are lots of chances for students at this school to get involved in sports, clubs, and other school activities outside of class.	.61	.03	.66	.13
I have lots of chances to be part of class discussions or activities.	.71	.02	.74	.09
Loadings of topic areas on the general engagement factor				
Cultural and linguistic competence	.86	.02	—	—
Relationships	.97	.02	—	—
Participation	.81	.02	—	—

Note. MCFA results are from a one-factor hierarchical confirmatory factor analysis estimated using weighted least squares means and variance. β = standardized factor loading; *SE* = standard error; MCFA = multilevel confirmatory factor analysis.

models that were most well-aligned with the original factor structure hypothesized in the original validation study.

Engagement Domain. The recommended multilevel measurement model for engagement uses the hierarchical factor structure for the within level and a single factor at the between level. One item, "People of different cultural backgrounds, races, or ethnicities get along well at this school" is dropped, due to a low factor loading of .37 at the between level. Model fit for the engagement domain improved over the alternative CFA model (CFI and TLI increased from .92 and .90 to .96 and .95, respectively, and RMSEA decreased from .07 to .03). The standardized factor loadings were at least .50 at the within and between levels, with the exception of one item at the between level with a factor loading of .47 (see Table 6).

Safety Domain. Although some multilevel measurement models appear to fit the data well according to the fit statistics (generally meeting both the thresholds used in the original validation study and the more stringent conventional

TABLE 7

Environment Domain: Standardized Item Loadings From an MCFA With Hierarchical Structure at the Within Level and a Single Factor at the Between Level (N = 3,416 Students)

Item text	Within level		Between level	
	β	SE	β	SE
Topic area: Physical environment				
The bathrooms in this school are clean.	.62	.06	.68	.18
The temperature in this school is comfortable all year round.	.63	.02	.60	.15
The school grounds are kept clean.	.73	.02	.67	.18
I think that students are proud of how this school looks on the outside.	.67	.06	.55	.31
Broken things at this school get fixed quickly.	.71	.02	.84	.14
Topic area: Instructional environment				
My teachers praise me when I work hard in school.	.72	.01	.54	.12
My teachers give me individual attention when I need it.	.75	.01	.70	.12
My teachers often connect what I am learning to life outside the classroom.	.66	.02	.95	.07
The things I'm learning in school are important to me.	—	—	—	—
My teachers expect me to do my best all the time.	.66	.01	.52	.18
Topic area: Mental health				
My teachers really care about me.	.81	.01	.87	.10
I can talk to my teachers about problems I am having in class.	.76	.01	.75	.12
I can talk to a teacher or other adult at this school about something that is bothering me.	.52	.02	.73	.09
Students at this school stop and think before doing anything when they get angry.	.57	.03	.71	.11
Students at this school try to work out their disagreements with other students by talking to them.	.82	.01	.82	.08
Topic area: Discipline				
My teachers make it clear to me when I have misbehaved in class.	.58	.01	.55	.18
Adults working at this school reward students for positive behavior.	—	—	—	—
Adults working at this school help students develop strategies to understand and control their feelings and actions.	.74	.02	.69	.11
School rules are applied equally to all students.	.64	.03	.97	.06
Discipline is fair.	.68	.04	.87	.08
Loadings of topic areas on the general environment factor				
Physical environment	.72	.02	—	—
Instructional environment	.92	.01	—	—
Mental health	.95	.02	—	—
Discipline	.96	.03	—	—

Note. MCFA results are from a one-factor hierarchical confirmatory factor analysis estimated using weighted least squares means and variance. β = standardized factor loading; *SE* = standard error; MCFA = multilevel confirmatory factor analysis.

thresholds), some models did not converge, and, among those that did, there were improper solutions at the between level across all tested models. These included negative residual variances, very high correlations across items in different practice areas, and very low factor loadings between the substance use practice area and the safety domain. We tried numerous ways to address these issues, including fixing negative variances to zero, dropping highly correlated items, fitting simpler models at the between level, fitting correlated models, and dropping substance use items but were not able to find an appropriate multilevel measurement model for the safety domain.

Environment Domain. The recommended multilevel measurement model for environment uses the hierarchical factor structure for the within level and a single factor at the between level, and drops two items due to low factor loadings at the between level: "The things I'm learning in school are important to me" (with a factor loading of .39) and "Adults working at this school reward students for positive behavior" (with a factor loading of .45). Model fit for the environment domain improved over the alternative CFA model (CFI and TLI improved from .94 and .93 to .98 and .97, respectively, and RMSEA decreased from .06 to .02).

