School Climate **IMPROVEMENT** | Resource Package

Data Interpretation Guide
on Making School Climate Improvements

JANUARY 2017

National Center on Safe Supportive Learning Environments

Safe Supportive Learning
Engagement | Safety | Environment

Disclaimer

This *Data Interpretation Guide* was designed and written under the U.S. Department of Education (Department) Contract Number EDESE12O0035 by American Institutes for Research, the National Center on Safe Supportive Learning Environments (NCSSLE). Rita Foy Moss served as the contracting officer's representative (COR) for the NCSSLE technical assistance center. This document contains resource materials that are provided for the user's convenience. The inclusion of these materials is not intended to reflect their importance, nor is it intended to endorse any views expressed, or products or services offered. These materials may contain the views and recommendations of various subject matter experts as well as hypertext links, contact addresses, and websites to information created and maintained by other public and private organizations. The opinions expressed in any of these materials do not necessarily reflect the positions or policies of the Department. The Department does not control or guarantee the accuracy, relevance, timeliness, or completeness of any outside information included in these materials.

January 2017

Availability of Alternate Formats

Requests for documents in alternate formats such as Braille or large print should be submitted to the Alternate Format Center by calling 202.260.0852 or by contacting the 504 coordinator via e-mail at om_eeos@ed.gov.

Notice to Limited-English-Proficient Persons

If you have difficulty understanding English, you may request language assistance services for Department information that is available to the public. These language assistance services are available free of charge. If you need more information about interpretation or translation services, please call 1-800-USA-LEARN (1.800.872.5327) (TTY: 1.800.877.8339) or e-mail us at ED.Language.Assistance@ed.gov. You also can write to U.S. Department of Education, Information Resource Center, LBJ Education Building, 400 Maryland Ave. SW, Washington, DC, 20202.

Contents

Introduction

Congratulations! You have made an important commitment to improve the climate in your state, district, or school.[1] Measuring and understanding how students, staff, and parents/guardians perceive the climate in their schools are key steps in making wise decisions on how to use resources to focus on areas in need of improvement. This document contains detailed information and resources to help you interpret and use results from the school climate surveys you have administered.

> **Who should read this guide?** The person(s) at your state, district, or school who will be most actively involved in the interpretation of the survey findings, who will translate the data into action, and who will communicate with those who will promote data-driven decision making for school climate improvement should read this guide. The information is technical in nature. Thus, if you or your team does not have technical expertise in interpreting data, you may wish to use this document as a reference while using the EDSCLS Data Analysis Worksheet instead.

We used the school climate model developed by the U.S. Department of Education (ED) to develop the ED School Climate Surveys (EDSCLS).[2] The model includes three key school climate domains composed of 13 subdomains or topic areas, as shown in Figure 1. Whether you used the EDSCLS or another survey that includes constructs similar to those in this model, this guide is for you! If you used the EDSCLS surveys, look for the logo to the right throughout to see information specific to them.

Figure 1. Model of School Climate: Domains and Topic Areas

Domain: Engagement	Domain: Safety	Domain: Environment
Topics: Cultural and linguistic competence	**Topics:** Emotional safety	**Topics:** Physical environment
	Physical safety	Instructional environment
Relationships	Bullying/cyberbullying	Physical health
	Substance abuse	Mental health
School participation	Emergency readiness/management	Discipline

---
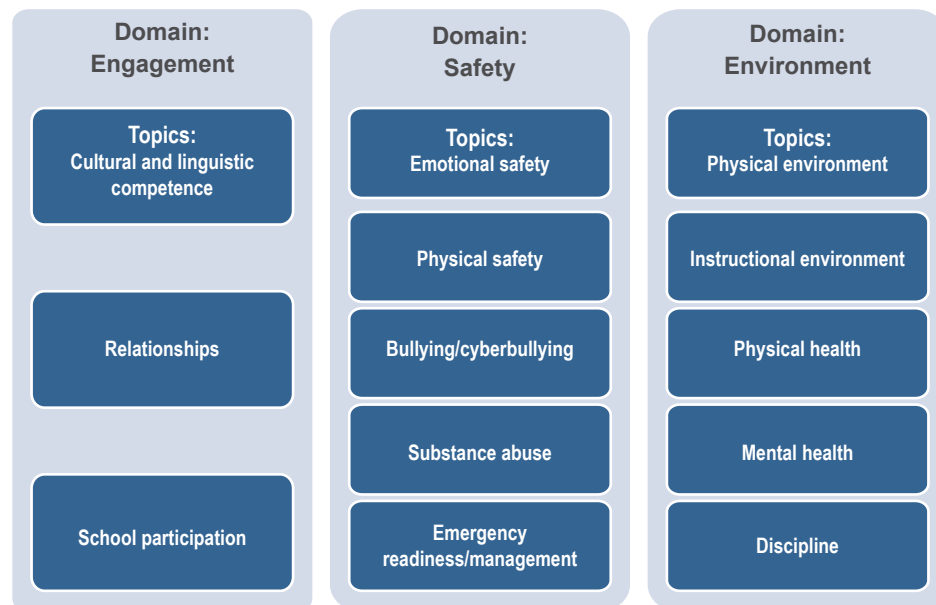
[1] This document provides strategies applicable to public schools and districts, including charter authorizers, charter management organizations, education management organizations, individual charter schools, and charter local educational agencies.

[2] The EDSCLS surveys and Web-based platform were developed by the National Center for Education Statistics in collaboration with the Office of Safe and Healthy Students (OSHS), both part of ED.

As shown in Figure 1, the domain of Engagement comprises several actionable topic areas—Cultural and Linguistic Competence, Relationships, and School Participation—while the other two domains, Safety and Environment, comprise actionable topic areas of their own.

Sites using the EDSCLS should note three things about this model and related results:

1. **Physical Health** is shown in the model as one of the topic areas within the Environment domain. However, although the data for instructional and non-instructional staff formed a Physical Health scale that can be used for analytical purposes, they did not form a Physical Health scale for students.

2. The data for the **parent/guardian survey** did not form a scale at the topic area or domain level because of the brevity of the survey. Thus, parent data should be examined at only the item level (i.e., analysis of individual survey questions), explained more fully below.

3. Although **Emergency Readiness/Management (ERM)** is shown as one of the topic areas within the Safety domain, it was not designed to form a scale; thus ERM data should be examined at only the item level.

Interpreting Survey Results

If you use the EDSCLS platform, it will generate scale scores and item-level data for you immediately after the close of data collection. Beginning in fall 2017, the EDSCLS platform also will integrate data from a benchmarking study. This guide will be updated to reflect those changes.

If you administered a survey other than the EDSCLS, you can use this guide as you interpret and use your survey results. Although the survey you administered may not use the same labels to identify topics as does the EDSCLS model (e.g., "emotional safety" or "cultural and linguistic competence"), many school climate surveys include constructs that are similar in composition and, most importantly, are actionable. You will find this guide and its link to a *Discussion Guide* for each topic area (along with other helpful resources) especially useful when you are considering appropriate actions your educational agency can take to improve climate based on your data.

Using EDSCLS data for multiple purposes: Schools and districts are also required to report information about school climate pursuant to the U.S. Department of Education's Civil Rights Data Collection (CRDC). Information collected by this survey may help schools and districts prepare their responses to the CRDC survey. More information about the CRDC can be found at ocrdata.ed.gov.

Discussion Guides and Additional Resources

At the end of each of the following sections, you will find links to a *Discussion Guide* for each topic area. It includes a brief description of the topic area and initial and deeper guiding questions that you can use to examine and interpret your results for that topic. The *Discussion Guide* also contains

links to webpages with interventions that you can begin to implement immediately and use as a foundation for further work as well as links to longer term intervention resources.

