AIR American Institutes for Research | SEL Solutions at American Institutes for Research

Are You

READY TO ASSESS

Social and Emotional Learning and Development?

STOP Think Act

Brief

About *Stop, Think, Act: Ready to Assess*

Families, educators, practitioners, employers, and policymakers alike recognize the importance of ensuring that youth develop the social and emotional (SE) competencies (i.e., knowledge, attitudes, and skills) necessary to succeed in school, careers, and life. There is also widespread recognition of the importance of creating and supporting the conditions that are conducive to social and emotional learning (SEL) and development. For example, the National Commission on Social, Emotional, and Academic Development (SEAD) underscores the importance of positive, supportive relationships and rich, stimulating environments in promoting learning and development (Aspen Institute, 2018). In addition, the Science of Learning and Development (SoLD) Alliance advocates for high-quality learning environments and experiences (Cantor, Osher, Berg, Steyer, & Rose, 2018; Darling-Hammond, Flook, Cook-Harvey, Barron, & Osher, 2019; Osher, Cantor, Berg, Steyer, & Rose, 2018). The emphasis on conditions for learning and development extends beyond classrooms and schools; it holds true across all settings and sectors, including out-of-school time, postsecondary education, and the workforce.

Ready to Assess **includes the following settings and stages:**

Settings

- Classrooms
- Schools
- Out-of-school time
- Postsecondary education
- Workforce

Stages

- Early childhood (0–5 years)
- Elementary school (Grades K-5, 5–10 years)
- Middle and high school (Grades 6–12, 11–18 years)
- Workforce (18+ years)

In this edition of *Stop, Think, Act: Ready to Assess* (referred to as *Ready to Assess* throughout this brief) we expand our suite of resources to respond to the evolving SEL landscape and the demand from educators, practitioners, and policymakers for clear guidance and information on whether and how to measure SE competence and the conditions for learning and development across settings.

This *Ready to Assess Brief* and the accompanying *Decision Tree* and *Tools Index* aim to provide an update on:

- The evolving SE assessment landscape.
- A framework for and guidance on deciding *if* and *when* you are **ready to assess** conditions for learning and development and SE competencies.

What's New in This Edition of *Ready to Assess?*

In 2015, AIR released *Ready to Assess,* a suite of resources designed to help educators, practitioners, and policymakers decide whether and how to assess SE competencies. The *Ready to Assess* collection of resources included a brief, a decision tree, and a tools index.

Since then, the SE assessment landscape and the needs of educators, practitioners, and policymakers have evolved. Today, we know more about the conditions for learning and development that are essential for SE competency development. In addition, and partially in response to federal guidance, there is an increased focus on student engagement, school climate, and safety (Blad, 2016).

In this second edition of *Ready to Assess,* we have expanded our framework for and guidance on deciding *if* and *when* you are ready to assess SE competencies by highlighting the importance of implementation readiness and the conditions that foster SEL and development. We also have expanded the *Tools Index* in two ways:

1. We included measures of learning conditions (i.e., school climate, SEL implementation, and program quality) in and out of school.

2. We added to the list of available SE competency assessments to reflect updates to existing tools; and we included newly developed assessments, including performance-based measures and tools that assess other developmental stages, such as adolescence.

It is important to note what *Ready to Assess* is and what it is not.

Ready to Assess is a suite of resources, including a brief, a decision tree, and a tools index focused on SEL.

Ready to Assess provides educators, practitioners, and policymakers with a framework for and guidance on deciding if and when they are ready to assess conditions for learning and development and SE competencies. In order for tools to be included in the *Tool Index,* they must be intended for use by educators and practitioners and readily available for use.

The *Tool Index* is not designed to evaluate the currently available SEL assessment tools, nor is this guide intended to be an exhaustive list of educational assessments (for example, it does not include assessments intended for clinical settings).

Assessment Landscape

During the past 20 years, the practice of assessing youth learning and development has changed at a rapid pace in terms of its scope, sophistication, intensity, and frequency across settings and sectors. Measuring youth SE competencies remains important, and now there is an increased emphasis on assessing and improving the conditions that support SEL.

