

Chapter Five. Summary of Challenge Findings and Discussion

The SEL Challenge was undertaken in pursuit of two ambitious goals: Identify promising practices for building SEL skills with vulnerable adolescents, and develop technical supports for use of these SEL practices at scale in thousands of OST settings¹. The promising practices are featured in an SEL Field Guide, *Preparing Youth to Thrive: Promising Practices for Social and Emotional Learning* (Smith, McGovern et al. 2016), this technical report, and a suite of tools and technical assistance (SELpractices.org).

In the Challenge, we used a qualitative methodology with expert practitioners to identify promising practices, or standards, for SEL practice. We also used performance studies to provide validation evidence for the standards and benchmarks for use as a comparison by other organizations. Both parts of the work provided insight into how SEL practices might be taken to scale in an OST field that is uniquely positioned to build SEL skills with all children and youth. The OST sector is already tuned to youth who can benefit from SEL – bringing vulnerable youth into settings designed to establish youth’s sense of safety and then engage youth’s interests. Further, the sector has invested extensively in continuous improvement policies and systems that are well suited to introduce, improve, and exemplify SEL practices.

The tools and technical assistance developed through the SEL Challenge are designed to help OST programs focus deeply on SEL practice, assess their strengths, and improve the quality and effectiveness of their services over multiple annual cycles. The suite of performance measures developed through the Challenge is designed to fit with an evidence-based, continuous improvement intervention used widely in the OST field. This means that the continuous improvement intervention can be used to improve the quality of contexts focused on SEL skill learning and, therefore, to increase SEL skill by employing the suite of SEL measures described in this report. In this chapter, we summarize four primary findings from the Challenge and discuss implications of these findings for SEL measurement and for taking SEL practices to greater scale.

¹ Again, the term out-of-school time is used to refer to settings variously labeled afterschool, expanded learning, extracurricular clubs, summer camps, and sports; many mentoring, tutoring, apprenticeship, and workforce development programs; programs for disconnected and homeless youth; and some alternative schools.

Summary of Challenge Study Findings

Findings from the SEL Challenge include the standards for SEL practice, information about the suite of SEL performance measures, and the benchmarks from the performance studies. These findings can be summarized as follows:

(1) The SEL Challenge methodology successfully identified exemplary SEL offerings and produced 34 standards, 78 practice indicators, and 327 vignettes for building SEL skills with vulnerable youth. The SEL Field Guide, Preparing Youth to Thrive: Promising Practices for Social and Emotional Learning, is the primary presentation of findings for the Challenge. The standards are spread over six domains of practice (i.e., emotion management, empathy, teamwork, responsibility, initiative, and problem solving) and a set of five curriculum features. The successful selection of expert practitioners and exemplary offerings, and the validity of standards that were produced in partnership with these experts, was evaluated through performance studies using a suite of SEL performance measures designed for this purpose. Results from these studies indicated that the offerings were indeed exemplary.

(2) The suite of SEL performance measures developed for the Challenge is feasible to implement and demonstrates sufficient reliability and validity for both continuous improvement and evaluation uses. The suite of performance measures was feasible to implement using technical supports (e.g., training raters, data entry portal, site performance reports) in the typical range for many other implementations in the OST field. With some important caveats, the data produced by the performance measures demonstrate sufficient reliability and validity for use as part of a lower-stakes continuous improvement intervention and for more evaluative purposes where it is necessary to reliably differentiate among settings, individuals, and time points.

A theory for SEL skill measurement was assembled to differentiate among several mental processes related to skill learning (e.g., automatic activation of mental contents by the context; intentional focusing of attention and awareness on the interactions between context, mental contents and behavior; automatic and intentional processing of affectively-charged schemas and scripts) that were directly targeted by the specific standards and curriculum features. This was an opportunity to fit theory about the multilevel person system to prior work on the multilevel context system, extending the continuous improvement intervention from the policy level through a cascade of effects on settings, and ultimately, to intra-individual SEL skill growth.