A Multi-Tiered Approach to Improving School Climate

The directions and resources in this document are presented using a "multi-tiered" approach. A multi-tiered system comprising various levels of support (universal, targeted, and intensive) can be conceptualized as a triangle, or a continuum within a triangle, as shown in Figure 2.[3]

Figure 2. Multi-Tiered System of Support: A Continuum Within a Triangle

Continuum of Support for All Students
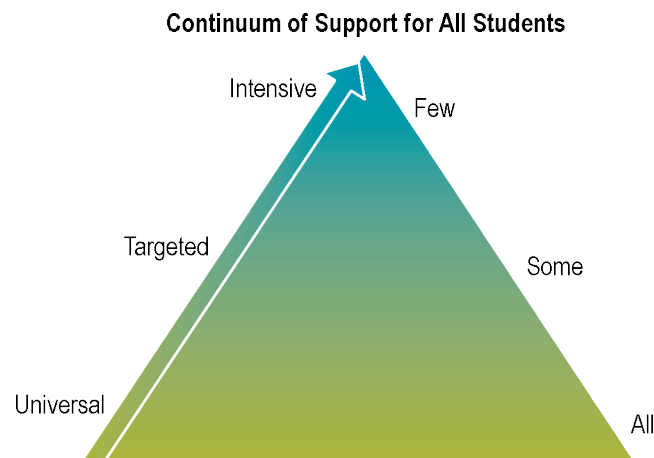
As illustrated in Figure 2, all students are served in a multi-tiered system, with more general support at the base, becoming gradually more intensive advancing up the triangle. The base, which is the largest tier, illustrates that universal interventions and supports are provided schoolwide, to all students and other stakeholders in the school community. The middle tier, which comprises a smaller area, suggests that interventions and supports are provided to a subset of those in need of more targeted help (often provided in small-group sessions). The top tier, which comprises the tip of the triangle, illustrates that interventions and supports are delivered to those in need of more intensive help (usually provided on a one-to-one basis).

Following this section, you will find information about interpreting your data as well as links to *Discussion Guides* that will help you use data in a universal or targeted way to improve school climate in specific topic areas. Click on the School Climate Improvement Resource Package (SCIRP) *Reference Manual* for additional information about multi-tiered systems of support. More information also is available on websites such as www.pbis.org/ and www.rtinetwork.org/.

Supports should be designed to improve school climate for the students who are most in need regardless of the subgroup(s) to which they belong. Targeting supports based on need as opposed to membership in a subgroup will support compliance with relevant civil rights laws.

---

[3] Office of Special Education Programs and OSHS Technical Assistance Center on Positive Behavior Interventions and Supports. (2014). *Multi-tiered behavior frameworks: Developing implementation capacity* [PowerPoint slides]. Retrieved from https://www.pbis.org/Common/Cms/files/pbisresources/mtbf_overview_sctg_28oct2014_GSugai.pptx

Interpreting Your Data

As you begin to review your survey results, you may feel overwhelmed and like you are "drowning in data." If so, you may want to look at your survey results one step at a time, perhaps in four or five "chunks," by respondent group (students, instructional staff, noninstructional staff, parents/guardians) or by school climate topic areas (e.g., physical safety, relationships). For example, EDSCLS users can use the EDSCLS Data Analysis Worksheet. However you decide to tackle your survey results, you can use the framework presented here as a guideline.

Contents of This Guide

Here you will find descriptions of and suggestions for interpreting results from different types of data:

- Scale scores (overall and by respondent characteristics),
- Item-level data, and
- Average (mean) topic area values (overall and by respondent characteristics).[4]

In each section, we present a brief description of the data type.

For EDSCLS users, we provide a link to example do's and don'ts when examining the results and making comparisons, and how those comparisons can help you improve climate in your district or school.

If you are a **district or local educational agency (LEA)** hosting the survey administration, you will find direction on interpreting your district data, comparing across schools in your district, and comparing individual school data to that of your district.

If you are one of a number of **schools** in a district that is hosting the survey administration, you will find suggestions for interpreting your school data, and how to compare them to your district if those data are available.

- If you are a **state educational agency (SEA)** hosting the survey administration, you can follow the support provided to districts in this guide, plus you can examine district-to-district comparisons as well as school-to-school comparisons.

At the end of each section, you will be able to link to universal/schoolwide or targeted topic area *Discussion Guides* that provide links to webpages and resources for interventions and strategies for each of the school climate topic areas.

> ### S ARE YOU A SINGLE SCHOOL HOSTING A SCHOOL CLIMATE SURVEY?
>
> Comparisons throughout this document focus on school-to-school and school-to-district comparisons. If you are a single school hosting the EDSCLS or another survey, you will not have other school and district data for comparison purposes. We recommend that you focus instead on the item-level data comparisons and respondent subgroup comparisons (see the section Interpreting Item-Level Data: Direction, Support, and Resources for Schools in Appendix B).

---

[4] Average (mean) topic area values have been calculated for and reported to EDSCLS pilot sites. EDSCLS users who administer surveys after the pilot may find it helpful to calculate and use these values. See Appendix C for further information on how to calculate and use average (mean) topic area values.

Scale Scores

Scale scores are the premier way that EDSCLS as well as many other school climate surveys measure school climate. A scale score, which combines multiple survey items related to different aspects of a topic area, is a more robust measure than attempting to measure that topic by asking about it with a single item.

For sites administering EDSCLS, scale scores are produced for each topic area as a number on a scale of 100 to 500. These scores can be thought of in the same way you might think about an SAT score or a credit score, in that higher numbers represent more positive results, but the real meaning is in how your score compares with others. Scale scores are produced by the platform overall for each respondent group and by respondent characteristics (e.g., male students and female students).

Districts will be able to view their scale scores at the district level or at the school level for any school that participated in the district's survey administration.

Schools will be able to see their own scale scores as well as those from their district, if the district makes them available to schools.

Overall Scale Scores

For each respondent group, overall scale scores are presented in a chart similar to the example in Figure 3.

This example shows that the average student scale score for this school for the Discipline topic area is 323 (on a scale of 100–500).[5]

Figure 3. Example of Overall Student Scale Score Chart for Discipline Topic Area

_____
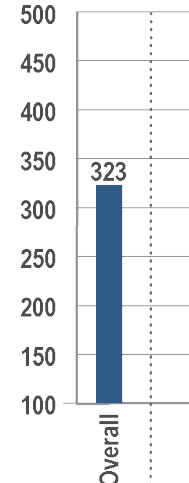
[5] Benchmarking data will be available in fall 2017. At that time, EDSCLS users will be able to use benchmark data to facilitate the meaning of their scale scores.

Scale Scores by Respondent Characteristics

Scale scores broken out by respondent subgroups provide a richer set of data, a way to see how perceptions of a topic area differ across subgroups of students and staff.

Scale scores are produced for EDSCLS users for the following subgroups:[6]

- Student scale scores per topic area can be examined by:
  - Gender,
  - Race/ethnicity, and
  - Grade.
- Staff (instructional and noninstructional) scale scores per topic area can be examined by:
  - Gender and
  - Race/ethnicity.

In the event of a possible disclosure risk that would allow a respondent or small subgroup of respondents to be identified (e.g., if there is only one Asian male teacher in the school), the EDSCLS platform will suppress the results for that subgroup (i.e., results for that subgroup will not be shown). If the data are suppressed (and thus, not shown) for a respondent subgroup that is important to your district or school, there are other ways that you can collect overall school climate or specific topic area data. For example, you may want to convene focus groups or conduct targeted interviews or student "fishbowls[7]." Click on the SCIRP *Reference Manual* to go to additional information about these methods of collecting data. (You also will have a chance to click on *Discussion Guides* to go to specific resources in topic areas of interest in Figure 5.)

For each respondent group, scale scores for subgroups by respondent characteristics can be viewed in a chart similar to the example in Figure 4.

Figure 4. Example of Student Scale Score Chart for Discipline Topic Area, Overall and by Gender
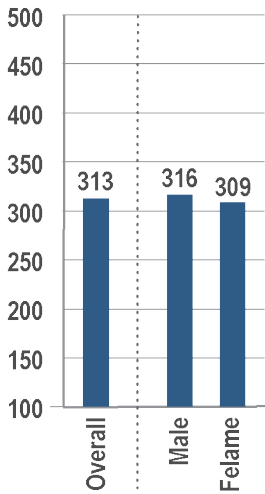
Interpreting Scale Scores

Appendices A (for districts) and B (for schools) include suggestions for interpreting overall scale scores and scale scores by respondent characteristics. They also include comparisons that can and cannot be made using scale scores (do's and don'ts) and how those comparisons can help you improve school climate.

If you are a **district**, click on Appendix A, Table A-1, to go to interpretation of scale scores, overall and by respondent characteristics.

If you are a **school**, click on Appendix B, Table B-1, to go to interpretation of scale scores, overall and by respondent characteristics.

---

[6] Note that the EDSCLS platform does not produce crossed demographics (e.g., Asian female students).

[7] A "Fishbowl" is a discussion style in which one group (students) has a dialogue, and another group (school staff) listens to their discussion. Afterwards both groups come together to discuss themes to work on together.

This example shows that the average student scale score for Discipline in this school is 313 (on a scale of 100–500), while the scale score for male students is 316 and the score for female students is 309.