Policy

The Every Student Succeeds Act (ESSA) provides language suggesting that student engagement and whole-child development should be central to education agencies' plans for supporting student learning and development. Although ESSA maintains a focus on state accountability and standards, it also provides flexibility regarding how those requirements are met in exchange for rigorous and comprehensive plans focused on increasing outcomes for all youth. To comply with ESSA, states must identify an indicator of their choosing that focuses on school quality or student success. In that context, some states have chosen to focus on the conditions for learning and development, such as school climate and safety. There are additional policy efforts at the state and local levels to support adult capacity in implementing and integrating SEL, the adoption of evidenced-based SEL programs and practices, and cross sector partnerships focused on SEL (Dusenbury & Weissberg, 2018).

Working Groups, Collaborations, and Reports

The National Commission on Social, Emotional, and Academic Development (SEAD) released the report titled *From a Nation at Risk to a Nation at Hope*, which calls for opportunities to foster whole-child development in school and out-of-school time settings, as well as for the integration of social, emotional, and academic competencies (Aspen Institute, 2018). Similarly, the Science of Learning and Development (SoLD) Alliance has published reports telling us that whole-child development is a personalized journey and that context and relationships are defining influencers of development (Cantor, Osher, Berg, Steyer, & Rose, 2018; Osher, Cantor, Berg, Steyer, & Rose, 2018; Darling-Hammond, Flook, Cook-Harvey, Barron, & Osher, 2019).

The Collaborative for Academic, Social, and Emotional Learning (CASEL) has established work groups to explore, define, and suggest strategies for practical assessment of SEL. For example, the *Establishing Practical Social-Emotional Competence Assessments of Preschool to High School Students* project (i.e., the Assessment Work Group) was a multidisciplinary collaborative of leading researchers and practitioners in the fields of PreK–12 education, assessment, SEL, and related fields. In addition, the National Practitioner Advisory Group on Using Data to Inspire SEL Practice is a group of approximately 30 educators, state and district education staff, and nonprofit leaders from across the country with expertise in SEL.

In addition to *Ready to Assess,* other tools are available for those seeking to learn about SEL measures. They include RAND's *Education Assessment Finder* and the *SEL Assessment Guide* from CASEL. These tools summarize SE competency assessments for use by practitioners and policymakers and complement *Ready to Assess.*

Ready to Implement Before You Are Ready to Assess

The first edition of *Ready to Assess* focused exclusively on assessing SE competencies. Effective assessment of SE competencies can yield significant benefits for youth, educators, practitioners, and policymakers. But if those assessments are to be implemented efficiently and productively, it is critical to first take a step back to understand your readiness for SEL implementation and whether and how you are creating the right conditions for SEL and development.

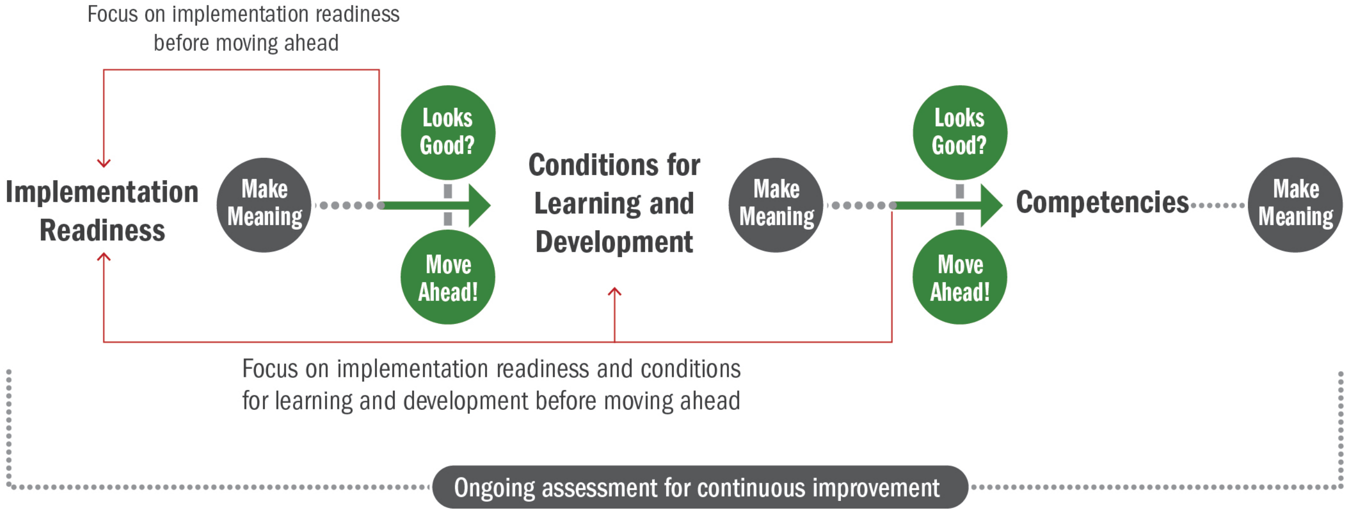
In this section, we expand on our framework for selecting and administering an assessment by introducing two considerations—**implementation readiness** and **conditions for learning and development.** The following logic model can help you to determine if:

1. You are ready to implement an SEL program or initiative.
2. You have created conditions that support learning and development (including school climate, intentional implementation of SEL, and program quality).