Discussion

As an increasing number of schools, districts, and states move toward measuring school climate as part of

TABLE 8

Model Fit Statistics and Reliability: EDSCLS Pilot Study, Compared With DC Sample

Domain	Pilot study					DC sample				
	N	RMSEA	CFI	TLI	α	*N*	RMSEA	CFI	TLI	α
Engagement	11,439	.10	.87	.89	.90	3416	.06	.91	.89	.88
Safety	11,494	.09	.91	.92	.91	3336	.06	.93	.92	.78
Environment	11,509	.08	.92	.93	.90	3120	.05	.95	.94	.90

Domain	DC sample: Alternate models dropping items with standardized loadings <.5[a]					DC sample: Alternate models with multiple imputation[a,b]				
	N	RMSEA	CFI	TLI	α	*N*	RMSEA	CFI	TLI	α
Engagement	3,416	.07	.92	.91	.88	3416	.07	.92	.90	.88
Safety	3,336	.06	.93	.92	.77	3416	.07	.92	.91	.75
Environment	No loadings ≤.5					3,416	.06	.94	.93	.90

Domain	DC sample: Multilevel models with hierarchical factor structure at within level and single factors at between level with multiple imputation[a,b,c]				
	N	RMSEA	CFI	TLI	α Within	α Between
Engagement	3,416	.03	.96	.95	.87	.95
Safety	No fitting model					
Environment	3,416	.02	.98	.97	.89	.94

Note. Cronbach's alpha values are presented as a measure of reliability. Values for the pilot study are from National Center for Education Statistics (2015). Values for the DC sample models were calculated in Mplus, treating the items as continuous and using all five data sets for the multiple imputation, where appropriate. Multilevel Cronbach's alphas were calculated using the between- and within-variance covariance matrices separately, using code adapted from online Appendix C of Geldhof et al. (2014). CFI = comparative fit index; RMSEA = root mean square error of approximation; TLI = Tucker–Lewis index; MCFA = multilevel confirmatory factor analysis.
[a]In the alternate models, the Engagement domain dropped the following two items from the Participation topic area: "I regularly attend school-sponsored events, such as school dances, sporting events, student performances, or other school activities" and "I regularly participate in extra-curricular activities offered through this school, such as, school clubs or organizations, musical groups, sports teams, student government, or any other extra-curricular activities." The Safety domain dropped the following item from the Physical Safety topic area: "I sometimes stay home because I don't feel safe at this school." [b]In CFA models on imputed data sets the Engagement domain dropped the following item from the Relationship topic area: "At this school, there is a teacher or some other adult who students can go to if they need help because of sexual assault or dating violence." The Safety domain dropped the following item from the Bullying topic area: "Students at this school are teased or picked on about their real or perceived sexual orientation." [c]Additional items were dropped from MCFA models due to low factor loadings: in the Engagement domain, Cultural and Linguistic Competence topic area: "People of different cultural backgrounds, races, or ethnicities get along well at this school;" In the Environment domain, Instructional Environment topic area: "The things I'm learning in school are important to me"; and in the Environment domain, Discipline topic area: "Adults working at this school reward students for positive behavior."

accountability and school improvement initiatives, it is critical to ensure the tools used are valid across populations and when used in normal administration conditions. The current study is the first to independently validate the structure of the EDSCLS outside of ED's original validation study and the first to use a multilevel CFA to explore not only the EDSCLS's structure at the individual level but also at the school level. The initial study leveraged pilot data from a diverse array of schools across the country and used a balanced incomplete block design whereby no student took the full survey tool (NCES, 2015). The present study used data from 3,416 students from 26 middle and high schools in DC which were collected under typical conditions—all students were presented with the full survey tool. We ran a series of hierarchical one-factor CFAs on each of EDSCLS' three domains (engagement, safety, and environment) to test whether items loaded on both their topic area scales as well as the overall domain. We then improved on

this model with multilevel CFA that accounts for the clustered nature of data collected from students within schools.

Our findings largely paralleled those of the original validation study for the single-level CFA. For each of the three domains, the RMSEA and CFI indicated that the measurement model fit the data according to the less rigorous standards used in the original study (>.90 for the CFI and the TLI [Bentler, 1990] and <.10 for the RMSEA [Browne & Cudeck, 1993]). However, for the engagement domain, the TLI was outside the recommended threshold, suggesting that the proposed measurement model is not an ideal fit to the data. Using stricter and more conventional thresholds(>.95 for the CFI and TLI and <.06 for the RMSEA [Hu & Bentler, 1998]), CFI and TLI for the engagement and safety domains indicated a less well-fitting model, with RMSEA right at threshold, and for the environment domain, all parameters were close to meeting these thresholds.