Discussion Guides for Interpreting Topic Area Scale Scores

You can click on any of the topic area labels in Figure 5 to go to that topic area's *Discussion Guide* for using scale scores. If you want to see the *Discussion Guide* for *universal/schoolwide* support for a specific topic area, go to "Overall" in the topic area box. It will guide you through interpreting the overall scale score for that topic area. If you want to see the *Discussion Guide* for *targeted* support, go to "By respondent characteristics," which guides you through interpreting scale scores by respondent characteristics for the topic area.

> **?** **WHAT IS IN EACH DISCUSSION GUIDE?**
>
> You will find a brief description of the topic area as well as a set of guiding questions and links to topic area webpages with shorter and longer term intervention resources pertinent to that topic area.

Figure 5. Links to Topic Area Discussion Guides for Using Overall Scale Scores or by Respondent Characteristics[8,9]

Domain: Engagement	Domain: Safety	Domain: Environment
Topics: Cultural and linguistic competence Overall by Respondent Characteristics	**Topics: Emotional safety** Overall by Respondent Characteristics	**Topics: Physical environment** Overall by Respondent Characteristics
	Physical safety Overall by Respondent Characteristics	**Instructional environment** Overall by Respondent Characteristics
Relationships Overall by Respondent Characteristics	**Bullying/cyberbullying** Overall by Respondent Characteristics	**Physical health** Overall by Respondent Characteristics
	Substance abuse Overall by Respondent Characteristics	**Mental health** Overall by Respondent Characteristics
School participation Overall by Respondent Characteristics	**Emergency readiness/management** Overall by Respondent Characteristics	**Discipline**

---

[8] ⚠ For EDSCLS users, no scale is produced for the emergency readiness and management topic. However, percentage distributions and item averages (means) are produced by the Web-based platform and, for pilot sites, appear in the District or School Report PDF.

[9] ⚠ For EDSCLS users, no physical health scale is produced for students; however, it is produced for instructional and noninstructional staff.

Item-Level Data

After scale scores have been examined, looking at item-level data may help a district or school dig deeper into the data to target specific areas or issues. Item-level results can often provide schools with concrete examples of the underlying topic area construct that may be more actionable, warranting more immediate implementation of interventions included in the topic area webpages. Item-level results can also suggest areas in which to begin planning and preparation for longer term interventions and strategies.

Item-level values reflect the EDSCLS response options (values of 1–4). Sites will receive the percentage distribution of those response options and their average (mean) for each item in the survey that is included in topic area scales produced by the platform. Item-level results will also be provided for important items that are not included in the final topic area scales. They will be marked for you as "stand-alone" items in the platform and, for pilot sites, they are included in your District or School Report PDF.

Item-level values reflecting the EDSCLS response options (values of 1–4) can be viewed in a chart like the example in Figure 6.

Figure 6. Example of Item Percentage Distribution and Mean Chart

This example shows that when presented with the item "Teachers understand my problems," 13.64 percent reported that they strongly agree (value 4), 50 percent reported that they agree (value 3), 27.27 percent reported that they disagree (value 2), and 9.09 percent reported that they strongly disagree (value 1). The average (mean) of this item's values, calculated as the average of all responses to the item (2.68), also is provided in the right column of the chart as displayed earlier.

Comparing Item Averages (Means)

The example in the chart is a positively valenced item ("Teachers understand my problems"). "Valenced" in a positive direction means that agreement with the statement is a positive perception. Out of all of the students who responded to this item, the average response was 2.68 (where *strongly agree* = 4 and *strongly disagree* = 1), indicating that students' perceptions, on average, were more positive than negative. (Click on Figure 8 for further information on interpreting item favorability.)

However, some items may be negatively valenced (e.g., "I sometimes stay home because I don't feel safe at this school."). "Valenced" in a negative direction means that agreement with the statement is a negative perception. Because the numeric values of both negative and positive items are on the same scale of *strongly agree* = 4 and *strongly disagree* = 1, the item means for negatively and positively valenced items should not be directly compared without first reverse coding the negative items.

For EDSCLS users, these negatively valenced items are marked for you with a footnote in the platform and, for pilot sites, in your District or School Report PDF.

Reverse Coding Negatively Valenced Items

Negatively valenced items have been reverse-coded for the purpose of calculating scale scores, *but the item-level results that you see are not reverse-coded*. If you want to compare item averages (means) across all items or a group of items (both positively and negatively valenced), you should reverse-code the values of negatively valenced items so that they are directed in the same way as positively valenced items. To reverse-code those items, use the following formula:

Reverse-Coded Item Mean = [5 - item mean]

For example, if the mean of a negatively valenced item is 3.4, the reverse-coded mean would be calculated as [5 - 3.4 = 1.6].

After negatively valenced items are reverse-coded, you can compare them with means of positively valenced items.

Interpreting Item-Level Data

Appendices A (for districts) and B (for schools) provide suggestions for using item-level data to assess how favorably (positively) or unfavorably (negatively) respondents perceive items. This information can be used by both EDSCLS and other survey users.

If you are a **district**, click Appendix A, Table A-2, to go to interpretation of item-level data.

If you are a **school**, click Appendix B, Table B-2, to go to interpretation of item-level data.

These tables include information to help you interpret item-level data, about how:

1. An item is perceived by individual respondent groups;

2. An item is perceived across respondent groups, but only for items *worded exactly the same* for each group (called "comparable items"); and

3. Items organized by content within a topic area (called "item-content groups") are perceived by individual respondent groups.

These tables also provide suggestions for districts and schools on how item-level data can be used to improve school climate, and limitations to be aware of when using item-level data.

Discussion Guides: Targeted Support for Using Item-Level Data

You can click on any of the topic areas in Figure 7 to go to that topic area's *Discussion Guide* for suggestions for using item-level data. You will find a brief description or literature review of the topic as well as guiding questions and links to webpages that include resources and strategies to help you improve school climate.

Figure 7. Links to Topic Area Discussion Guides for Using Item-Level Data

Domain: Engagement	Domain: Safety	Domain: Environment
Topics: Cultural and linguistic competence	**Topics:** Emotional safety	**Topics:** Physical environment
Relationships	Physical safety	Instructional environment
	Bullying/cyberbullying	Physical health
	Substance abuse	Mental health
School participation	Emergency readiness/management	Discipline
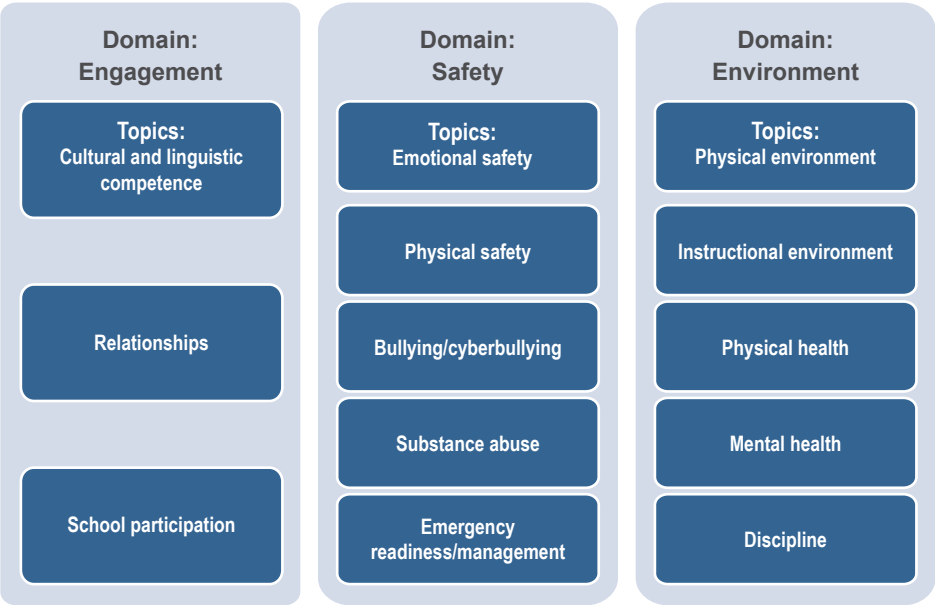
Conclusion

Your commitment to using a data-driven approach to improving school climate is a key first step in your journey toward a positive educational environment for students, staff, and parents. In this document, we have presented information and resources for interpreting and using results of survey data to identify areas for improvement and to work toward providing universal/schoolwide support (using overall scale scores and item-level data). We also have provided assistance on interpreting and using data to help you identify specific groups within the school community who may need more targeted support using scale scores parsed out by respondent characteristics. In addition, we have included links to a *Discussion Guide* for each of 13 school climate topics that include (1) an introduction; (2) guiding questions for you to consider as you work through your data; and (3) links to webpages with resources that you can implement immediately and in the longer term. In Appendix C, we present additional information about calculating, using and interpreting average (mean) topic area values—for pilot sites that received such results in their District and School report PDFs—and for other users who are interested in using these types of data.