   If you have crossed both of those milestones,

3. You are ready to assess youth SE competencies.

SEL Logic Model



As our logic model indicates, implementation readiness and conditions for learning and development are essential to establish before assessing SE competencies. In the next sections, we describe these two considerations in more detail.

Implementation Readiness

Before jumping into assessment, it is critical to explore your *readiness to implement.* This means you must first ensure that professionals at your school or organization are ready to implement intentional strategies to support SEL in conditions that are conducive to SEL.

Implementation readiness involves:

1. The motivation to implement something new
2. The general capacities of an organization to support implementation
3. The specific capacities that are needed for this particular innovation

Scaccia et al. (2015) suggest a useful way to heuristically conceptualize implementation readiness by applying the equation $R=MC^2$, where *R* refers to readiness, *M* refers to motivation, and *C* refers to the two kinds of capacity—general and specific. Readiness considerations are useful prior to implementation; however, the road to readiness is not linear. Rather, it can be cyclical in nature, and readiness-focused thinking is intended to inform continuous improvement. Agencies, schools, and organizations may move back and forth in their readiness to implement as they recalibrate and improve the quality of their implementation. As such, even if your school or organization is already engaged in implementation, it may be useful to consider the current state of readiness (i.e., motivation and capacity) to see if there are opportunities for improvements that will ultimately ensure successful implementation of SEL.

Although a deeper exploration of readiness is beyond the purview of this brief, we encourage you to carefully consider and reflect on your organizational readiness and change management processes (and stay tuned for more readiness thinking in the next edition of *Ready to Assess*).

Conditions for Learning and Development

If and when you have determined that your organization is ready (or sufficiently close to being ready) for implementation, it is critical to consider the conditions for learning and development. Extensive research (Berg, Osher, Moroney, & Yoder, 2017; Cantor, Osher, Berg, Steyer, & Rose, 2018; Osher, Cantor, Berg, Steyer, & Rose, 2018) points to the critical role of the environment in supporting learning and development. We know that young people, as well as adults, are better able to learn in safe and supportive spaces, where relationships can flourish and there is freedom to learn, practice, make mistakes, and receive feedback (see, for example, Center on the Developing Child, 2016; Jones & Bouffard, 2012; Newman, Dymnicki, Fergus, Weissberg, & Osher, 2018).

We also know that competence building is most effective when the process of SEL is intentional (i.e., implemented on purpose); is sequenced, active, focused, and explicit (better known as "SAFE"; see Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011); is implemented across settings; and is embedded into everyday practice (Cantor, Osher, Berg, Steyer, & Rose, 2018; Osher, Cantor, Berg, Steyer, & Rose, 2018).

The *Tools Index* includes assessments of school climate, SEL implementation, and program quality. These assessments help you understand whether and how you are creating the conditions for SEL and development. They can be used separately (we recommend starting here) or concurrently with an assessment of youth SE competencies (after you have determined that the right conditions are in place). When the conditions for learning and development are in place and you are implementing intentional SEL, it then becomes reasonable to explore how young people are developing. At this point, the use of an SE competency assessment may be beneficial.

Are You Ready to Assess?

In the sections that follow, we review four considerations in selecting and implementing an assessment: purpose, rigor, practicality and burden, and ethics. The *Tools Index* discussed in this edition of *Ready to Assess* includes assessments of conditions for learning and development (school climate, SEL implementation, and program quality) and SE competencies. The four considerations presented here apply to assessments of all types, including those that focus on implementation readiness.

We recommend starting with the following question: Why are we using this assessment?

Purpose of the Assessment

The decision to use any assessment must be grounded in a clear, well-founded purpose. The purpose(s) of an assessment may be highly nuanced and can easily become confounded (e.g., competing interests, mixed messaging, or mismatched expectations). Three common assessment purposes are (1) information, (2) communication, and (3) accountability.