Three items, all focused on personal behaviors (e.g., "I sometimes stay at home . . ."), had particularly low standardized loadings. By dropping these items, all of the fit statistics pass thresholds for acceptable fit using the original parameters but continued to fall short of the stricter thresholds. Given the EDSCLS's general focus on broad perceptions of school climate, the fact that the three personal behavior items had low loadings is not surprising. On their face, these items are measuring a different, although highly correlated, concept. Although a student's decision to participate in school activities is necessarily associated with the opportunities provided by a school, there are many other external and intrapersonal factors that contribute to a student's decision to engage (Feldman & Matjasko, 2005).

Although our findings closely replicated those of the original validation study, we note that the model falls short when using rigorous cutoffs. Marsh et al. (2004) argue that there is no "golden rule" for cutoff scores and using too stringent of cutoffs can incorrectly reject well-fitting models (i.e., Type I error). Given how closely our fit indices align with those of the original validation, and given that the RMSEA met the more rigorous threshold for each of the EDSCLS domains even as the others fell short, we argue that there remains continued support for the validity of the underlying factor structure at the individual level.

Prior to this study, EDSCLS was only examined at the individual level; its use as a school-level measure had not previously been tested. Models replicating the individual within-school model—where each domain consists of three or four subdomains—at the between-school level did not fit the data well. Instead, for engagement and environment, a more simplified model consisting of a single higher order factor at the between level fit the data well, with fit statistics reaching more rigorous standards. This is consistent with several prior studies suggesting that Level 1 and Level 2 structures are not often consistent, and typically demonstrate simpler structures at higher levels (Huang & Cornell, 2016).

For safety, however, we did not identify a model that had acceptable fit and was free of statistical violations (e.g., negative residual variances). Although we urge caution in this finding given our relatively low power at the between level ($N = 26$ schools), its inconsistency with the other domains raises critical questions about the use of student perceptions of safety to generate an overarching school safety score. Only a few previous studies have used multilevel CFA to validate student school climate surveys (e.g., Huang & Cornell, 2016; Konold & Cornell, 2015) and these only include some aspects of school safety as defined by the EDSCLS (e.g., bullying and teasing). Given the EDSCLS's inclusion and focus on physical safety and substance use alongside bullying, and the relatively low prevalence of physical violence and substance use but high levels of bullying, particularly, at the middle school level (Musu et al., 2019), it may be that although a domain-level factor fits at the individual level, there is insufficient consistency between subdomains at the school level to create a higher order factor. Further research with a larger sample of schools is needed.

Our between-school findings do raise questions, however, about the push to include school climate survey data as part of states' accountability plans (Jordan & Hamilton, 2020; Temkin & Harper, 2018). Although our models for environment and engagement ultimately fit well, they differ substantively from the individual-level structures identified through single-level CFA. This means that states risk calculating scores that may not accurately reflect schools' climates. Given that many school climate measures have not yet been validated using a multilevel CFA, they may simply not be ready to use as an accountability tool.

Still, our study provides broad support for using the EDSCLS to examine within-school, individual-level differences in perceptions of school climate across all domains and for comparing schools on engagement and environment. Understanding how perceptions of school climate vary between individuals is important for schools to ensure that their interventions are reaching all students. Users should take note of the item modifications (e.g., removal of the three student behavior items) identified in the course of our analyses. This may mean that standard scale scores populated by the EDSCLS platform may need adjustment, although further replication using broader samples is necessary.

Limitations and Future Directions

There are a few key limitations to this study. First, the sample predominately consisted of seventh and eighth graders, limiting its ability to generalize to high schools. EDSCLS contains two items that are designated as "high school only." Because our sample predominately consisted of seventh- and eighth-grade students, and because only one school administered these items to seventh- and eighth-grade students, there was substantial missing data on these items. However, our factor loadings for these items in the single-level CFA were similar to those for the original validation study and as such, the missing data do not seem to have affected our findings.

Additionally, although the research team provided proctoring in order to help protect student confidentiality, because students completed these surveys on tablets or computers, their answers may have been more visible to classmates than through paper-and-pencil surveys. This may have affected how truthful students were in their responses, and the instrument did not include any validation items to allow us to account for social desirability bias or identify mischievous responders. Schools should use technology such as screen shields to help protect student privacy when administering online surveys such as the EDSCLS.

This study primarily focused on confirming the findings from the original EDSCLS validation study using CFA. Future studies, however, should explore whether there are alternative models that may better fit the EDSCLS data and provide a more nuanced understanding of school climate, particularly at the school level and for the safety domain.