Appendix A. Comparisons for Districts

Interpreting Scale Scores: Direction, Support, and Resources for EDSCLS Districts

Table A-1 provides direction, support, and resources to **districts** on interpreting scale scores overall and by student and staff characteristics. It includes comparisons that can and cannot be made using scale scores (do's and don'ts), how these comparisons can help districts improve school climate, and cautions or limitations when interpreting the data. Scale scores may be difficult to interpret by themselves since scores across scales may not mean the same things (e.g., a score of 300 on the physical safety scale may not be comparable to a score of 300 on the emotional safety scale) in terms of what excellent schools look like. For sites using the EDSCLS, benchmark data will be available in the platform in fall 2017 to facilitate the meaning of scale scores.

Table A-1. Scale Scores: Do's and Don'ts for Student and Staff Data

Do Compare:

- The district's scale scores for topic areas within domains (i.e., "within-domain topic areas").

  Example: Within the Engagement domain, you can compare scores for its topic areas: Cultural and Linguistic Competence (CLC), Relationships, and Participation.

- Each school's overall scale score for each domain or each topic area with the district's overall score.

  Example: You can compare a school's score on Engagement with the district's score on that same domain.

- Scale scores produced for each domain or topic area across schools in the district.

  Example: You can compare scores for all district schools with each other for each of the 12 topic areas.

- The district's scale scores for each domain and topic area between subgroups of respondents.

  Example: You can compare district scores for female students versus male students on Physical Safety (a topic area) or on Safety (a domain).

- The district's scale scores for within-domain topic areas for a single subgroup of respondents.

  Example: You can compare district scores for Black or African-American students on Emotional Safety, Physical Safety, Bullying/Cyberbullying, and Substance Abuse within the Safety domain.

- Each school's scale scores by respondent characteristics for domains or topic areas with the district score(s).

  Example: You can compare each school's score with the district score for female students on CLC.

- Scale scores by respondent characteristics for domains/topic areas across schools in the district.

  Example: You can compare each school's score with other schools' scores for male staff on CLC.

Do Not Compare:

● Scale scores for domains/topic areas across respondent groups.

Example: Scores on CLC for students should not be compared with CLC scores for staff.

● Scale scores for subgroups on topic areas across domains.

Example: Scores for students on the Physical Safety topic area (within the Safety domain) should not be compared with the student score for the Relationships topic area (within the Engagement domain).

● Scale scores by respondent characteristics for domains/topic areas for subgroups across respondents.

Example: Scores on CLC for White students should not be compared with CLC scores for White staff.

● Scale scores by respondent characteristics for topic areas across domains.

Example: Scores for Hispanic noninstructional staff on the Physical Safety topic area (within the Safety domain) should not be compared with Hispanic noninstructional staff's score for the Relationships topic area (within the Engagement domain).

Do Use These Comparisons To:

● Provide universal/schoolwide support across the district to students and staff for topic areas that have low scores relative to other topics within the same domain.

Example: If the district's score on student perceptions of CLC is low compared with student scores on other Engagement topics (Relationships and Participation), the district can focus supports on improving CLC for students in schools across the district.

● Provide universal/schoolwide supports to schools for topic areas that have low scores relative to the district score.

Example: For schools that have staff scores on CLC that are low compared with the district's score on CLC, the district can work with those staff in schools on improving CLC.

● Provide universal/schoolwide supports to schools for topic areas that have low scores relative to other schools in the district.

Example: For schools that have lower staff scores on CLC than other schools in the district, the district can help those schools improve conditions related to CLC for staff.

● Provide targeted support across the district to subgroups of students and staff for topic areas that have low scores relative to that subgroup's score on other topics within the same domain.

Example: If the district's score for male students on CLC is low compared with male students' scores on other Engagement topics (Relationships and Participation), the district can focus supports on improving CLC for boys in the district.

● Provide targeted support on topic areas to schools that have low scores for specific subgroups relative to that **subgroup's district score on the topic area**.

> **Example:** For schools that have female staff scores on CLC that are low compared with the districts' female staff scores on CLC, the district can help those schools improve CLC-related issues for their female staff.

● Provide targeted supports on topic areas to schools that have low subgroup scores relative to scores for that subgroup in other schools in the district.

> **Example:** For schools where Hispanic noninstructional staff report lower scores on CLC than in other schools in the district, the district can focus help in those schools to improve the CLC climate for Hispanic noninstructional staff.

d Interpreting Item-Level Data: Direction, Support, and Resources for Districts

Table A-2 provides direction, support, and resources for **districts** on interpreting item-level data overall and by respondent characteristics. It includes comparisons that can and cannot be made using item-level data (do's and don'ts), how these comparisons can help districts improve school climate, and cautions or limitations when interpreting the data.

> Note that parent/guardian item-level results can be interpreted in the same way as for other respondent groups (students, instructional staff, and noninstructional staff). That is, parent/guardian data can be examined using item-level percentage distributions and averages (means), across respondent groups, if items are worded the same, and they can be included in item-content groupings.

Table A-2. Do's and Don'ts for Item-Level Data

Do Examine:

1. **For items per individual respondent group (students, instructional staff, noninstructional staff, and parents/guardians)**

● How favorably a single item is perceived, on average (rated 1–4), across the district within respondent groups.

> **Example**: A district can see how favorably students across the district perceive the single item, "This school provides instructional materials (e.g., textbooks, handouts) that reflect my cultural background, ethnicity, and identity."

● How favorably an item is perceived, on average, in each school (rated 1–4), compared with the district and with other schools in the district, per respondent group.

> **Example:** A district can see how favorably, on average, students in a single school perceive the item, "This school provides instructional materials (e.g., textbooks, handouts) that reflect my cultural background, ethnicity, and identity," compared with other schools in the district or with the district average.

2. **For individual items that are worded *exactly the same* across respondent groups of interest (called "comparable items")**

● How a comparable item is perceived, on average, across the district, across respondent groups.

> **Example:** All four respondent groups (students, instructional staff, noninstructional staff, and parents) are asked, "At this school, how much of a problem is student drug use?" Districts can examine how much the perception of drug use might differ across these groups, on average, or by looking at the distribution across response options: not a problem, small problem, somewhat a problem, large problem.

● The extent of the difference in ratings of a comparable item between respondent groups in a specific school, compared with other schools in the district and with the district average (mean).

> **Example:** Do any of the district schools show a substantially greater difference between students and staff on the extent of student drug use at school, compared with the extent of the difference between students and staff in other district schools or compared with the district average (mean)?

3. **For items within a topic area grouped by content ("item-content group")**

It can sometimes be daunting to examine item-level data because of the sheer number of items and response options; therefore, it is useful to organize items by content area and examine these groupings. (Suggested item-content groupings can be found in each topic area's *Discussion Guide*.) In this way, specific areas within topics with particularly low favorability ratings may be more readily identified; districts and schools may want to target these content areas for immediate interventions.

> For example, items in the Bullying/Cyberbullying topic area can be grouped by content areas such as bullying is a problem; bullying related to race, ethnicity, and religion; cyberbullying; bullying related to sexuality; bullying related to physical and mental disability; bullying prevention: student reporting and stopping of bullying; and bullying prevention: bullying knowledge and formal prevention strategies.

Item groupings for each topic area are suggested in each topic area *Discussion Guide*.

● How favorably an item-content group is perceived overall across the district, compared with other item-content groups in the topic area.

> **Example:** A district can see how favorably all respondent groups in the district perceive the item-content group called "cyberbullying," compared with the other bullying item-content groups.

● How favorably an item-content group is perceived overall in each school, compared with the district and with other schools in the district.

> **Example:** A district can see how favorably all respondent groups in a specific school perceive the cultural diversity of materials compared with the other schools in the district and with the district average.

▼ **Cautions/Limitations:**

1. **For items per individual respondent group (students, instructional staff, noninstructional staff, and parents/guardians)**

● It is important to consider that items represent various aspects of the overall topic area that is being measured; each has been selected to reflect only one aspect of respondent perceptions.

Example: If examining Environment items suggests that bathroom cleanliness in this school is an area for improvement, this information might help contextualize perceptions of the environment in the school. However, although improving bathroom cleanliness might lead to more favorable responses about that particular problem, it may have less of an effect on the important underlying perception of the overall environment in the school.

