PREPARING YOUTH TO THRIVE

Methodology and Findings from the Social and Emotional Learning Challenge

Charles Smith, Gina McGovern, Stephen C. Peck, Reed Larson, Barbara Hillaker, Leanne Roy

Acknowledgements

The SEL Challenge was made possible by the financial support of the Susan Crown Exchange (SCE) www.scefdn.org. We thank the staff at Susan Crown Exchange (SCE)—Susan Crown, Haviland Rummel, and Hannah Baptiste—for invaluable vision, input, and facilitation at each stage of the study.

The authors also wish to thank and acknowledge the expert practitioner teams at the eight visionary organizations listed below. In addition to being the expert informants in developing the standards, they also helped us design and implement a suite of SEL performance measures and an improvement cycle at each of their sites.

AHA! (Attitude, Harmony, Achievement)

Boys & Girls Clubs of Greater Milwaukee

The Possibility Project

Philadelphia Wooden Boat Factory

Voyageur Outward Bound School

Wyman

Youth on Board

YWCA Boston

The SEL Challenge was conducted as a partnership between the eight exemplary agencies listed above, SCE, the David P. Weikart Center for Youth Program Quality at the Forum for Youth Investment, and additional technical consultants. Charles Smith was the principal investigator for the Challenge. Gina McGovern was the Challenge project manager. Reed Larson (University of Illinois) was the primary consultant for qualitative methodology and analyses. Stephen C. Peck (University of Michigan) is the Weikart Center's technical consultant for measurement and psychological theory. Barb Hillaker and Leanne Roy supported the research teams.

Suggested citation: Smith, C., McGovern, G., Peck., S.C., Larson, R., Hillaker, B., Roy, L. (2016). *Preparing youth to thrive: methodology and findings from the social and emotional learning challenge*. Forum for Youth Investment, Washington, D.C.

All inquiries regarding the contents of this report should be directed to Charles Smith at charles@cypq.org.

Table of Contents

List of Tables

List of Figures

Glossary of Key Terms

Agency – "To be an agent is to influence intentionally one's functioning and life circumstances" (Bandura, 2006, p. 164). We distinguish between two forms of agency. *Type 1 agency* refers to actions that result from the automatic activation of previously formed beliefs (e.g., goals and plans). *Type 2 agency* refers to the conscious process of formulating new beliefs (aka, *self-reflection*).

Benchmark – Benchmarks describe the type and level of performance on a performance measure used in the performance studies conducted at each program site as part of the Challenge.

Context – The term context is used to refer to the physical, social, and informational features of the environments in which youth live and develop, including OST settings. Perhaps most importantly in this work, context refers to places where adults and youth come together to carry out the work of the SEL offering and includes other persons, adults, and youth. Descriptions of context are frequently referred to as context, environment, situation, setting, and many other terms.

Co-regulation – Adult leaders provide the "nurturing, instruction, coaching, and support that will promote optimal self-regulation by the child, while simultaneously buffering against environmental stressors that might diminish regulatory capacity." (Murray et al, 2015, p. 19)

Curriculum – Description of (a) a sequence of content and planned experiences fit to the developmental and learning needs of the learner, and (b) the supports necessary for the instructional staff to plan and implement that sequence.

Offering – An offering is characterized by the same group of youth and adults meeting over multiple sessions for a planned learning purpose. The target offerings in the SEL Challenge are those exemplary offerings using SEL practices and curriculum to grow youth social and emotional skills.

Out-of-school time (OST) – 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.

Performance study – Use of selected performance measures at an individual OST site to describe the quality of management and instructional practices as well as change in youth skills to produce a performance report used during a continuous improvement cycle. The performance report includes comparisons to normative performance benchmarks.

Practice indicator – The lowest level descriptor for an SEL standard in this guide. Practice indicators describe specific youth experiences, staff behaviors, or other objective conditions that occur during out-of-school time offerings.

Promising practice – Staff behaviors and program structures for social and emotional skill learning that are both theoretically defined and have supporting evidence of effectiveness from expert practitioners. See Endnote i.

Quality improvement system (QIS) – In the OST field, a QIS typically consists of four elements: Standards for good performance, performance measures and reports, an annual continuous improvement cycle, supports and incentives necessary to implement the prior elements.