Information

One of the great promises of using assessment data is that the information will contribute to well-informed decision making at the local, state, and federal levels. Gathering information using assessments can serve many purposes: it may be broadly exploratory or it can inform a need for changes to policies or practices to better support conditions that facilitate the learning and development of SE competencies; it can help improve performance or practice by providing formative feedback to youth and practitioners, build capacity, or determine further professional development efforts; and it can provide proof of the effectiveness of policy or practice. Finally, assessments can be used to identify youth needs for more intensive support (e.g., individualized education programs), to provide intentional instruction, or to guide intervention.

Communication

Educators, practitioners, and policymakers communicate regularly with diverse audiences and for many different purposes. At the various levels (e.g., school, organization), educators, practitioners, and policymakers may need supporting evidence to make their case for a need, to satisfy a request for information, or to communicate with external stakeholders (e.g., parents, industry, the community at large). In such cases, the type of assessment will vary with the audience and intended outcomes. Demonstrating compelling evidence to advocate for or against a given policy may be a relatively high-stakes objective and may require large-scale, highly rigorous assessments. Smaller-scale, tailored assessments can be used to meet lower stakes, isolated requests for information or explanation; or, they can be used for "storytelling" purposes.

Accountability

When you select an assessment to use for accountability purposes, the type of assessment you choose may vary depending on the stakeholders to whom you are accountable and the implications or consequences of meeting established requirements. Policymakers concerned with accountability to standards and establishing funding for various programs face a high-stakes effort that may demand the most rigorous evaluations requiring large-scale, standardized assessment. Educators and practitioners concerned with demonstrating more local impact or general improvement objectives may face lower stakes and thus may be better served by smaller scale, customizable assessments.

STOP

After considering the many purposes that assessments might serve and then choosing one or more as the basis for using an assessment, stop and consider whether the purpose necessitates assessment. Although using an assessment may be highly appealing because of its potential power in validating results, often youth-, school-, organization-, district-, or state-level data already exist to meet these and many other needs. It is worthwhile to ask whether more data are needed or whether the need can be addressed in alternative ways; for example, by using existing data and assessments or by using a proxy (e.g., assessing elements of school climate). Finally, at each stage of the decision process, consider the risks and benefits of pursuing these goals with assessment methods. If assessment is the best option, then the next consideration in the *Ready to Assess* framework is rigor.

Rigor

After firmly establishing the rationale for using an assessment and the stakes involved, it is important to determine the rigor of the prospective assessment. We encourage practitioners to rigorously implement any assessment under consideration; however, when considering the term "rigor" in *Ready to Asses,* we are referring to the:

1. Comprehensiveness of the assessment.
2. Degree to which the assessment is a well-established, valid, and reliable measure.
3. Assessment's implementation and use.

Assessment Type

The first dimension of rigor to consider is the assessment type, which may vary depending on whether the purpose is relatively high stakes or low stakes. Less rigorous assessments may be appropriate for lower-stakes purposes, such as information gathering and communication—especially at the local level by a single school or program. More rigorous assessments may be required for higher-stakes accountability purposes, especially at organization-, district-, and state-wide levels of reporting. However, in some cases, the stakes involved may not match the desired level of rigor. It may be appropriate, or even highly desirable, to include some lower-rigor assessments for high-stakes purposes, especially if those assessments are implemented rigorously and if they provide depth to the assessment results. Similarly, higher-rigor assessments can be used effectively in low-stakes contexts if certain practical implementation considerations—such as capacity and program maturity—can be ensured.

9311a_09/19

Technical Properties of the Assessment

The second dimension of rigor to consider relates to the technical properties of the assessment, which depend largely on its levels of validity and reliability. Ideally, the assessment under consideration has a record of successful use by other state and local actors for similar purposes. This will help stakeholders determine the significance of the assessment results within the larger picture of all other education data. An assessment should have high levels of reliability and validity to be considered for use on a large scale or for decision making. In some cases, multiple assessments or assessment types may be necessary to achieve the established purposes.

Administration of Assessment

The third dimension of rigor to consider is administration, which refers to the rigor with which an assessment is implemented and used. Some assessments are validated for external assessment, while others may also be used for self-reflection. Some assessments require multiple observations or raters, while others function well with a single (i.e., point-in-time) rating. It is important to consider whether a chosen assessment can be administered with fidelity—consistently and in accordance with its guidelines—and based on those guidelines, the ways in which the data may be interpreted and used.