● You can think of excessive focus on improving responses to a single item as a physician attempting to alleviate symptoms without curing the underlying disease or condition.

● Perhaps as important, when measuring any topic area, the indicators (items) should reflect a representative sample of behaviors and attitudes about the topic. If schools focus change solely on behaviors and attitudes as defined by specific items, they risk a form of unintentional "teaching to the test" in which schools are able to show growth with respect to specific items even though perceptions of the underlying topic area have not changed.

2. **For individual items that are worded <u>exactly the same</u> across respondent groups (called "comparable items")**

● In addition to the cautions detailed earlier, item-level data cannot be compared across respondent groups or subgroups of interest unless they are worded exactly the same.

3. **For items within a topic area grouped by content ("item-content group")**

● In addition to the cautions and limitation described for item-level results, data users should recognize that item groupings are intended solely as a way to organize a large amount of information and do not represent psychometrically validated scales.

Do Use These Results To:

For items per individual respondent group (students, instructional staff, noninstructional staff, and parents/guardians)

In general, if the rating on a specific item is less favorable than you think acceptable and is an important issue to your district and community, you may consider this a specific area to target interventions.

● Look at items that have low favorability ratings compared with other items in the same domain, within specific respondent groups to target interventions.

Example: If, on average, students rate the single item, "This school provides instructional materials (e.g., textbooks, handouts) that reflect my cultural background, ethnicity, and identity," lower than other items in the CLC topic area, a district can, for instance, support its teaching staff to start including more culturally diverse examples in class materials and lessons.

● Look at items that have low favorability ratings compared with other items across domains, within specific respondent groups, to target interventions.

Example: If examining Safety domain items suggests that there is a problem with students being teased or picked on about race or ethnicity, the school may want to examine other items from other domains to see if related issues are being reported there. For instance, examining Engagement domain items may show that students report few examples of different racial, ethnic, or cultural backgrounds in class lessons. Targeting a quick adjustment on the latter issue, such as instructing teachers to incorporate more culturally diverse examples in materials and lessons, may help to bring a sense of cultural sensitivity to the district.

1. **For individual items that are worded exactly the same across respondent groups (called "comparable items")**

   In general, if the rating on a comparable item is more unfavorable for a specific respondent group than you think acceptable, or if it is more unfavorable for a respondent group in your school than that group in the district, and this issue is important to your school and community, you may consider this area one to specifically target for interventions.

   ● Analyze percentage distributions on comparable items across respondent groups to investigate differences in perceptions of a potential problem or specific area.

   > **Example:** Look at the percentage distribution across response options for the student drug-use item (not a problem, small problem, somewhat a problem, large problem). If most students say that it is somewhat or a large problem, but instructional staff are saying it's not a problem or a small problem, you can infer that teachers are not seeing what the students are seeing in school. You can then begin targeting support for staff in regard to this issue.

   ● Compare averages (means) of comparable items across respondent groups to investigate differences in perceptions of a potential problem or specific area.

   > **Example:** Look at the reported average (mean) for the student drug use item for students and for instructional staff. You can think of the response options as ranging from favorable to unfavorable (1 = not a problem, 2 = small problem, 3 = somewhat a problem, 4 = large problem).
   >
   > Thus, the average (mean) that is reported for the drug-use comparable item will range from 1 (most favorable) to 4 (most unfavorable). If the average (mean) of student reports is 3.5 and the average (mean) for instructional staff is 2, you can see that students see this more unfavorably than do teachers. Again, you can use this information to target supports concerning this issue to appropriate stakeholders in the school, (in this case, teachers).

2. **For items within a topic area grouped by content ("item-content group")**

   Looking at groups of items that pertain to similar substantive content is another way of looking at item-level data. This approach may be especially helpful when digging deeper into the data to see which areas might be ripe for immediate intervention.

   For each school climate topic area, we have consulted with content experts to arrive at helpful ways of grouping items. When you click on links to the *Discussion Guides*, you will see these item-content groupings for specific topic areas.

   ● Using this approach, districts can see where they can target interventions and where they might be able to wait to implement longer term interventions.

   > **Example:** In digging deeper into the CLC Physical Environment topic area data, it may become evident that the cultural diversity reflected in learning materials is wanting for students and staff. A quick adjustment can immediately be implemented, such as instructing teachers to incorporate more culturally diverse examples in materials and lessons. If, conversely, a school's respondents report that the climate of respect and equality in the school is favorable, efforts to improve respect and equality may wait until more urgent CLC needs are taken care of.

Appendix B. Comparisons for Schools

Interpreting Scale Scores: Direction, Support, and Resources for EDSCLS Schools

Table B-1 provides direction, support, and resources to **schools** on interpreting scale scores overall and by student and staff characteristics. It includes comparisons that can and cannot be made using scale scores (do's and don'ts), how these comparisons can help schools improve school climate, and cautions or limitations when interpreting the data. Scale scores may be difficult to interpret by themselves since scores across scales may not mean the same things (e.g., a score of 300 on the physical safety scale may not be comparable to a score of 300 on the emotional safety scale) in terms of what excellent schools look like. For sites using the EDSCLS, benchmark data will be available in the platform in fall 2017 to facilitate the meaning of scale scores.

Table B-1. Scale Scores: Do's and Don'ts for Student and Staff Data

Do Compare:

- The school's scale scores for within-domain topic areas.

  Example: Within the Engagement domain, you can compare scores for its topic areas: Cultural and Linguistic Competence, Relationships, and Participation.

- The school's overall scale scores with the district overall scores for each domain and each topic area produced.

  Example: You can compare a school's score on Engagement or Safety or Environment or each of the 12 topic areas with the district's score on that same domain or topic area.

- The school's scale scores for each domain and topic area between subgroups of respondents.

  Example: You can compare a school's scores for female students versus male students on Physical Safety (a topic area) or on Safety (a domain).

- The school's scale scores for within-domain topic areas for a single subgroup of respondents.

  Example: You can compare school scores for Black or African-American students on Emotional Safety, Physical Safety, Bullying/Cyberbullying, and Substance Abuse within the Safety domain.

- The school's scale scores with the district scale scores for each domain or topic area by respondent characteristics.

  Example: You can compare your school's score for Black or African-American students on CLC with the district's score for Black or African-American students on CLC.

Do Not Compare:

Prior to the availability of benchmark data, scale scores may be difficult to interpret by themselves because scores across scales may not mean the same things (e.g., a score of 300 on the physical safety scale may not be comparable to a score of 300 on the emotional safety scale in terms of what excellent schools look like).

● Scale scores for domains/topic areas across respondent groups.

> **Example:** Scores on CLC for students should not be compared with CLC scores for staff.

● Scale scores for subgroups on topic areas across domains.

> **Example:** Scores for students on the Physical Safety topic area (within the Safety domain) should not be compared with the student score for the Relationships topic area (within the Engagement domain).

● Scale scores by respondent characteristics for domains/topic areas for subgroups across respondents.

> **Example:** Scores on CLC for White students should not be compared with CLC scores for White staff.

● Scale scores by respondent characteristics for topic areas across domains.

> **Example:** Scores for Hispanic noninstructional staff on the Physical Safety topic area (within the Safety domain) should not be compared with Hispanic noninstructional staff's score for the Relationships topic area (within the Engagement domain).

Do Use These Comparisons To:

● Provide universal/schoolwide support to students and staff for within-domain topic areas that have relatively low scores in the school.

> **Example:** If a school's student score on CLC is low compared with the student scores on other Engagement topics (Relationships and Participation), you can focus supports on improving conditions in the school pertaining to student CLC.

● Provide universal/schoolwide support to students and staff for within-domain topic areas that have relatively low scores compared with the district scores.

> **Example:** If the school's student score on CLC is low compared with the other Engagement topics (Relationships and Participation), the school can focus supports on improving conditions in the school pertaining to student CLC.

● Provide targeted support to subgroups of students and staff for within-domain topic areas that have relatively low scores in the school.

> **Example:** If a school's score for Asian students on CLC is low compared with Asian students' scores on other Engagement topics (Relationships and Participation), the school can target supports on improving CLC for Asian students in the school.

● Provide targeted support to subgroups of students and staff for within-domain topic areas that have relatively low scale scores compared with the district scale scores.

> **Example:** If a school's score on CLC is lower for Black or African-American instructional staff than the district's score on CLC for Black or African-American instructional staff, the school can target efforts on improving issues related to CLC for those staff.

● Provide targeted support to subgroups of students and staff for topic areas that have relatively low scores for one subgroup compared with another.