(3) The performance studies indicate that the exemplary offerings were exceptionally high quality compared to other OST programs and that youth skills improved in all six SEL domains. Skill growth also occurred for the higher risk groups. Benchmarks for SEL performance include:

(3.a) Diverse staff and youth, intensive participation, and expert adult guidance. The Challenge offerings were diverse in terms of ethnicity and risk. The program staff intentionally recruited ethnically-diverse youth; overall, the Challenge cohort was 48% African American, 30% Hispanic, 14% White, and 8% Asian or Pacific Islander. In more ethnically-homogenous offerings, staff's ethnicity reflected the youth's ethnicity. All of the offerings targeted vulnerable youth, and these youth also represented diversity of SEL strengths and more difficult SEL histories. A total of 35% of the youth were identified as higher risk, but only 5% had more than one risk indicator.

All of the SEL offerings were intensive commitments for the youth, ranging between 20 and 75 sessions and between 39 and 370 contact hours. Almost all staff had a college degree and, in over half of the organizations, at least one team member had an advanced degree. The Challenge organizations had what we would characterize as lower staff turnover, and lead instructors' tenure ranged between eight months and 20 years, with one third of the staff in their current position for five years or more. Staff reported having expertise in their offering, but not necessarily in SEL, where only 50% of program managers and 11% of lead instructors rated themselves as experts.

(3.b) Highly collaborative organizational cultures. SEL Challenge organizations performed higher than the comparison group in all measures of culture and climate. In particular, substantially higher performance on both staff-to-manager and staff-to-staff collaborative practices reflect the importance of staff supports identified in the curriculum features and opportunities to model SEL skills identified in the standards.

(3.c) Exceptionally high-quality instruction and youth engagement. Challenge offerings were exceptionally high-performing contexts for two types of instructional quality: The quality of the project curriculum (Growth and Mastery scale), and the quality of the staff SEL practices (Instructional Total Score), were substantially higher than the comparison group. Staff SEL practices and youth engagement were assessed at three time points. Almost all youth reported very high engagement with the context at all time points.

(3.d) A consistent pattern of positive SEL skill growth across measures, offerings, and risk status. Youth SEL skills, as indicated by youth beliefs and behaviors, increased during the offering cycle. Three-time-point growth models demonstrated positive change on almost all measures in all six domains. These models also indicated that youth who entered the program at higher risk (i.e., in a low-skill subgroup at baseline or in a subgroup indicated by measures of attachment-related anxiety, avoidance, and social phobia) also improved as much or more, on average, than students who started out with higher SEL skills.

(4) The exemplary offerings shared an OST-SEL intervention design: project-based learning with co-regulation. In addition to using the SEL practices identified by the standards, the exemplary offerings shared several curriculum features: intensive participation in challenging project curricula; SEL curricula

that include responsive practices and structured check-ins; the cycle-in, cycle-out sequence focused on deeper engagement with youth; and a broad and integrated approach to implementation of the SEL practices in the six domains identified by the standards. Together, these curriculum features constitute an OST-SEL intervention design – project based learning with co-regulation – for offerings with a primary purpose to build SEL skills with vulnerable adolescents.

Discussion

Efforts to define high-quality instructional practices for vulnerable youth are not new ([American Psychological Association Coalition for Psychology in Schools and Education, 2015](#); [Eccles & Gootman, 2002](#); [Gideonse, 1988](#)), and we are fairly certain that there are few new practices or ideas named in the Challenge study. Rather, the primary significance of the study was the opportunity to systematically describe the point of service level as a consistently granular set of elements (e.g., in terms of adult behavior, basic instructional processes, and short-term youth skill change) and then align to these elements both vignettes in the voice of expert practitioners and performance measures implemented at the Challenge sites. We believe that this is one of the best ways to move the policy agenda for high-quality OST forward because good measurement requires adequate description, and what gets measured can be moved—and funded.