Validity: The degree to which an assessment can be said to measure what it is supposed to be measuring (e.g., communication skills), instead of something else (e.g., mathematics ability).

Reliability: The degree to which an assessment can be expected to produce the same results after being administered multiple times to the same population.

A theory of action: An "if–then" statement that articulates the mechanisms by which the desired outcomes will be achieved via the selected means. A theory of action should clearly state how using the assessment—including its implementation and the analysis of its results—will lead to achieving your identified purposes.

After all the elements of rigor are determined to be sufficient for the purposes identified, it is critical to develop a theory of action and a concrete plan for how to use the assessment results. Ultimately, this plan will determine whether even highly significant and meaningful results can be used to serve the intended purpose of the assessment. For example, if the assessment does not have strong evidence of validity and reliability, then that potential issue should be clearly communicated, and the results should be considered exploratory rather than final or definite. Conversely, the decision to assess at all may be postponed until higher levels of validity and reliability can be ensured.

THINK

After reviewing the key elements of rigor, think critically about how the rigor of your chosen assessment maps onto your purposes. It may become clear that there is not a good match between the level of rigor of the available assessments and your intended purpose. In some cases, you might be able to take less formal measures, such as teacher reports, parent-teacher conferences, out-of-school time supports, or routine counseling or other services, to achieve the same outcomes. If these options are not adequate, and if the rigor of your chosen assessment matches the purpose you have identified, then the next step is to consider practicality and burden.

Assessment Practicality and Burden

After developing a clear sense of purpose and of the assessment's rigor, it is paramount to outline the relevant practical considerations and estimate burden (or implementation costs). Even with the most clearly articulated purpose and correspondingly rigorous assessments, a disconnect between those factors and real-world practicalities and burdens could derail your process.

Practicality

Consider two key practical considerations before estimating burden: (1) the age of the SEL program or initiative and, if you are assessing youth SE competencies, (2) the number of youth engaged in the SEL program or initiative. Each of these factors may vary independent of one another, so it is important to consider the possible permutations between them. In the case of new programs serving small numbers of youth, it may not be feasible to implement high-rigor assessments. Beyond feasibility, high-rigor assessment results may not be used with maximum effectiveness when applied to small groups or with little opportunity to reveal change over time. In such cases, it may be more appropriate to use a relatively low-rigor assessment or to refrain from assessing altogether until the program or initiative is more mature and a larger number of youth can be evaluated.

On the other hand, an assessment of a mature program with a large number of youth may yield a much greater return on investment when using a high-rigor assessment instead of implementing a relatively low-rigor evaluation only to save costs. These general guidelines follow from the fact that high-rigor assessments typically have higher levels of validity and reliability but are more complex and costly to implement, while comparatively low-rigor assessments are less likely to have high, well-established levels of validity and reliability but are much less complex and costly to implement (Soland, Hamilton, & Stecher, 2013). Ensuring a proper articulation of purpose, rigor, and practicality will allow for the most accurate and productive comparison of potential benefits against implementation costs (or burden).

Burden

Elements of burden include staff capacity, infrastructure requirements, data use, budget, and risks to teachers, staff, youth, and families. These barriers to implementation can limit the maximum return on investment already identified via the establishment and articulation of purpose, rigor, and assessment practicality. It is important to consider each element of burden carefully before continuing with assessment plans.

Ensuring staff capacity requires training staff on assessment administration and data collection, data analysis, and reporting—or contracting those services to another agency. When evaluating infrastructure, it is important to consider mechanisms for data collection (e.g., computerized assessments), data storage, and data analysis tools. The use of assessment data requires a plan for data analysis, as well as the continued use of that data and data from future assessments. All of these factors involve unique costs, both monetary and relative to labor and time. Finally, consider the burden on teachers, staff, youth, and families given the implementation costs and identified capacity as potential barriers to achieving maximum return on investment. After weighing the burdens you have identified against the potential benefits of assessing, the final checkpoint before acting is a consideration of ethics.

Ethics

Once the you have established the purpose, rigor, practicalities, and burden of using an assessment and developed a plan for its rollout has been developed, it is time to stop and think before acting—to do a final check for ethics and to consider the big picture. At this time, convene your team members and evaluate how use of the assessment and its associated potential risks and benefits will align with the larger mission of the institution and the community. Key ethical questions you may ask include:

1. Does this assessment align with our organizational mission, values, and purpose?
2. Does the benefit of having the data outweigh the risks to participants of collecting it?
3. Are we administering an assessment to a group already burdened with surveys and observations?