> **Example:** If a school's score for female students in Emotional Safety is lower than the school's score for male students in Emotional Safety, the school can target supports on improving Emotional Safety for girls in the school.

Interpreting Item-Level Data: Direction, Support, and Resources for Schools

Table B-2 provides direction, support, and resources for **schools** on interpreting item-level data overall and by respondent characteristics. It includes comparisons that can and cannot be made (do's and don'ts), how these comparisons can help schools improve school climate, and cautions or limitations when interpreting the data.

> Note that parent/guardian item-level results can be interpreted in the same way as for other respondent groups (students, instructional staff, and noninstructional staff). That is, parent/guardian data can be examined using item-level percentage distributions and averages (means), across respondent groups if items are worded the same, and can be included in item-content groupings.

Table B-2. Do's and Don'ts for Item-Level Data

Do Examine:

1. **For items per individual respondent group (students, instructional staff, noninstructional staff, and parents/guardians)**

● How favorably an item is perceived, on average, in the school (rated 1–4), per respondent group.

> **Example:** A school can see how favorably, on average, students perceive the single item, "This school provides instructional materials (e.g., textbooks, handouts) that reflect my cultural background, ethnicity, and identity." Because this item is positively valenced, higher average (mean) values represent more favorable perceptions.

● How favorably an item is perceived, on average, in the school (rated 1–4), compared with the district, per respondent group.

> **Example:** A school can see how favorably, on average, students perceive the item, "This school provides instructional materials (e.g., textbooks, handouts) that reflect my cultural background, ethnicity, and identity" compared with the district average. Because this item is positively valenced, higher average (mean) values represent more favorable perceptions.

2. **For individual items that are worded exactly the same across respondent groups (called "comparable items")**

● How a comparable item is perceived across respondent groups.

> **Example:** A school can examine how much the perception differs across students, instructional staff, noninstructional staff, and parents on the comparable item, "At this school, how much of a problem is student drug use?"

● The extent of the difference in responses to a comparable item between respondent groups in the school, compared with the extent of the difference in the district.

> **Example:** Do any of the district schools show a substantially greater difference between students and staff on the extent of student drug use at school, compared with the extent of the difference between students and staff in other district schools, or compared with the district average (mean)?

3. **For items within a topic area grouped by content ("item-content group")**

It can sometimes be daunting to examine item-level data because of the sheer number of items and response options; therefore, it is useful to organize items by content area and examine these groupings. (Suggested item-content groupings can be found in each topic area's _Discussion Guide_.) In this way, specific areas within topic areas with particularly low favorability ratings may be more readily identified; schools may want to target these content areas for immediate interventions.

> For example, items in the Bullying/Cyberbullying topic area can be grouped by content areas such as bullying is a problem; bullying related to race, ethnicity, and religion; cyberbullying; bullying related to sexuality; bullying related to physical and mental disability; bullying prevention: student reporting and stopping of bullying; and bullying prevention: bullying knowledge and formal prevention strategies.

Item groupings for each topic area are suggested in each topic area's _Discussion Guide_.

● How favorably an item-content group is perceived by all respondent groups in the school, compared with other item-content groups in the same topic area.

> **Example:** A school can see how favorably all respondent groups in the school perceive the item grouping "cyberbullying," compared with the other bullying item-content groups.

● How favorably an item-content group is perceived by all respondent groups in the school, compared with the district.

> **Example:** A school can see if its respondent groups, as a whole, perceive the cultural diversity of materials more or less favorably than the district as a whole.

Cautions/Limitation

1. **For items per individual respondent group (students, instructional staff, noninstructional staff, and parents/guardians)**

● It is important to consider that items represent various aspects of the overall topic area that is being measured; each has been selected to reflect only one aspect of respondent perceptions.

> **Example:** If examining Environment items suggests that bathroom cleanliness in this school is an area for improvement, this information might help contextualize perceptions of the environment in the school. However, although improving bathroom cleanliness might lead to more favorable responses about that particular problem, it may have less of an effect on the important underlying perception of the overall environment in the school.

● One can think of excessive focus on improving responses to a single item as a physician attempting to alleviate symptoms without curing the underlying disease or condition.

● Perhaps as important, when measuring any topic area, the indicators (items) should reflect a representative sample of behaviors and attitudes about the topic. If schools focus change solely on behaviors and attitudes as defined by specific items, they risk a form of unintentional "teaching to the test" in which schools are able to show growth with respect to specific items even though perceptions of the underlying topic area have not changed.

2. **For individual items that are worded <u>exactly the same</u> across respondent groups (called "comparable items")**

● In addition to the cautions detailed earlier, item-level data cannot be compared across respondent groups or subgroups of interest unless they are worded exactly the same.

3. **For items within a topic area grouped by content ("item-content group")**

● In addition to the cautions and limitation described for item-level results, data users should recognize that item groupings are intended solely as a way to organize a large amount of information and do not represent psychometrically validated scales.

Do Use These Results To:

1. **For items per individual respondent group (students, instructional staff, noninstructional staff, and parents/guardians)**

In general, if the rating on a specific item is less favorable than you think acceptable and is an important issue to your district and community, you may consider this a specific area to target interventions.

● Look at items that have low favorability ratings compared to other items in the same domain, within specific respondent groups to target interventions.

> **Example:** If, on average, students rate the single item, "This school provides instructional materials (e.g., textbooks, handouts) that reflect my cultural background, ethnicity, and identity," lower than other items in the CLC topic area, a school can, for instance, support its teaching staff to start including more culturally diverse examples in class materials and lessons.

● Look at items that have low favorability ratings compared with other items across domains, within specific respondent groups, to target interventions.

> **Example:** If examining Safety domain items suggests that there is a problem with students being teased or picked on about race or ethnicity, the school may want to examine other items from other domains to see if related issues are being reported there. For instance, examining Engagement domain items may show that students report few examples of different racial, ethnic, or cultural backgrounds in instructional materials. Targeting a quick adjustment on the latter issue, such as instructing teachers to incorporate more culturally diverse examples in materials and lessons, may help to bring a sense of cultural sensitivity to the school.

2. **For individual items that are worded exactly the same across respondent groups (called "comparable items")**

In general, if the rating on a comparable item is more unfavorable for a specific respondent group than you think acceptable, or if it is more unfavorable for a respondent group in your school than that group in the district, and this issue is important to your school and community, you may consider this a specific area to target interventions.

● Analyze percentage distributions on comparable items across respondent groups to investigate differences in perceptions of a potential problem or specific area.

> **Example:** Look at the percentage distribution across response options for the student drug-use item (not a problem, small problem, somewhat a problem, large problem). If most students say that it is somewhat or a large problem, but instructional staff are saying it's not a problem or a small problem, you can infer that teachers are not seeing what the students are seeing in school. You can then begin targeting support for staff in regard to this issue.

● Compare averages (means) of comparable items across respondent groups to investigate differences in perceptions of a potential problem or specific area.

> **Example:** Look at the reported average (mean) for the student drug use item for students and for instructional staff. You can think of the response options as ranging from favorable to unfavorable (1 = not a problem, 2 = small problem, 3 = somewhat a problem, 4 = large problem).
>
> Thus, the average (mean) that is reported for the drug-use comparable item will range from 1 (most favorable) to 4 (most unfavorable). If the average (mean) of student reports is 3.5 and the average (mean) for instructional staff is 2, you can see that students see this more unfavorably than do teachers. Again, you can use this information to target supports concerning this issue to appropriate stakeholders in the school (in this case, teachers).

3. **For items within a topic area grouped by content ("item-content group")**

Looking at groups of items that pertain to similar substantive content is another way of looking at item-level data. This approach may be especially helpful when digging deeper into the data to see which areas might be ripe for interventions.

For each school climate topic area, we have consulted with content experts to arrive at helpful ways of grouping items. When you click on links to the *Discussion Guides*, you will see these item-content groupings for specific topic areas.

● Using this approach, schools can see where they can target interventions and where they might be able to wait to implement longer term interventions.

> **Example:** In digging deeper into the CLC topic area data, it may become evident that the cultural diversity reflected in learning materials is wanting for students and staff. A quick adjustment can immediately be implemented, such as instructing teachers to incorporate more culturally diverse examples in materials and lessons. If, conversely, a school's respondents report that the climate of respect and equality in the school is favorable, efforts to improve respect and equality may wait until more urgent CLC needs are taken care of.

Appendix C. Average (Mean) Topic Area Values

This appendix is intended for those who are interested in using average (mean) topic area values for interpreting their survey results. This includes (1) EDSCLS pilot sites that received average (mean) topic area values in their District and School Report PDFs in spring 2015; (2) any EDSCLS users who administer surveys after spring 2015 and wish to calculate average (mean) topic area values; and (3) users of other surveys who employ items with a 1–4 response option scale.

