Appendix A: Full Evaluation Tool

Questions to ask	Does Not Meet Criteria	Partially Meets Criteria	Meets Criteria
Is the assessment rooted in equity?			
1. Do the developers center equity in the purpose and intended use of the assessment?	Developers do not consider the equity implications of the items or the assessment. • Developers may reference equity but do not center it in the assessment questions or other materials, or the assessment questions and materials do not match the developer's stated purpose. • Developers do not recognize systemic inequities or discuss how their assessment can be used to reduce inequities. • Developers do not clearly identify the purpose of the assessment with an asset-based lens recognizing the assets of historically marginalized students. **Assessments not meeting criteria:** • DESSA	Developers consider the equity implications but do not center an equity lens in the creation of the assessment. • Developers may reference equity but do not clearly center it in the assessment questions or other materials. • Developers may discuss how assessment results can be used to address disparities, but do not clearly articulate an anti-racist lens for the assessment overall or do not explicitly recognize systemic inequities facing marginalized students. **Assessments partially meeting criteria:** • 5Essentials • Inventory of School Climate • NJ School Climate Improvement Survey • Panorama SEL Survey • Elevate • USED School Climate Survey	The developers center an equity lens (e.g. anti-racist or liberatory) in the purpose and intended use of the assessment. • Developers explicitly recognize systemic inequities facing historically marginalized student groups and include how the assessment can be used to address them. • Developers identify the purpose of the assessment with an asset-based lens, such as identifying the assets of historically marginalized students rather than seeking out deficits. • Developers provide clear examples of how responses to assessment results, such as policy and practice changes, can address disparities. **Assessments meeting criteria:** • Catalyze • Cultivate • School Climate for Diversity • SAESECI
2. Do all or most items in the assessment use a strength-based lens?	Most items have a deficit lens. • Items have a deficit lens and are framed negatively, such as "Does the student have difficulty with relationships?" or "Student possession of weapons occurs often at this school." **Assessments not meeting criteria:** • n/a	Some items use a strength-based lens, but it's not consistent throughout the assessment. • Items mostly center on what students are doing well or on students' experiences, but there are also some deficit-focused or negatively framed items. **Assessments partially meeting criteria:** • Inventory of School Climate • NJ School Climate Improvement Survey • Panorama SEL Survey • School Climate for Diversity • USED School Climate Survey	All or almost all items use a strength-based lens. • Items center on what students are doing well even in the face of struggles, e.g., "How often does the student reach out to help other students?" vs "Does the child have difficulty with relationships?" • Some assessments may have mostly items focusing on student experiences rather than their own competencies and actions. For this type of item, look for framing that focuses on the student's experience rather than negative behaviors of other students, e.g. "I feel stereotyped for my race at school" rather than "Other students have a lot of stereotypes about my race." **Assessments meeting criteria:** • 5Essentials • Catalyze • Cultivate • DESSA • Elevate • SAESECI

Questions to ask *Continued*	Does Not Meet Criteria *Continued*	Partially Meets Criteria *Continued*	Meets Criteria *Continued*
3. Does the assessment include an array of items to capture the full picture of students' experiences?	The assessment is narrowly focused on students' social and emotional skills without attention to experiences that impact students' social, emotional, and academic development. **Assessments not meeting criteria:** • DESSA	The assessment includes some items about students' experiences and how they impact their social, emotional, and academic development. • The assessment may be missing some categories of students' experiences, such as focusing only on academic experiences and not including any on belonging or cultural identity. **Assessments partially meeting criteria:** • 5Essentials • Catalyze	The assessment is designed to identify how students' experiences have impacted their social, emotional, and academic development. • Assessment items focus on students' experiences and not primarily on their competencies and skills. • The assessment includes items about students' feelings of safety, acceptance, and belonging at school. • The assessment includes items about students' experiences related to their race/ethnicity or cultural identity, including whether they feel reflected in academic curriculum. **Assessments meeting criteria:** • Cultivate • Inventory of School Climate • NJ School Climate Improvement Survey • Panorama SEL Survey • Elevate • School Climate for Diversity • SAESECI • EDSCLS