Conclusion

Our findings' close replication of the original validation study suggests that the EDSCLS functions as expected during real-world administration for assessing individuals' perceptions of school climate. We find more limited support for using EDSCLS to compare school climate between schools. Schools can confidently continue to use, or begin using, this freely available tool to assess school climate at the individual level and at the between-school level for the engagement and environment domains. However, given low-factor loadings for behavioral items in the CFA, and at the between-school level in the multilevel CFA, there should be continued refinement of the model, including separating student perceptions from student behavior, and investigations into differences in the model at the school and student levels.

ORCID iD

Deborah Temkin (iD) https://orcid.org/0000-0002-1998-2732

References

Asparouhov, T. (2008). *Scaling of sampling weights for two level models in Mplus 4.2*. http://www.statmodel.com/download/Scaling3.pdf

Bear, G., Yang, C., Mantz, L., Pasipanodya, E., Hearn, S., & Boyer, D. (2014). *Technical manual for Delaware School Survey: Scales of school climate, bullying victimization, student engagement, and positive, punitive, and social emotional learning techniques*. Delaware Positive Behavior Support (DE-PBS), Center for Disabilities Studies, and Delaware Department of Education.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Berger, R., Berman, S., Garcia, J., & Deasy, J. (2019). *A practice agenda in support of how learning happens*. Aspen Institute, National Commission on Social, Emotional, & Academic Development. http://nationathope.org/research-practice-and-policy-agendas/practice/

Bradshaw, C. P., Waasdorp, T. E., Debnam, K. J., & Johnson, S. L. (2014). Measuring school climate in high schools: A focus on safety, engagement, and the environment. *Journal of School Health*, *84*(9), 593–604. https://doi.org/10.1111/josh.12186

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (Vol. 154, pp. 136–162). Sage.

Carle, A. C. (2009). Fitting multilevel models in complex survey data with design weights: Recommendations. *BMC Medication Research Methodology*, *9*, Article 49. https://doi.org/10.1186/1471-2288-9-49

Chen, F., Bollen, P., Curran, P. J., & Kirby, J. B. (2001). Improper solutions in structural equation models. *Sociological Methods & Research*, *29*(2), 468–508. https://doi.org/10.1177/0049124101029004003

Consortium on Chicago School Research. (n.d.). *2007 Consortium survey measures*. University of Chicago.

Cornell, D. (2014). *Overview of the Authoritative School Climate Survey*. Curry School of Education, University of Virginia.

Cornell, D., & Huang, F. (2019). Collecting and analyzing local school safety and climate data. In M. J. Mayer, & S. R. Jimerson (Eds.), *School safety and violence prevention: Science, practice, and policy* (pp. 151–175). American Psychological Association. https://doi.org/10.1037/0000106-007

District of Columbia Public Schools. (n.d.). *DCPS at a glance: Enrollment*. https://dcps.dc.gov/page/dcps-glance-enrollment

Feldman, A. F., & Matjasko, J. L. (2005). The role of school-based extracurricular activities in adolescent development: A comprehensive review and future directions. *Review of Educational Research*, *75*(2), 159–210. https://doi.org/10.3102/00346543075002159

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, *9*(4), 466–491. https://doi.org/10.1037/1082-989X.9.4.466

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*(1), 72–91. https://doi.org/10.1037/a0032138

Government of the District of Columbia Office of the State Superintendent of Education. (2017). *Schedules of student enrollment and independent accountant's examination reports thereon*. https://osse.dc.gov/sites/default/files/dc/sites/osse/publication/attachments/2016-17%20School%20Year%20Enrollment%20Audit%20Report_0.pdf

Harrington, D. (2009). *Confirmatory factor analysis*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195339888.001.0001

Haynes, N. M., Emmons, C., & Ben-Avie, M. (1997). School climate as a factor in student adjustment and achievement. *Journal of Educational and Psychological Consultation*, *8*(3), 321–329. https://doi.org/10.1207/s1532768xjepc0803_4

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structural equation modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453. https://doi.org/10.1037/1082-989X.3.4.424

Huang, F. L., & Cornell, D. G. (2016). Using multilevel factor analysis with clustered data: Investigating the factor structure of the Positive Values Scale. *Journal of Psychoeducational Assessment*, *34*(1), 3–14. https://doi.org/10.1177/0734282915570278

Jakobsen, J. C., Gluud, C., Wetterslev, J., & Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomized clinical trials—A practical guide with flowcharts. *BMC Medical Research Methodology*, *17*(1), 162–172. https://doi.org/10.1186/s12874-017-0442-1