Average (Mean) Topic Area Values

The scale scores provided for EDSCLS sites prior to fall 2017 will not have benchmark data on which to ground the meaning of their results. Therefore, we provide information here about average (mean) topic a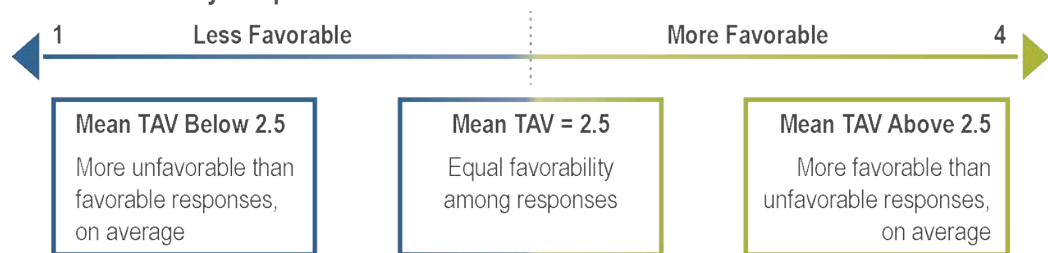rea values that can be used as points of comparison. These values were included in the District and School Report PDFs that were provided to pilot study sites. After the benchmark data are released in fall 2017, EDSCLS users can use benchmark data to facilitate the meaning of their scale scores.

Overall Average (Mean) Topic Area Values

Average (mean) topic area values can be used to gauge the degree to which respondents agree with the items that make up a content group or topic area (i.e., how favorably they perceive the topic area).

Because all of the response options in the average (mean) topic area values are on a scale of 1–4 (where 1 is less favorable and 4 is more favorable), you can think of the favorability of the item compared with standard benchmarks as shown in Figure 8.

Figure 8. Item Favorability Interpretation

1	Less Favorable	More Favorable	4

Mean TAV Below 2.5	Mean TAV = 2.5	Mean TAV Above 2.5
More unfavorable than favorable responses, on average	Equal favorability among responses	More favorable than unfavorable responses, on average

Note: Topic area value is abbreviated TAV.

As illustrated in Figure 8, an average (mean) topic area value of less than the benchmark of 2.5 would indicate that there are more unfavorable than favorable responses, on average, to the items comprising the topic area, suggesting that this topic area is perceived relatively unfavorably. Conversely, an average (mean) topic area value of more than 2.5 indicates more favorable than unfavorable responses, on average, suggesting that the topic area is perceived favorably, while an

average (mean) topic area value of 2.5 indicates equal favorability among responses, suggesting that perception of the topic area is "neutral" (neither favorable nor unfavorable).

Average (Mean) Topic Area Values Broken Out by Respondent Characteristics

As with scale scores, average (mean) topic area values can be broken out by respondent subgroups to provide a deeper understanding of the data. For example, districts and schools can see how the favorability of a topic area differs across subgroups of students and staff.

For sites using the EDSCLS, average (mean) topic area values can be examined by the following respondent characteristics:[10]

- Student average (mean) topic area values by:
  - Gender,
  - Race/ethnicity, and
  - Grade.
- Staff (instructional and noninstructional) average (mean) topic area values by:
  - Gender and
  - Race/ethnicity.

For EDSCLS pilot sites, for each respondent group except parents, overall average (mean) topic area values have been provided for you in a District or School Report PDF. They can be viewed in a chart like the example shown in Figure 9.

Figure 9. Example of Student Average (Mean) Discipline Value Chart, Overall and by Gender, Grade, and Race/Ethnicity
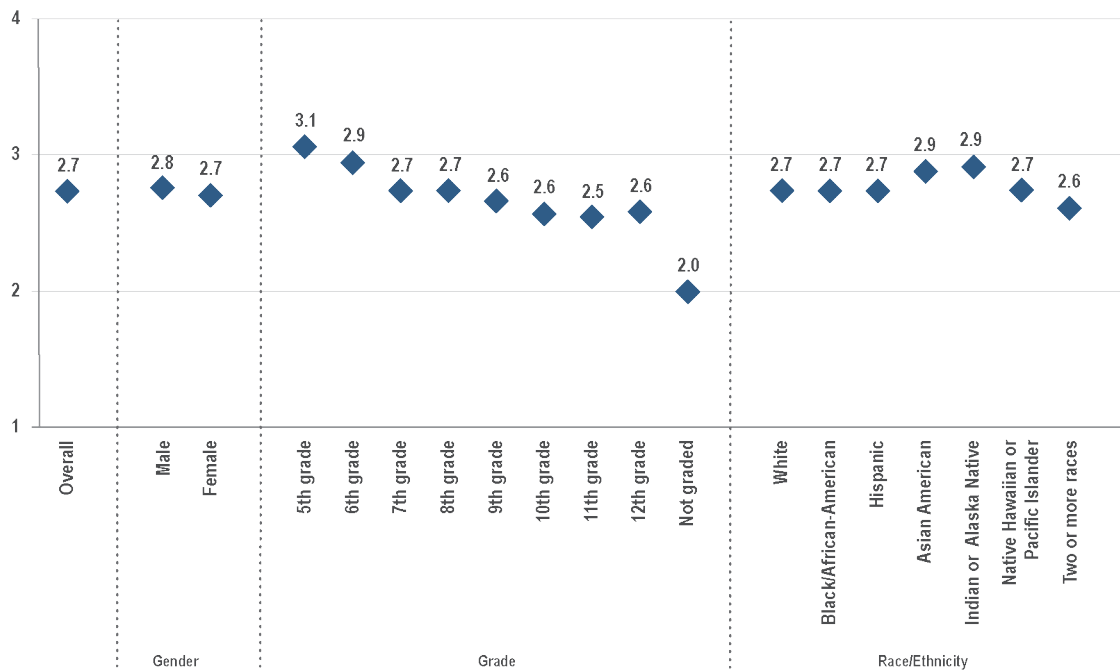
10 Note that the EDSCLS platform does not produce crossed demographics (e.g., Asian female students).

In this example, the overall school average (mean) for the Discipline topic area is 2.7 (see first column). The chart also shows, for example, that the average (mean) for the Discipline topic area for student grade level ranges from a low of 2.00 for ungraded students to a high of 3.1 for fifth graders. The average (mean) for the Discipline topic area for race/ethnicity ranges from 2.6 for students of two or more races to 2.9 for Indian or Alaskan Native students.

Interpreting Average (Mean) Topic Area Values

Tables C-1 (for districts) and C-2 (for schools) provide suggestions for EDSCLS and other survey users for interpreting average (mean) topic area values overall and by respondent characteristics. They also include comparisons that can and cannot be made using average (mean) topic area values (do's and don'ts) and how these comparisons can help you improve school climate.

If you are a **district**, click Appendix C, Table C-1, to go to interpretation of average (mean) topic area values overall and by respondent characteristics.

If you are a **school**, click Appendix C, Table C-2, to go to interpretation of average (mean) topic area values overall and by respondent characteristics.

Instructions for Calculating Average (Mean) Topic Area Values or Item-Content Groups (Mean of Means)[11]

Averages (means) of groups of items, whether they are topic area values, item-content groups, or any other item groupings, can be calculated as the mean of the means of the associated items. For example, if a topic area or item-content group has four associated items, the mean would be calculated as follows:

Example Items	Item 1	Item 2	Item 3	Item 4
Item Mean	2.5	2.5	2.4	2.6

In this example, the averages (means) for items 1–4 are 2.5, 2.5, 2.4, and 2.6. The average (mean) of these averages (means) is calculated as the sum of the item average (mean) divided by the number of items:

Average (mean) topic area value = (2.5+2.5+2.4+2.6)/4 = 2.5

Here, the average (mean) of the items 1–4 is 2.5. Before calculating an average (mean) value for a group of items, be sure to reverse code any negatively valenced items in the group, as described in the section Reverse Coding Negatively Valenced Items.

If you are an EDSCLS pilot site, the reverse coding has already been done to calculate the average (mean) topic area values in the District or School Report PDF produced for you, as well as in the scale scores produced by the Web-based platform. However, you will have to reverse code negative items yourself if you want to calculate the average (mean) of item-content groups (see Table 1 of the *Discussion Guides* for your specific topic of interest) or any other groups of items.