Questions to ask *Continued*	Does Not Meet Criteria *Continued*	Partially Meets Criteria *Continued*	Meets Criteria *Continued*
Is the use case of the assessment clear?			
4. Are developers clear about the context in which the assessment should be used?	Developers do not clearly identify how the assessment should be used, or the stated purpose does not match the assessment and materials. **Assessments not meeting criteria:** • DESSA	Developers vaguely discuss the purpose or context in which the assessment should be used. • The developer may give a clear purpose for the assessment, or clear suggestions about contexts in which it should be used, but not both. **Assessments partially meeting criteria:** • Inventory of School Climate • School Climate for Diversity	Developers clearly identified how the assessment should be used, including by whom and for what purpose, and that purpose is clearly reflected in the assessment and materials. • The developer clearly identifies who should use an assessment, such as district leaders or classroom teachers. • The developer gives a specific purpose for the assessment and provides details on how it can be used for this purpose. • The assessment itself and associated materials clearly align with the purpose identified by the developer. • The developer offers clear information about what level an assessment should be used at and what level of data will be provided, such as whether data will be for a whole school only or whether it can be disaggregated by classroom or demographic subgroup. **Assessments meeting criteria:** • 5Essentials • Catalyze • Cultivate • NJ SCI Survey • Panorama SEL Survey • Elevate • SAESECI • EDSCLS
5. Does the developer offer a set of assessments to capture multiple perspectives?	The developer only offers one assessment with one perspective. **Assessments not meeting criteria:** • Catalyze • Cultivate • Inventory of School Climate • Elevate • Self-Assessing Social and Emotional Instruction and Competencies • School Climate for Diversity	The developer offers multiple assessments with varied perspectives, but they're not designed to work together to provide a cohesive picture, or developer offers only two perspectives. **Assessments partially meeting criteria:** • DESSA • Panorama SEL Survey	Developer offers multiple assessments to capture multiple perspectives, including students, families, and educators, and these assessments are designed to work together to provide a cohesive picture. • The developer provides resources or tools to help interpret the data from multiple perspectives together. **Assessments meeting criteria:** • 5Essentials • NJ SCI Survey • EDSCLS

Questions to ask *Continued*	Does Not Meet Criteria *Continued*	Partially Meets Criteria *Continued*	Meets Criteria *Continued*
Is the assessment valid and reliable?			
6. Were assessment items probed/validated via interviews with diverse groups of the intended respondent type?	Items were not validated via interviews or focus groups, or the developer does not report conducting interviews or focus groups. **Assessments not meeting criteria:** • 5Essentials • Inventory of School Climate • NJ School Climate Improvement Survey • Self-Assessing Social and Emotional Instruction and Competencies	Items were validated via interviews or focus groups, but the interviewees were not racially/ethnically diverse, or demographics of interviewees were not reported. • The developer may also report conducting interviews with diverse interviewees, but only for some sections or items from the assessment. **Assessments partially meeting criteria:** • Catalyze • Elevate • Panorama SEL Survey • USED School Climate Survey	Items were validated via interviews or focus groups with a racially/ethnically diverse group. • The developer gives detailed information about the development process of the assessment, including interviews. • The developer conducts cognitive interviews with diverse participants and provides demographic information. **Assessments meeting criteria:** • Cultivate • DESSA • School Climate for Diversity
7. Are the items written with language appropriate for the intended audience?	The developer does not provide information on how the language was developed to be appropriate for the target audience. • The items may seem to be of an appropriate reading level, but the developer does not provide any information about how age-appropriateness was determined. **Assessments not meeting criteria:** • Self-Assessing Social and Emotional Instruction and Competencies	The developer makes some consideration about the age-appropriateness of items, but does not provide extensive information about how it was determined. • Developers may discuss how language was adjusted after pilot testing. • Developers may have different versions of the assessment for different age groups, but do not provide information about how the language was adjusted to be . **Assessments partially meeting criteria:** • 5Essentials • Catalyze • DESSA • NJ School Climate Improvement Survey • USED School Climate Survey	Items are written at an appropriate reading level and developers considered the age-appropriateness of the content of the items. • The developer provides information on how they determined age appropriateness, such as consulting research and experts or testing with students of different ages. • The developer provides the specific reading level of the survey items. **Assessments meeting criteria:** • Inventory of School Climate • Cultivate • Elevate • Panorama • School Climate for Diversity

| Questions to ask
Continued | Does Not Meet Criteria
Continued | Partially Meets Criteria
Continued | Meets Criteria
Continued |
|---|---|---|---|
| **8.** Was the assessment piloted with diverse participants? | The assessment was not piloted with diverse participants, or there is no information about the demographics of the participants.

• The developer may still be in the process of piloting the assessment.

Assessments not meeting criteria:

• Catalyze
• Self-Assessing Social and Emotional Instruction and Competencies | The assessment was piloted with some diversity in the sample of students.

• The developer reports demographics such as race or income level, but does not report demographics such as English Language Learner status, gender, or LGBTQ+ status.

Assessments partially meeting criteria:

• 5Essentials
• DESSA
• Inventory of School Climate
• School Climate for Diversity
• USED School Climate Survey | The assessment was piloted with racially and ethnically diverse participants. If the assessment included reports from students, developers made sure to include students from low-income backgrounds, LGBTQ+ students, and English learners. If the assessment was developed for educators, it was piloted with diverse educators who also work with a diverse group of students.

• The developer conducted pilot testing with diverse participants (including racially /ethnically and economically diverse students, as well as other student groups such as students who identify as LGBTQ+, English learners, and students with disabilities) and reports data on the diversity of pilot participants.
• If the assessment was developed for educators, it was piloted with diverse educators who also work with a diverse group of students.

Assessments meeting criteria:

• Cultivate
• NJ School Climate Improvement Survey
• Panorama SEL Survey
• Elevate |
| **9.** Have latent variables been psychometrically validated with a statistical measure showing a good fit? | Latent variables were not psychometrically validated, or the developer does not report any information about validation of latent variables.