Jordan, P. W., & Hamilton, L. S. (2020). Walking a fine line: School climate surveys in state ESSA plans. *FutureEd.*

https://www.future-ed.org/wp-content/uploads/2020/01/FutureEdSchoolClimateReport.pdf

Konold, T. R., & Cornell, D. (2015). Measurement and structural relations of an authoritative school climate model: A multi-level latent variable investigation. *Journal of School Psychology*, *53*(6), 447–461. https://doi.org/10.1016/j.jsp.2015.09.001

Lippman, L., Guzman, L., & Moore, K. A. (2012). Measuring flourishing among youth: Findings from the flourishing children positive indicators project. *Child Trends*. https://www.childtrends.org/wp-content/uploads/2013/05/FlourishingChildren.pdf

Marsh, H. W., Hau, K., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's 1999 findings. *Structural Equation Modeling*, *11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

Musu, L., Zhang, A., Wang, K., Zhang, J., & Oudekerk, B. A. (2019). *Indicators of school crime and safety: 2018* (NCES 2019-047/NCJ 252571). National Center for Education Statistics, U.S. Department of Education, & Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. https://nces.ed.gov/pubs2019/2019047.pdf

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

National Center for Education Statistics. (2015). *ED School Climate Surveys (EDSCLS): National benchmark study 2016. Appendix D. EDSCLS Pilot Test 2015 Report* (ED575971). ERIC. http://files.eric.ed.gov/fulltext/ED577461.pdf

National Center for Education Statistics. (2016). *Education Department School Climate Survey (EDSCLS)*. https://safesupportivelearning.ed.gov/edscls/administration

National Center on Safe Supportive Learning Environments. (n.d.). *School climate survey compendium*. https://safesupportive-learning.ed.gov/topic-research/school-climate-measurement/school-climate-survey-compendium

Pedersen, A. B., Mikkelsen, E. M., Cronin-Fenton, D., Kristensen, N. R., Pham, T. M., Pedersen, L., & Petersen, I. (2017). Missing data and multiple imputation in clinical epidemiological research. *Clinical Epidemiology*, *9*, 157–166. https://doi.org/10.2147/CLEP.S129785

Rubin, D. B. (1987). *Introduction in multiple imputation for nonresponse in surveys*. John Wiley. https://doi.org/10.1002/9780470316696

Schweig, J. (2014). Cross-level measurement invariance in school and classroom environment surveys: Implications for policy and practice. *Educational Evaluation and Policy Analysis*, *36*(3), 259–280. https://doi.org/10.3102/0162373713509880

Steffgen, G., Recchia, S., & Viechtbauer, W. (2013). The link between school climate and violence in school: A meta-analytic review. *Aggression and Violent Behavior*, *18*(2), 300–309. https://doi.org/10.1016/j.avb.2012.12.001

Temkin, D., Belford, J., McDaniel, T., Stratford, B., & Parris, D. (2017). Improving measurement of sexual orientation and gender identity among middle and high school students. *Child Trends*. https://www.childtrends.org/wp-content/uploads/2017/06/2017-22LGBTSurveyMeasures-1.pdf

Temkin, D., & Harper, K. (2017, September 20). Some states are missing the point of ESSA's fifth indicator. *Child Trends*. https://www.childtrends.org/states-missing-point-essas-fifth-indicator

Temkin, D., Solomon, B. J., Katz, E., & Steed, H. (2019). Creating healthy schools: Students, educators, and policymakers name priorities. *State Education Standard*, *19*(1), 11–17.

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research*, *83*(3), 357–385. https://doi.org/10.3102/0034654313483907

Wang, M. T., & Degol, J. L. (2016). School climate: A review of the construct, measurement, and impact on student outcomes. *Educational Psychology Review*, *28*(2), 315–352. https://doi.org/10.1007/s10648-015-9319-1

Ye, C., & Wang, Y. (2018). *ED School Climate Surveys (EDSCLS): Psychometric benchmarking technical report*. National Center for Education Statistics, U.S. Department of Education. https://safesupportivelearning.ed.gov/sites/default/files/SCIRP/EDSCLS_Psychometric_Benchmarking_Technical_Report_2018-04-25.pdf

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*(2), 139–152. https://doi.org/10.1177/0734282909344205

Authors

RENEE RYBERG, PhD, is a research scientist in the Education Research Program Area at Child Trends. She studies conditions of learning that support youth as they make the transition from adolescence to adulthood.