Responsive practices – These practices—modeling, scaffolding, coaching, facilitating—are delivered in the moment as youth experience the project curriculum, in particular during moments of challenge and represent specific methods of co-regulation.

Skill – "The capacity to act in an organized way in a specific context" (Fischer and Bidell, 2006, p. 321), where *capacity* refers specifically to the mental contents and processes (e.g., plans and strategies) that structure mental engagement with, and influence behavior enacted in relation to, the context.

SEL practice – Our definition of staff practices includes both staff behaviors (e.g., modeling appropriate use of emotion) and program structures that the staff put into place (e.g., curriculum, recruitment policy). Key youth experiences (e.g., taking on roles and obligations) point to staff practices and program structures necessary to produce those youth experiences. In this sense, the staff practices and key youth experiences are both included as standards for SEL practices.

Social and emotional learning (SEL) – SEL is defined in the following way by the Collaborative for Academic, Social, and Emotional Learning (CASEL): the process through which children and adults acquire and effectively apply the knowledge, attitudes, and skills necessary to understand and manage emotions, set and achieve positive goals, feel and show empathy for others, establish and maintain positive relationships, and make responsible decisions.

Standard – We use the phrase "standards for SEL practice" and the term "standards" to describe practices that (a) appear across the Challenge offerings, (b) were described as important by the expert practitioners, and (c) were supported in the evidence base. In this guide, each standard consists of a sentence describing the practice and multiple practice indicators describing different facets of the standard.

Executive Summary

The SEL Challenge was undertaken in pursuit of two ambitious goals: To identify promising practices for building SEL skills with vulnerable adolescents, and to develop technical supports for use of these SEL practices at scale in thousands of out-of-school time (OST) settings. The study design included a qualitative methodology, expert practitioners, and performance studies at each of eight exemplary programs. The products of the Challenge—standards for SEL practice and the suite of SEL performance measures—is designed to help OST programs focus deeply on SEL practice, assess their strengths, and improve the quality and effectiveness of their services using a continuous improvement approach.

By focusing systematically at a granular level of adult and youth behavior, the Challenge content supports use in practice-oriented settings and systems—youth programs, school day classrooms, mentorships, residential treatment, apprenticeships, workplace, families—where the qualities of adult-youth interaction and learning are a primary concern. 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.

The promising practices are featured in a Field Guide, *Preparing Youth to Thrive: Promising Practices for Social and Emotional Learning* (Smith, McGovern, et al., 2016), a companion website, and a suite of tools and technical assistance (SELpractices.org). This report, *Preparing Youth to Thrive*: *Methodology and Findings from the SEL Challenge*, describes how the partnership carried out the work of the Challenge and what we learned as a result. Findings from the SEL Challenge include:

1) The Challenge methodology successfully identified exemplary SEL offerings and produced *34 standards, 78 practice indicators, and 327 vignettes* for building SEL skills with vulnerable youth.
2) The *suite of performance measures* developed for the Challenge is feasible to implement and demonstrates sufficient reliability and validity for both continuous improvement and evaluation uses.
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: (a) Diverse staff and youth, intensive participation, and expert adult guidance; (b) Collaborative organizational cultures; (c) Exceptionally high quality instruction and youth engagement; (d) A consistent pattern of positive SEL skill growth across measures, offerings, and risk status.
4) The offerings shared an OST-SEL intervention design: *project-based learning with intensive co-regulation.*

The Discussion section addresses generalizability of findings, cautions about SEL measurement, and study limitations.

Introduction to the SEL Challenge

The SEL Challenge is a partnership among the Susan Crown Exchange (SCE), staff teams from eight exemplary out-of-school time (OST)[1] programs, the David P. Weikart Center for Youth Program Quality (CYPQ), and technical consultants. The partnership was created for two purposes: (a) identify promising practices[i] for building social and emotional learning (SEL) skills with vulnerable adolescents, and (b) develop technical supports for use of these SEL practices at scale in several thousand OST settings. The promising practices are featured in an SEL Field Guide, *Preparing Youth to Thrive: Promising Practices for Social and Emotional Learning* (Smith, McGovern, Larson, Hillaker, & Peck, 2016), companion website, and a suite of tools and technical assistance (SELpractices.org). This Technical Report, *Preparing Youth to Thrive*: *Methodology and Findings from the SEL Challenge*, describes how the partnership completed the work of the Challenge and what we learned as a result.