In relation to description and measurement, during our work on the Challenge it quickly became clear that two areas of SEL research require greater investment and investigation. First, more effort is required to develop a consensus framework for thinking about the nature of, and interrelations among, contexts, skills, and behavior related to SEL. We hope that our introduction of a multilevel person-in-context framework and, in particular, the Basic Levels of Self model (see Chapter 3, Figure 3, and Appendix C) to the OST field will help make a contribution in this regard. In particular, we believe that many discussions of SEL, or skill learning, can be usefully framed in relation to two types of agency experience, automatic and intentional. Automatic forms of agency tend to flow naturally from encounters with contexts that promote feelings of safety, support, and efficacy. Intentional forms of agency require active self-reflection on thoughts and feelings about goals, skills, and behavioral options in relation to the context. Given that SEL skills, like all skills, are learned in context as youth practice self-regulation for their own specific purposes, we suspect that any consensual framework sufficient for integrating the diverse literature on SEL skill growth and transfer will need to attend explicitly to both forms of agency and their relations to context quality and behavior.

Second, integrated with the effort to develop a consensus framework for SEL context and skills, more attention needs to be given to measurement issues. In particular, given the relatively loose coupling between our theory of the intra-individual dynamics of SEL skill building and the suite of measures that

we recommend in this report, our ability to evaluate and document empirically the cascade of causal effects from context quality to skill growth and transfer depends critically on developing more specific measures of the mental contents and processes hypothesized to mediate these effects. These issues—clarification of theories and improvement of measures—go hand-in-hand and, we believe, substantial improvement could be made in a fairly short amount of time with sufficient investments.

Finally, we believe that local communities have critical roles to play in SEL skill building, beyond the simple dictum that community-based providers should use evidence-based programs identified by researchers. Because high-fidelity implementation matters and has a cost, we hope that local policy makers and funders will use the Challenge as a template for identifying the exemplary SEL services already available in their communities and make sure that they are adequately recognized, resourced, and replicated. Conversely, we also hope that the many organizations already doing high-quality SEL work can use the products of the Challenge to make the case for their work to local leaders. We know from experience that there are SEL experts in every community, and while it can take many years and substantial resources to retrain a regional workforce, identification of currently existing expertise is likely a cost-effective first step. Further, it is likely that different SEL practices are important in different communities with different SEL histories and that local experts may be best able to design interventions and select performance measures that are most fit for their youth.

This last point is particularly important when considering our findings related to discrete staff practices and intervention design. First, the standards for SEL practice present the opportunity to consider any single standard as a discrete area of practice that an organization may want to assure or improve. In this sense, a program director might select a single standard in the Emotion Management domain (e.g., “Staff model healthy strategies for dealing with emotion within the context of caring mutually respectful relationships with youth”) as an area for improvement in their programming. Given that there are many standards to potentially focus on, there will likely be wide variation across places and settings on which standards are selected.

However, in Chapter Two, we also identified a set of common curriculum features, an OST-SEL intervention design, which we referred to as *project-based learning with intensive co-regulation*. Despite differences in skills learned through the Challenge offerings’ project curricula (e.g., building boats, community organizing through the arts), the offerings had common features that included an SEL curriculum with responsive practices and planned check-ins; sequencing of SEL content over time; the cycle in, cycle out method; the safe space where staff come to know youth deeply; *and a broad and integrated use of SEL practices in all six domains*. For vulnerable youth who have had difficult SEL histories (e.g., exposure to traumatic events or chronic stressors), this intervention design may be of

particular importance. We know from experience that most communities have offerings that aspire to deliver the OST-SEL intervention and that every community has vulnerable youth. We believe that the

OST-SEL intervention has wide applicability in all communities, although not necessarily in all OST settings. In both cases (i.e., the selection of specific standards for improvement and/or the adoption of the OST-SEL intervention design), decisions about selection or adoption should be influenced by local decision makers who know the experiences of the youth and the resources available in the OST program setting and its immediate community.

Generalizability of Findings

The study design used in the Challenge supports extension to other settings and populations in two ways. First, by focusing systematically on the details of adult and youth behavior, the Challenge content supports interpretation, adoption, and adaptation by expert practitioners – professionals who are trained to think about OST and other learning contexts where vulnerable youth participate. Second, we identified a comparison sample of OST programs to provide a comparative reference for levels of performance described using the suite of SEL measures assembled for the study.