SARAH HER, MA, is a research analyst at Child Trends. Her research interest is in analyzing and evaluating factors influencing academic and career readiness.

DEBORAH TEMKIN, PhD, is the vice president for youth development and education research at Child Trends. Her research focuses on the intersections between education policy and healthy social and emotional development.

REBECCA MADILL, PhD, is a research scientist in the Early Childhood Program Area at Child Trends. Her current research focuses on measuring and improving low-income families' access to high-quality child care, with a focus on the federal child care subsidy program.

CLAIRE KELLEY, MA, is a senior data scientist at Child Trends, where she conducts and supports research across all program areas. Her primary research interests focus on the intersection of machine learning and social science, particularly, in the domains of health and education.

JOY THOMPSON, PhD, is a research scientist with Child Trends. Her research interests include equitable access to educational opportunities and supports for students' entry and persistence into STEM pathways.

ALEXANDER GABRIEL, MPP, is a senior research analyst at Child Trends. He is a quantitative researcher, focusing on policies and programs that affect equity in schools.

APPENDIX TABLE A1

Multilevel Confirmatory Factor Analysis Models, by Domain

Engagement

	N	RMSEA	CFI	TLI	Notes
Model 1: Higher order factor model with factor loadings free to vary between levels	3,416	.03	.96	.95	Between level: *People of different cultural backgrounds, races, or ethnicities get along well at this school* (SENGCLC7) loads at .37. *I have lots of chances to be part of class discussions or activities* (SENGPAR48) loads on Participation at greater than 1.00 and has a negative residual variance. Cultural and Linguistic Competence and Relationships load on domain at greater than 1.00 and have negative residual variances.
Model 1a: Higher order factor model, dropping one-half of each pair of highly correlated items (>.90; SENGCLC3 and SENGREL12)	3,416	.03	.965	.95	Between level: *People of different cultural backgrounds, races, or ethnicities get along well at this school* (SENGCLC7) loads at .44. *I have lots of chances to be part of class discussions or activities* (SENGPAR48) loads on Participation at greater than 1.00 and has a negative residual variance. Relationships loads on the domain at greater than 1.00 and has a negative residual variance. Two pairs of items have cross-subdomain correlations of greater than .90.
Model 1c: Higher order factor model at the within level, single factor at the between level	3,416	.03	.96	.95	Between level: *People of different cultural backgrounds, races, or ethnicities get along well at this school* (SENGCLC7) loads at .37.
Model 1d: Higher order factor model at the within level, single factor at the between level, dropping SENGCLC7 at both levels	3,416	.03	.96	.95	Between level: *Students like one another* (SENGREL21) loads at .47. Note that this model is identical to Model 3a, as it is just identified with three factors. **This is the final model presented in the text.**
Model 2: Correlated factor model with subdomains at both levels	3,416	.03	.96	.95	Within level: Subconstructs are correlated from .73 to .87. Between level: Subconstructs are correlated from .94 to 1.05. *People of different cultural backgrounds, races, or ethnicities get along well at this school* (SENGCLC7) loads at .37 and *I have lots of chances to be part of class discussions or activities* (SENGPAR48) loads at greater than 1.00 and have negative residual variances. Correlations between subconstructs range from .94 to .98, but Cultural and Linguistic Competence and Relationships are correlated at greater than 1.00.
Model 2a: Correlated factor model with subdomains at both levels, dropping one-half of each pair of highly correlated items (>.90; SENGCLC3 and SENGREL12)	3,416	.03	.956	.95	Within level: Correlations between subconstructs range from .67 to .85. Between level: Correlations between subconstructs range from .78 to 1.05. *People of different cultural backgrounds, races, or ethnicities get along well at this school* (SENGCLC7) loads at .37 and *I have lots of chances to be part of class discussions or activities* (SENGPAR48) loads at greater than 1.00 and have negative residual variances. Correlations between subconstructs range from .78 to 1.05, with correlations between Relationships and Cultural and Linguistic Competence and REL and Participation being greater than 1.00.
Model 3: Correlated factor model with subdomains at individual levels and a single factor at the between level	3,416	.03	.96	.95	Within level: Correlations between subconstructs range from .73 to .87. Between level: *People of different cultural backgrounds, races, or ethnicities get along well at this school* (SENGCLC7) loads at .37.
Model 3a: Correlated factor model with subdomains at the individual level and a single factor at the between level. Removing SENGCLC7 at both levels	3,416	.03	.96	.95	Within level: Correlations between subconstructs range from .70 to .83. Between level: *Students like one another* (SENGREL21) loads at .47.