Although there are many ways to define and discuss the importance of SEL skills for vulnerable adolescents, a great deal can be summarized using the terms *self-regulation* and *agency*. In general, an adolescent's ability to self-regulate—to manage emotions, attention, motivation, and behavior to achieve specific purposes—is related to a wide range of positive outcomes. Perhaps more importantly, the ability to intentionally shift focus away from environmental cues that cause reactive or negative emotional responses, or to choose to be in environments already free from these cues, is a powerful step toward transcendence of contexts that limit potential. When adolescents use self-regulatory powers to ignore distractions, inhibit impulses, or choose environments that have higher developmental potential, they are often referred to as having agency.[ii] SEL skills are action skills for navigating and negotiating complicated real-world situations.

Although we have much to learn about social and emotional skills, we also have a great deal of evidence. For example, meta-analyses[iii]—studies that summarize across findings from many prior studies—have suggested that SEL skill-building curricula delivered in both OST and school settings have substantively important impact on a wide range of skills and outcomes (Durlak & Weissberg, 2010; Durlak et al., 2011). Further, the wider literature on SEL suggests that SEL skills transfer across settings and improve skill learning in other content areas (Durlak, 2015). In particular, this literature is consistent with the idea that SEL skills are also learning skills as both SEL interventions and SEL skills are

[1] 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.

associated with successful outcomes in settings where learning academic and other content is the central purpose.[iv]

There is growing consensus about the many positive effects of SEL, but access to settings that build SEL skills are not equally available to all youth (Putnam, 2015). Because these skills are critical for healthy functioning across the life course, lack of access to environments that build these skills constitutes a developmental risk factor (Bailey, Duncan, Odgers, & Yu, 2015; Cunha & Heckman, 2006). Many American youth are vulnerable in ways similar to the Challenge participants: relentlessly exposed to a popular culture of violence and aggression; experiencing social exclusion, poverty, and instability in their neighborhoods and (sometimes) households; attending substandard and stressful schools; and exposed to environmental contaminants (Murphey et al., 2014). Chronic exposure to these stressful and traumatic experiences can produce negative effects across the life course (Blair & Raver, 2012; Evans & Fuller-Rowell, 2013; Jaffee & Christian, 2014). These are precisely the young people who most need settings designed to foster SEL skills - and for whom they are often least available. The SEL Challenge was designed to help address these unmet needs.

In the SEL Challenge, we focused on descriptions of practices used by professional staff and performance benchmarks demonstrated by exemplary SEL organizations and offerings. These critical aspects of implementation are often not adequately described in the aforementioned SEL impact literature (Durlak & DuPre, 2008; Durlak & Weissberg, 2007, 2010; Durlak et al., 2011). By focusing on granular descriptions of adult behavior and youth experience at the point-of-service level, the Challenge content supports point-of-service level application in OST programs, regular school-day classrooms, mentorships, residential treatment, apprenticeships, workplace, families, and other contexts where the quality of adult-youth interaction is a primary concern. We hope that local policy makers and funders will use the SEL Challenge as a template for identifying local networks of expert practitioners and their exemplary programs, forms of social capital already available in many communities, and make sure that they are adequately recognized, resourced, and replicated.

Science of SEL Practice

In the Challenge, and in this technical report, we address the complicated issue of how settings can be organized for the purpose of youth's SEL skill learning. Recent treatments of this subject matter cover the broad developmental arguments and evidence linking participation in OST contexts to youth skills and outcomes (Collaborative for Academic Social and Emotional Learning, 2015; Corcoran & Slavin, 2016; Farrington et al., 2012; Jones & Bouffard, 2012; Li & Julian, 2012; Nagaoka, Farrington, Ehrlich, & Heath, 2015; Yoder, 2013). Herein, we focus more narrowly on proximal descriptions of youth SEL skills that are demonstrated in the OST context and staff practices and curriculum features

that help to structure the OST context for skill learning. We then use those descriptions of practice and skills to guide selection of performance measures for use in a continuous improvement cycle in organizations trying to improve the quality and effectiveness of their SEL practices.