• The developer may still be in the process of piloting and validating the assessment.

Assessments not meeting criteria:

• Catalyze
• DESSA
• Elevate
• Self-Assessing Social and Emotional Instruction and Competencies | Latent variables were validated but did not have a good fit (I.e., RMSEA is not less than .05/.08).

• The developer may provide information about how the assessment was validated but does not provide specific measures that meet a high standard of fit.

Assessments partially meeting criteria:

• Cultivate | Latent variables were psychometrically validated with a good fit (e.g. RMSEA less than 0.5/0.8 or mean square residual close to 1.0).

• The developer provides extensive information about how the assessment was validated.
• The developer provides a measure of validation for latent variables (i.e., a set of questions aimed at measuring the same construct), such as an RMSEA or mean square residual, that meets a high standard of fit.

Assessments meeting criteria:

• 5Essentials
• Inventory of School Climate
• NJ School Climate Improvement Survey
• Panorama SEL Survey
• School Climate for Diversity
• USED School Climate Survey |

Questions to ask *Continued*	Does Not Meet Criteria *Continued*	Partially Meets Criteria *Continued*	Meets Criteria *Continued*
10. Is the assessment reliable across diverse groups of students?	The developer provides analysis showing that assessment items are not reliable across diverse groups of students, or the developer does not provide information on how items are valid and reliable across different groups. • The developer may still be in the process of piloting the assessment. **Assessments not meeting criteria:** • 5Essentials • Catalyze • DESSA • NJ School Climate Improvement Survey • Panorama SEL Survey • Elevate • Self-Assessing Social and Emotional Instruction and Competencies • School Climate for Diversity	The developer provides analysis showing that some assessment items function reliably across subgroups, while others do not. • The developer may provide a measure such as an RMSEA or Rasch analysis for each student subgroup, or measures such as Cronbach's alpha or differential item functioning, but the measures show that some items do not function similarly across groups. **Assessments partially meeting criteria:** • n/a	The developer provides analysis showing that the assessment items function similarly for subgroups of students, including race/ethnicity, students from low-income backgrounds, LGBTQ+ students, and English learners. • Developers include information about validation for each subgroup of students, such as an RMSEA or Rasch analysis for each student subgroup. • Developers provide a measure such as Cronbach's alpha of 0.7 or higher showing internal consistency across groups or differential item functioning at 0.64 or below. **Assessments meeting criteria:** • Inventory of School Climate • Cultivate • USED School Climate Survey
Does the assessment empower decisionmakers to improve outcomes?			
11. Do the developers offer a tool to easily interpret and understand the assessment data?	Developers do not provide support for interpreting assessment data. • The developer may have designed the assessment in a more academic setting and is not able to provide support through a commercial developer. **Assessments not meeting criteria:** • Inventory of School Climate • School Climate for Diversity	Developers have some guidance for interpreting assessment data, but do not have an easy-to-use tool that visualizes the assessment data, or they provide a tool with limited function. • The developer may provide a tool, platform, or reports, but they do not allow for the disaggregation of data. **Assessments partially meeting criteria:** • Self-Assessing Social and Emotional Instruction and Competencies	Developers have an easy-to-use tool that allows educators or decisionmakers to access and understand the assessment's data. This tool ensures the data is disaggregated to understand impacts on various student groups.. • The developer provides a tool, platform, or reports with data visualizations that can also be generated for specific subgroups of students. **Assessments meeting criteria:** • 5Essentials • Catalyze • Cultivate • DESSA • NJ School Climate Improvement Survey • Panorama SEL Survey • Elevate • USED School Climate Survey

Questions to ask	Does Not Meet Criteria *Continued*	Partially Meets Criteria *Continued*	Meets Criteria *Continued*
12. Have developers identified how assessment results are connected to policy and practice decisions?	Developers do not identify how outcomes are connected to policy or practice decisions. **Assessments not meeting criteria:** • n/a	Developers have some suggestions about how outcomes are connected to policy or practice decisions, but they are not clearly rooted in equity or do not promote evidence-based decisions. • The developer may provide resources for connecting outcomes to practice, but they are only available to subscribers, so it is not possible to tell if they are rooted in equity or evidence-based. **Assessments partially meeting criteria:** • 5Essentials • DESSA • Inventory of School Climate • Panorama SEL Survey • School Climate for Diversity	Developers clearly identify how outcomes are connected to policy or practice decisions and provide guidance or resources to support decision-making based on assessment data. The guidance/resources are rooted in equity (e.g., identify and address adult biases/inequitable practices) and promote evidence-based decision-making. • Developers link to research in their guidance on responding to assessment outcomes. • The guidance explicitly calls out biases and systemic inequities. • The guidance gives specific practice and policy guidance for each assessment item or section of assessment items. **Assessments meeting criteria:** • Catalyze • Cultivate • NJ School Climate Improvement Survey • Elevate • Self-Assessing Social and Emotional Instruction and Competencies • USED School Climate Survey