Extending from the mixed-methods study design, we offer several informed opinions about the generalizability of the findings to a wider field of practice-oriented settings and systems—OST programs as well as school day classrooms and settings in the behavioral health and juvenile justice fields—where the qualities of adult-youth interaction and learning are a primary concern. The most generalizable aspect of the Challenge is the common sense lesson that it reinforces about youth development: All youth need opportunities to develop social and emotional skills but, in particular, vulnerable youth with difficult SEL histories need adults who care about them and support them to have repeated experiences of agency (i.e., practicing self-regulation) in its more passive and active forms. Across the standards, and across the offerings from the SEL Challenge, the adults got to know the youth well, learned about their lives and experiences using safe and professional methods, and co-created opportunities with youth to engage and extend work projects with real world significance. SEL skill learning occurs in settings with these general qualities – described by the standards – that emphasize practicing at self-regulation, using both forms of agency, to generate mastery experiences that promote skill growth and transfer.

We also used the Challenge performance data to develop benchmarks for the range of performance in the exemplary offerings and then compared these benchmarks to performance levels from other, more generic, OST offerings. On average, 44% of the offerings in the comparison group were operating in the benchmark range on at least one of the benchmark measures, suggesting that (a) not all OST programs are trying to be OST-SEL interventions for vulnerable youth and (b) some programs are operating in this high range of performance that describes the OST-SEL intervention. These findings are consistent with our

thinking that, although many of the SEL practices can be successfully used in almost all settings in almost all communities, the OST-SEL intervention design may be applicable in only some of the settings in all communities.

The SEL standards, performance measures, and other Challenge products were designed to fit with an evidence-based continuous improvement intervention for OST organizations. The objectives of the intervention are a cascade of effects over multiple levels of setting to culminate in high-quality instruction, youth skill mastery, and skill transfer. This approach to building SEL skills using a continuous improvement intervention is already widely used in the OST field, and there are clear analogies to other sectors. For example, much of the infrastructure for SEL performance measurement described in this report is already available in local and state education agencies and supported in state laws and education-agency policies. Further, the SEL practices described in the standards demonstrate substantial overlap with best practices for teaching and learning (see Endnotes iv and viii). We hope that novel solutions lie ahead that use these existing education policies (e.g., evidence-based school reform provisions in the Every Student Succeeds Act) to achieve continuous improvement goals by integrating lower-stakes teacher evaluation, observation-based assessment of teacher practices, and teacher ratings of youth SEL behavior.

While the discrete practices named in the standards have high alignment with best practices in teaching, we believe that the more intensive OST-SEL intervention design – and the performance benchmarks developed for the suite of SEL measures – have direct applicability in the behavioral health and juvenile justice sectors. As residential treatment solutions continue to phase out, and more intensive community-based solutions are developed, this OST-SEL intervention design may be a useful tool for professionals seeking prevention interventions with broad applicability for youth who have had difficult SEL histories. Further, as the juvenile justice field struggles with the decoupling of court disposition and confinement-based policies, the OST-SEL intervention may be a viable alternative treatment to prevent development of core criminological factors.

Recent meta-analytic evidence suggests that features of effective programs have alignment with many of the SEL standards and curriculum features described in the Challenge offerings (Lipsey and Howell et al, 2010). In particular, the OST-SEL intervention design explicitly supports the more automatic and fast moving experience of agency by creating a context where youth can access and practice the skills they already have without experiencing barriers posed by lack of interest or cues that trigger strong emotion-behavior scripts. Second, the design also supports a deeper form of agency, more intentional and extended through time, as the curriculum brings specific aspects of the context and youth’s mental engagement with the context into the conscious focus of attention and awareness.