(continued)

Safety

	N	RMSEA	CFI	TLI	Notes
Model 1: Higher order factor model with factor loadings free to vary between levels	3,416	.03	.95	.95	Between level: *I feel safe going to and from this school* (SSAFPSAF60) loads at .45 and *Students at this school carry guns or knives to school* (SSAFPSAF65) loads at .46. Conversely, *Students at this school try to stop bullying* (SSAFBUL80) and *Students at this school think it is okay to smoke one or more packs of cigarettes a day* (SSAFSUB92) load at greater than 1.00 and have negative residual variances. Physical Safety loads on the domain at greater than 1.00 and has negative residual variance. Substance Abuse loads on the domain at .17.
Model 1a: Higher order factor model, dropping one-half of each pair of highly correlated items (>.95; SSAFBUL80)	3,416	.02	.97	.96	Between level: I feel safe going to and from this school (SSAFPSAF60) loads at .45 and Students at this school carry guns or knives to school (SSAFPSAF65) loads at .46. Conversely, Students at this school try to stop bullying (SSAFBUL80) and Students at this school think it is okay to smoke one or more packs of cigarettes a day (SSAFSUB92) load at greater than 1.00 and have negative residual variances. Physical Safety loads on the domain at greater than 1.00 and has negative residual variance. Substance Abuse loads on the domain at .15.
Model 1b: Higher order factor model, dropping one-half of each pair of highly correlated items (>.90; SSAFBUL80, SSAFEMO0, SSAFBUL1, SSAFPSAF69)	3,416	.02	.97	.96	Between level: Students at this school try to stop bullying (SSAFBUL80) loads at greater than 1.00 and have negative residual variance. Physical Safety loads on the domain at greater than 1.00 and has negative residual variance. Substance Abuse loads on the domain at .17.
Model 1c: Higher order factor model at the within level, single factor at the between level, with substance abuse items removed					No convergence.
Model 1d: Higher order factor model at the within level, single factor at the between level, with substance abuse items removed at between level only					No convergence.
Model 2: Correlated factor model with subdomains at both levels	3,416	.02	.96	.95	Within level: Correlations between subconstructs range from |.61| to .78. Between level: Correlations between subconstructs range from |.09| to |.99|. I feel safe going to and from this school (SSAFPSAF60) loads at .45 and *Students at this school carry guns or knives to school* (SSAFBUL80) loads at .45. Conversely, *Students at this school try to stop bullying* (SSAFBUL80) and *Students at this school think it is okay to smoke one or more packs of cigarettes a day* (SSAFSUB92) load at greater than 1.00 and have negative residual variances. Correlations between subconstructs range from −.99 to .90. Substance abuse is barely correlated with the others (−.09 to .25), but Emotional Safety and Physical Safety are correlated at −.99.
Model 2a: Correlated factor model with subdomains at both levels, dropping one-half of each pair of highly correlated items (>.90; SSAFBUL8, SSAFEMO0, SSAFBUL1, SSAFPSAF69)	3,416	.02	.97	.96	Within level: Correlations between subconstructs range from |.37| to .74. Between level: Correlations between subconstructs range from |.81| to |.99|. Substance abuse is barely correlated with the others (from −.08 to .26), but Emotional Safety and Physical Safety are correlated at −.99.

(continued)

Safety

	N	RMSEA	CFI	TLI	Notes
Model 2b: Correlated factor model with subdomains at both levels, excluding substance abuse	3,416	.03	.94	.93	Within level: Correlations between subconstructs range from |.61| to .78.
					Between level: Correlations between subconstructs range from |.81| to |.99|. *I feel safe going to and from this school* (SSAFPSA0) loads at .44 and *Students at this school carry guns or knives to school* (SSAFPSA2) loads at .42. Conversely, *Students at this school try to stop bullying* (SSAFBUL8) load at greater than 1.00, and has a negative residual variance. Correlations between subconstructs range from −.99 to .90. Emotional Safety and Physical Safety are correlated at −.99.
Model 3: Correlated factor model with subdomains at individual levels and a single factor at the between level	3,416	.02	.96	.95	Within level: Correlations between subconstructs range from |.37| to .78.
					Between level: *I feel safe going to and from this school* (SSAFPSA0) loads at .45 and *Students at this school carry guns or knives to school* (SSAFPSA2) loads at .46. Conversely, *Students at this school try to stop bullying* (SSAFBUL8) loads at greater than 1.00 and has a negative residual variance.
Model 3a: Correlated factor model with subdomains at the individual level and a single factor at the between level. Removing substance abuse items.	3,416	.03	.94	.94	Within level: Correlations between subconstructs range from |.61| to .78.
					Between level: *I feel safe going to and from this school* (SSAFPSA0) loads at .43 and *Students at this school carry guns or knives to school* (SSAFPSA2) loads at .42.