4. Is our purpose aligned with the rigor of the assessment?

5. Does this team comprise people with the appropriate knowledge, skills, and training necessary to select an appropriate assessment?

6. Do we have the capacity to ensure that the assessment is administered in a standardized manner, with appropriate accommodations for people of all abilities and languages?

7. Is there evidence for using the assessment with our intended groups?

8. Will our interpretation of the assessment results fit within the intended purpose and use of the assessment?

9. How will we use data from the assessment and make it meaningful to the people from whom it was gathered?

10. How will we ensure that we highlight strengths and areas for growth (avoiding stigmatizing language) when communicating assessment results?

Consider any other options to reduce the risk and maximize the benefit of the outlined plan. Use the *Decision Tree*—to help ensure that after deeply considering these four components, the decision to use (or not to use) an assessment is of greatest potential value.

ACT

After fully mapping the connections between your assessment purpose, rigor, and burden, and after determining that the potential value of the assessment in question matches your ability to implement it, ethically and at relatively low cost, it is time to take action. The *Decision Tree* and *Tools Index* can be useful references as you begin to implement or refine your plan. Keeping all of the above considerations in mind throughout the process of developing and implementing your assessment plan can help to ensure that assessment goals truly meet education needs.

References

Aspen Institute. (2018). *From a nation at risk to a nation at hope.* Retrieved from http://nationathope.org/wp-content/uploads/2018_aspen_final-report_full_webversion.pdf

Berg, J., Osher, D., Moroney, D., & Yoder, N. (2017, February). *The Intersection of school climate and social and emotional development.* Washington, DC: American Institutes for Research. Retrieved from https://www.air.org/sites/default/files/downloads/report/Intersection-School-Climate-and-Social-and-Emotional-Development-February-2017.pdf

Blad, E. (2016). *ESSA law broadens definition of school success.* Retrieved from https://www.edweek.org/ew/articles/2016/01/06/essa-law-broadens-definition-of-school-success.html

Cantor, P., Osher, D., Berg, J., Steyer, L., & Rose, T. (2018, January 24). Malleability, plasticity, and individuality: How children learn and develop in context. Retrieved from https://tdlc.ucsd.edu/global/images/Cantor_etal.pdf

Center on the Developing Child. (2016). *From best practices to breakthrough impacts: A science-based approach to building a more promising future for young children and families.* Cambridge, MA: Harvard University, Center on the Developing Child.

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2019). Implications for educational practice of the science of learning and development, *Applied Developmental Science,* doi:10.1080/10888691.2018.1537791

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82,* 405–432.

Dusenbury, L., & Weissberg, R. (2018, June). *Emerging insights from states' efforts to strengthen social and emotional learning.* Chicago, IL: Collaborative for Academic, Social, and Emotional Learning. Retrieved from https://casel.org/wp-content/uploads/2018/06/CSI-Insights.pdf

Jones, S. M., & Bouffard, S. M. (2012). Social and emotional learning in schools: From programs to strategies. *Society for Research in Child Development Social Policy Report. 26*(4), 1–33. Retrieved from http://files.eric.ed.gov/fulltext/ED540203.pdf

Newman, J. Z., Dymnicki, A., Fergus, E., Weissberg, R. P., & Osher, D. (2018). Social and emotional learning matters. In D. Osher, D. A. Moroney, & S. Williamson (Eds.), *Creating safe, equitable, engaging schools: A comprehensive, evidence-based approach to supporting students* (pp. 213–222). Cambridge, MA: Harvard Education Press.

Osher, D., Cantor, P., Berg, J., Steyer, L., & Rose, T. (2018). Drivers of human development: How relationships and context shape learning and development. doi:10.1080/10888691.2017.1398650

Scaccia, J. P., Cook, B. S., Lamont, A., Wandersman, A., Castellow, J., Katz, J., & Beidas, R. S. (2015). A practical implementation science heuristic for organizational readiness: $R=MC^2$. *Journal of Community Psychology, 43*(4), 484–501.

Soland, J., Hamilton, L. S., & Stecher, B. M. (2013). *Measuring 21st century competencies: Guidance for educators.* Washington, DC: RAND Corporation.

9311a_09/19