Finally, the OST-SEL intervention design requires staff with sufficient skills and with sufficient organizational supports to use these skills intentionally. Almost all of the exemplary offerings had at least

one staff member trained as a social worker or counselor in the setting at all times, and almost all had staff-to-student ratios at or below 1-to-8. Conducting SEL practice at high quality will require both investments in staff preparation and, in many cases, investments in more staff. Further, almost all of the expert practitioner teams in the Challenge had at least one member who was a career practitioner of their craft. This suggests, again, that part of the short-term path to improving SEL skills at scale is in identifying expert practitioners who are already working in most communities and who have already been developing curricula with groups of youth who present with SEL needs.

Cautions for SEL Measurement

The concept of taking SEL practices to scale using a continuous improvement intervention asks and answers relevant questions about what constitutes a good SEL measure. We argue that SEL measures are “good” to the extent that they fit into a coherent theory for the use of SEL data produced with the measures and that the use of the SEL data results in positive change in the well-being of youth. This thinking reflects the concept of consequential validity (Messick, 1995) and common sense. Recent questions about SEL measurement (Hoffman, 2009), and current debates over how to best measure SEL in schools (Martin, 2016; McEvers, 2016), may miss this logic given the climate of achievement testing and higher stakes accountability policies. In particular, individual SEL measures designed similarly to (a) maximize reliable differentiation between students who might be very near in actual performance and (b) predict future performance on similar tests (e.g., like an achievement test) may fall short on several counts.

First, one logical use of such information is to produce negative social comparisons of students and to create an artificial sense of certainty about assessment of individual performance. The over-reliance on achievement-test types of measures (i.e., point in time, context independent, youth self-report) for SEL invites policies that include higher-stakes social comparison of schools, teachers, and students, almost assuring a negative experience of judgement and distrust of the assessment process. Further, an over-valuing of the precision and predictive validity of SEL measures is likely to produce self-fulfilling prophecies. In this circumstance, the youth least well supported by the context do the worst on the tests, and the tests are there to prove it. Then, the aspects of educational service that the professional staff can actually address—quality of the context and the youth’s mental engagement with that context—are all but ignored by the accountability system. Our theory work in Chapter Three was specifically designed to raise this point—the press of context really matters and, if that press is not adjusted correctly for some youth to mentally engage, there will be clear winners and losers.

Second, use of SEL data will only improve SEL skill learning if the data means something to the staff implementing SEL practices, the people who themselves are critical parts of the context in which

SEL skills grow. To understand and work on SEL skill learning, our experience and evidence suggest that we need data that first and foremost describes the contexts well as context-specific mental engagement and behavior. Our response to this logic led us to develop the suite of SEL measures described in this report, which include staff reports about the setting and their supports, independent observations of the setting, youth reports of mental engagement with the setting, and teacher ratings of youth behavior demonstrated in the setting. All of these measures balance the need for precision with the need for the items to describe objective conditions and behaviors that support action by adults. Youth workers and teachers need useful feedback about their performance, the context of learning that they are accountable to create, in order to affect change.

Study Strengths and Limitations

The findings from the Challenge must be understood within the strengths and limitations of the study design. Most importantly, the qualitative method and performance measures were implemented on the basis of strong theory about the parts and processes that should be present in exemplary SEL offerings. The qualitative study to produce the standards drew from an extensive qualitative evidence base focused on staff practice and the standards reflect consensus among a diverse group of exemplary providers. The performance studies also drew upon strong theory and included measures with a validation history and a reference database for comparison, bolstering our interpretation of the pattern of results. Specifically, the combination of exceptionally high quality of instruction and youth engagement provides a strong expectation that skills will grow in these settings and supports our finding that skills did in fact increase over time.

Although the study design focused on rich description anchored by strong theory, there are several limitations on interpretation.. First, the Challenge performance study did not include a comparison group of young people who attended moderate- and low-quality offerings, so our assumption that SEL skill growth is the result of participation in the offerings is only an assumption and not the product of a strong research design. Rigorous answers to questions about how much growth occurs as a result of participation in the offerings—questions about impact—await a next study.