Environment

	N	RMSEA	CFI	TLI	Notes
Model 1: Higher order factor model with factor loadings free to vary between levels	3,416	.02	.98	.98	Between level: *Broken things at this school get fixed quickly* (SENVPENV107) and *My teachers often connect what I am learning to life outside the classroom* (SENVINS114) load greater than 1 and have negative residual variances with SENVPENV107's negative residual variance being fairly large (−.30). *The things I'm learning in school are important to me* (SENVINS115) loads at .43 and *Adults working at this school reward students for positive behavior* (SENVDIS143) loads at .48. Instructional Environment and Discipline load on the domain at greater than 1.00 and have negative residual variances.
Model 1a: Higher order factor model with factor loadings free to vary between levels, dropping one-half of each pair of highly correlated items (>.90; SENVDIS143, SENVINS114, SENVMEN130, SENVDIS147)	3,416	.02	.98	.97	Within level: Discipline load on the domain at greater than 1.00.
					Between level: *Broken things at this school get fixed quickly* (SENVPENV107), *Discipline is fair* (SENVDIS147C), Instructional Environment and Discipline have factor loadings greater than 1 and negative residual variances (3 of them greater than .10 in magnitude). *My teachers praise me when I work hard in school* (SENVINS111) loads at .46. *The things I'm learning in school are important to me* (SENVINS115) loads at .24 and *My teachers expect me to do my best all the time* (SENVINS121) loads at.46.

(continued)

Environment

	N	RMSEA	CFI	TLI	Notes
Model 1b: Higher order factor model at the within level, single factor at the between level	3,416	.02	.98	.98	Between level: *The things I'm learning in school are important to me* (SENVINS115) loads at .39 and *Adults working at this school reward students for positive behavior* (SENVDIS143) loads at .45.
Model 1c: Higher order factor model at the within level, single factor at the between level, dropping SENVINS115 and SENVDIS143.	3,416	.02	.98	.97	**This is the final model presented in the text.**
Model 2: Correlated factor model with subdomains at each level	3,416	.02	.98	.98	Within level: Correlations between subconstructs range from .64 to .90.
					Between level: Correlations between subconstructs range from .51 to 1.11. *Broken things at this school get fixed quickly* (SENVPENV107) loads at greater than 1.00 and has a negative residual variance. *The things I'm learning in school are important to me* (SENVINS115) loads at .45 and *Adults working at this school reward students for positive behavior* (SENVDIS143) loads at .49. Correlations between subconstructs range from .51–.91, but Instructional Environment and Discipline are correlated at 1.11.
Model 2a: Correlated factor model with subdomains at each level, dropping one-half of each pair of highly correlated items (>.90; SENVDIS143, SENVINS114, SENVMEN130, SENVDIS147)	3,416	.02	.98	.97	Within level: Correlations between subconstructs range from .63 to .96.
					Between level: Correlations between subconstructs range from .58 to 1.44. *Broken things at this school get fixed quickly* (SENVPENV107) and has a negative residual variance. Instructional Environment has a factor loading greater than 1.00. *The things I'm learning in school are important to me* (SENVINS115) loads at .30, *My teachers expect me to do my best all the time* (SENVINS121) loads at .491. Correlations between subconstructs range from .78–.94, but Instructional Environment and Discipline are correlated at greater than 1.00.
Model 3: Correlated factor model with subdomains at individual levels and a single factor model at the between level	3,416	.02	.98	.98	Within level: Correlations between subconstructs range from .64 to .90.
					Between level: *The things I'm learning in school are important to me* (SENVINS115) loads at .39 and *Adults working at this school reward students for positive behavior* (SENVDIS143) loads at .44.
Model 3a: Correlated factor model with subdomains at the individual level and a single factor model at the between level, dropping SENVINS115 and SENVDIS143.	3,416	.02	.98	.97	Within level: Correlations between subconstructs range from .62 to .89.

Note. This table presents the main series of models that were tested in each domain. Additional models were tested that, for example, dropped items or fixed negative residual variances to zero. They are not presented here for the sake of parsimony. All models are based on multiply imputed data and exclude high school–only items as well as the behavioral items that were ill-fitting in the CFA. CFA = confirmatory factor analysis; CFI = comparative fit index; RMSEA = root mean square error of approximation; TLI = Tucker–Lewis index.