---

[11] For more information on calculating mean (average) topic area values with EDSCLS data, please contact the EDSCLS Help Desk at edscls@air.org.

d Interpreting Average (Mean) Topic Area Values: Direction, Support, and Resources for Districts

Table C-1 provides direction, support, and resources to **districts** on what comparisons can be made with average (mean) topic area values overall and by respondent characteristics, how these comparisons can help schools and districts improve school climate, and cautions and limitations when interpreting these data.

Table C-1. Do's and Don'ts for Average (Mean) Topic Area Values

▲ Do Compare (in addition to the scale score comparisons detailed in Table A-1):

- The favorability of a topic area (rated as 1–4) within one domain versus a topic area in another domain, as perceived by a single respondent group.

  > **Example:** You can compare your school's average (mean) topic area values on the Relationship topic area (within the Engagement domain) with the Mental Health topic area (within the Environment domain) to determine how favorably respondents perceive each of these topic areas.

- The favorability of a topic area (rated as 1–4) within a single domain as perceived by a respondent group versus the favorability perceived by another respondent group.

  > **Example:** You can compare your school's average (mean) topic area values on Emotional Safety and Physical Safety (both within the Safety domain) across students and staff.

- The favorability of a topic area (rated as 1–4) within one domain versus a topic area in another domain, as perceived by subgroups of a single respondent group.

  > **Example:** You can compare your school's average (mean) topic area value on the Relationship topic area (within the Engagement domain) with the Mental Health topic area (within the Environment domain) to determine how favorably female students versus male students perceive each area.

- The favorability of a topic area (rated as 1–4) within a single domain as perceived by a specific subgroup of respondents versus the favorability perceived by that same subgroup of a different respondent group.

  > **Example:** You can compare your school's average (mean) topic area value on the Emotional Safety versus Physical Safety topic areas (both within the Safety domain) to determine how favorably female students versus female instructional staff perceive each area.

▼ Do Not Compare:

- These average (mean) scale scores will provide information about how positively or negatively respondents feel about the topic area or domain, but even though topic areas have similar averages (means), they still may not be "equal" (i.e., respondents may feel more or less favorably about various topics to start with or have more or less difficulty answering questions about them). Readers should therefore use caution when interpreting comparisons across respondent groups.

- As is the case for scale scores, when data are disaggregated by subgroup, they will indicate how positively or negatively subgroups feel about the topic area or domain, but even though topic areas

have similar averages (means), they still may not be "equal" (i.e., the importance of a topic or how intense the desire to achieve excellence in that topic area may vary across respondent subgroups). Readers should therefore use caution when interpreting comparisons across respondent subgroups.

Do Not These Comparisons To:

● Focus immediate universal/schoolwide support pertaining to those topic areas perceived as the most unfavorable across all domains, for a specific respondent group.

> **Example:** You can compare your school's average (mean) topic area values on each topic area in the survey and prioritize supports based on the topic area(s) with the lowest average (mean) student values.

● Focus immediate targeted support to a specific respondent group that perceives a topic area more unfavorably than do other respondents.

> **Example:** If instructional staff report highly favorable perceptions of CLC in the school building but students report less favorable perceptions, this difference could indicate that staff's intended attention to cultural sensitivity and inclusion is not being executed in a way that speaks to students, and improvement could be focused on staff execution.

● Focus immediate universal/schoolwide support pertaining to topic areas perceived as the most unfavorable across all domains, for a specific subgroup of respondents.

● Focus immediate targeted support to a specific subgroup of respondents that perceives a topic area more unfavorably than does the same subgroup of other respondents.

Interpreting Average (Mean) Topic Area Values: Direction, Support, and Resources for Schools

Table C-2 provides direction, support, and resources for **schools** on what comparisons can be made with average (mean) topic area values overall and by respondent characteristics, how these comparisons can help schools to improve school climate, and cautions and limitations when interpreting these data.

Table C-2. Do's and Don'ts for Average (Mean) Topic Area Values

Do Compare (in addition to the scale score comparisons detailed in Table A-1):

● The favorability of a topic area (rated as 1–4) within one domain versus a topic area in another domain, as perceived by a single respondent group.

> **Example:** You can compare your school's average (mean) topic area values on the Relationship topic area (within the Engagement domain) with the Mental Health topic area (within the Environment domain) to determine how favorably respondents perceive each of these topic areas.

● The favorability of a topic area (rated as 1–4) within a single domain as perceived by a respondent group versus the favorability perceived by another respondent group.

> **Example:** You can compare your school's average (mean) topic area values on Emotional Safety and Physical Safety (both within the Safety domain) across students and staff.

● The favorability of a topic area (rated as 1–4) within one domain versus a topic area in another domain, as perceived by subgroups of a single respondent group.

> **Example:** You can compare your school's average (mean) topic area value on the Relationship topic area (within the Engagement domain) with the Mental Health topic area (within the Environment domain) to determine how favorably female students versus male students perceive each area.

● The favorability of a topic area (rated as 1–4) within a single domain as perceived by a specific subgroup of respondents versus the favorability perceived by that same subgroup of a different respondent group.

> **Example:** You can compare your school's average (mean) topic area value on the Emotional Safety versus Physical Safety topic areas (both within the Safety domain) to determine how favorably female students versus female instructional staff perceive each area.

▼ Do Not Compare:

● These average (mean) scale scores will provide information about how positively or negatively respondents feel about the topic area or domain, but even though topic areas have similar averages (means), they still may not be "equal" (i.e., respondents may feel more or less favorably about various topics to start with or have more or less difficulty answering questions about them). Readers should therefore use caution when interpreting comparisons across respondent groups.

● As is the case for scale scores, when data are disaggregated by subgroup, they will indicate how positively or negatively subgroups feel about the topic area or domain, but even though topic areas have similar averages (means), they still may not be "equal" (i.e., the importance of a topic or how intense the desire to achieve excellence in that topic area may vary across respondent subgroups). Readers should therefore use caution when interpreting comparisons across respondent subgroups.

▲ Do Use These Comparisons To:

● Focus immediate universal/schoolwide support pertaining to those topic areas perceived as the most unfavorable across all domains, for a specific respondent group.

> **Example:** You can compare your school's average (mean) topic area values on each topic area in the survey and prioritize supports based on the topic area(s) with the lowest average (mean) student values.

● Focus immediate targeted support to a specific respondent group that perceives a topic area more unfavorably than do other respondents.

> **Example:** If instructional staff report highly favorable perceptions of CLC in the school building but students report less favorable perceptions, this difference could indicate that staff's intended attention to cultural sensitivity and inclusion is not being executed in a way that speaks to students, and improvement could be focused on staff execution.

● Focus immediate universal/schoolwide support pertaining to topic areas perceived as the most unfavorable across all domains, for a specific subgroup of respondents.

● Focus immediate targeted support to a specific subgroup of respondents that perceives a topic area more unfavorably than does the same subgroup of other respondents.

Appendix D. Glossary of Terms

Domain: Refers to a major overarching school climate scale, constructed psychometrically, often containing subdomains. The ED School Climate Surveys model of school climate contains three domains—Engagement, Safety, and Environment.

Mean: Refers to the *average* value of a group of scores. Mean scores are calculated by adding all of the scores and dividing by the number of scores. For example, the mean of three scores (2, 3, and 4) would be calculated as: *Mean: (2 + 3 + 4) / 3 = 3*.

Mean topic area value: Refers to the *average* value of a group of items that comprise a topic area. Mean topic area values are presented on a scale of 1–4 to reflect the response options in the EDSCLS (see Appendix C for more information).

Scale score: Refers to a numeric measurement that represents a survey respondent's answers to a group of related survey questions. Psychometric analysis is often used to produce scale scores. One common type of scale score is called a "theta" value, which ranges from −3 to +3 but is often transformed. For example, scale scores for the EDSCLS are derived from theta values but are transformed to a scale of 100–500.

Subdomain: Refers to a concept of school climate that is part of a larger domain. For example, in the EDSCLS model of school climate, there are three subdomains within the Engagement domain. These subdomains are Cultural and Linguistic Competence, Relationships, and Participation. Subdomains also are referred to *topics* or *topic areas*.

Topic area: Refers to a concept of school climate that is part of a larger domain. For example, in the EDSCLS model of school climate, there are three topics or topic areas within the Engagement domain. These topics are Cultural and Linguistic Competence, Relationships, and Participation. Topics also are referred to as *subdomains*.

Valenced: Refers to the intrinsic direction of a term or phrase. For example, "I feel safe" is considered positively valenced because agreeing with the statement indicates good or positive perceptions. In contrast, "I don't feel safe" is considered a negatively valenced phrase because agreeing with the statement indicates bad or negative perceptions.