Second, although our theory of SEL skill measurement focuses on the mental contents that mediate between context and behavior, our measures were only poorly mapped onto these mental contents. Specifically, we included a set of measures for attachment-related anxieties, but other fast-moving emotion-laden scripts and schemas could be at work blocking or enhancing the relationship between context and behavior. Similarly, most of the SEL belief measures contain a mixture of objects of measurement - youth beliefs about their own behavior, their mental states, and their opinions about causal processes connecting states of the world and their own behaviors and beliefs. We treated these measures

as efficacy indicators, reflecting beliefs about skill competence, but they do not represent a systematic approach to measurement of mental contents in a way that differentiates between individuals on objects of measurement which are defined and agreed upon in advance.

Third, the Challenge study is limited by small sample size. In both the qualitative and quantitative sides of the study, rich description of exemplars was the focal method. For the performance studies in particular, small sample sizes reduced statistical power and increased the likelihood of type-2 error. For these reasons, we were unable to sufficiently investigate critically important issues such as reliability and validity of measures, relations among variables, and change across time.

Finally, and we believe most importantly, the lack of a consensual framework regarding how context, mental contents, and behavior interact in regard to development of SEL skills makes it difficult to map our pattern of results onto other's results. We believe that our results are in accord with others in the literature but the level of effort necessary to produce a cumulative statement is well beyond the resources available for the current work.

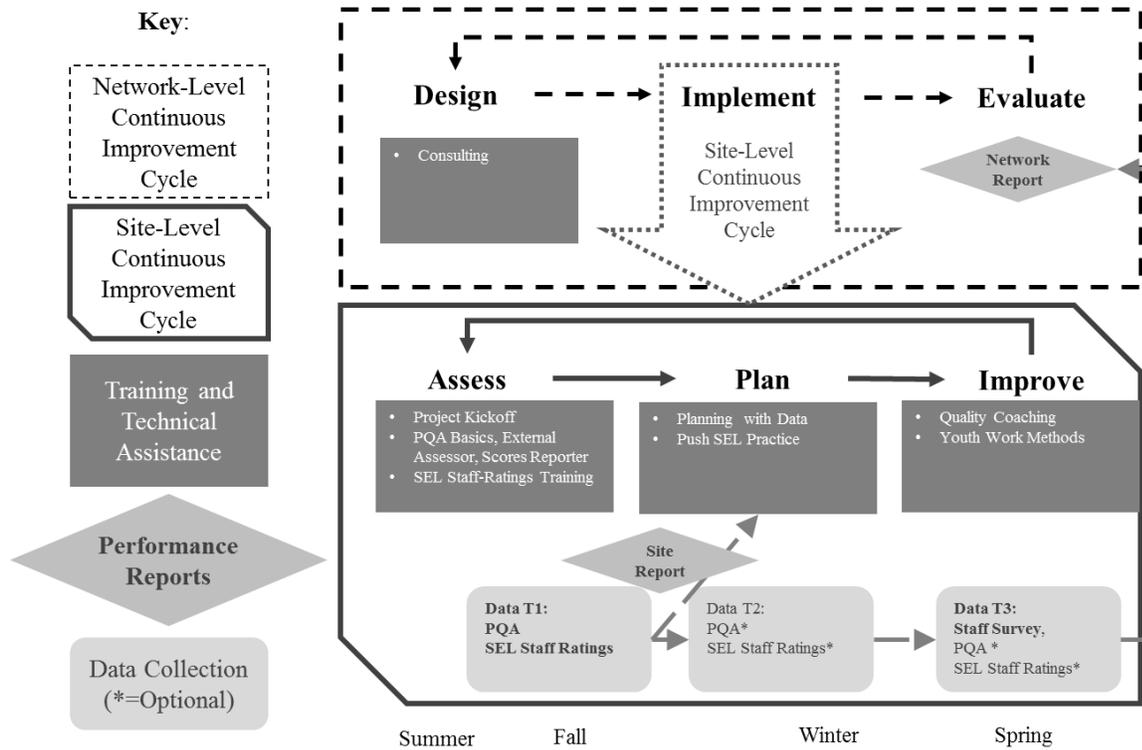


Figure A-1. Logic Model for the Continuous Improvement Intervention