The barriers to attaining these objectives—taking SEL practice to greater scale—are high because much of the available information about SEL practice is difficult to use. There are many ways to talk about SEL practices and skills, making it difficult to know if we are focused on the same thing when we use the same words. This is particularly true of the language of skill-building and the many definitions for words that describe how people and contexts come together to produce individual skills (e.g., situations, transactions, ecologies).[v] Further, intervention designs are often privatized in specific curricula and tools that do not support translation and adaptation. Finally, much of the scientific description of SEL is focused on individual youth skills and outcomes rather than the SEL practices necessary to initiate and scaffold skill growth.[vi] In short, we lack a cumulative base of scientifically-organized expertise about SEL practice.

The lack of a cumulative base of expertise about SEL practice has hindered the development of useful theories about practice which, in turn, has hindered development of useful tools for practice (e.g., standards, curricula, logic models, clinical checklists, performance measures) that can be applied by both expert and novice practitioners over wide variations in local context.[vii] More useful theories about practice can be used by professionals to align generic SEL practices with the unique profiles of youth experiences and community resources in which their offerings "make sense" with and for youth. With an adequate theory of practice, performance measures can also be aligned to help leaders understand how well the SEL practices are being delivered and how the youth are responding in terms of engagement and skill growth. One of the primary aims of the Challenge Study was to assemble suitable theory—theory about how youth develop skill, how specific aspects of context (i.e., standards for SEL practice) support skill development, and how organizational systems promote use of SEL practices in the many settings they are accountable for—so that more powerful tools could be designed and applied in the work of creating more effective SEL offerings.

SEL Skill, SEL Practice, and the Continuous Improvement Intervention

SEL Skill. In our use, *youth skill* refers to mental content and processes that (a) structure mental engagement with the context and (b) influence behavior enacted in relation to context. These basic elements—context, mental processes, and behavior—are reflected in a definition from Fischer and Bidell (2006) that "skill is the capacity to act in an organized way in a specific context" (p. 321), where the terms *capacity* and *organized* imply the centrality of mental content and processes. Mental skills are also often described in behavioral terms, as behavioral indicators of the application of mental skills to

the context. The final definitions for the six domains of youth SEL skill, which are addressed in the Challenge, describe both *mental skills* (e.g., ability, disposition) and *behavioral indicators* of mental skills enacted in relation to the OST context (e.g., take action, persevere):

- *Emotion Management* – Abilities to be aware of and constructively handle both positive and challenging emotions
- *Empathy* – Relating to others with acceptance, understanding, and sensitivity to their diverse perspectives and experiences
- *Teamwork* – Abilities to collaborate and coordinate action with others
- *Responsibility* – Dispositions and abilities to reliably meet commitments and fulfill obligations of challenging roles
- *Initiative* – Capacities to take action, sustain motivation, and persevere through challenge toward an identified goal
- *Problem Solving* – Abilities to plan, strategize, and implement complex tasks

Youth are active builders of their own skills, and the language of the standards includes many descriptions of active mental engagement, such as practice, identify and name, explore, own, articulate, develop, share, work together, manage, build knowledge, plan, brainstorm, think strategically, etc. The implications of this language, and the literature on skill development,[viii] is that the individual youth is the ultimate and only skill-builder and that self-regulation occurs when youth self-organize their mental skill (Fischer & Bidell, 2006).

As described in Chapters Two and Three of this report, we assemble theory that emphasizes youth's experience of self-regulation, which we refer to as *agency* and which we see as a primary purpose for building contexts that focus on SEL. Consistent with more and less volitional modes of self-regulation, the term *agency* can be thought of in two senses (see Endnote i): First, youth can experience increased agency because they enter a context that elicits positive emotion and offers opportunities to use existing skills (e.g., playing to youth's strengths and interests). Agency in this sense is more automatic, unconscious, and fast moving. In the second sense, youth experience increased agency as they train the conscious focus of their attention and awareness on specific aspects of (a) the context, (b) their own mental engagement with the context, (c) their own behavior in relation to the context, and (d) the wide array of meanings that accompany each of these parts of youth experience in an OST setting.

Young people who have been exposed to traumatic or chronically-stressful experiences may require more intensive supports for successful skill building in either the more passive or more active senses of agency. In particular, mental processes related to emotion may block or enhance (i.e., mediate)

skill learning, as cues in the context trigger emotion-laden responses that may deactivate mental skills that could be engaged by the context, or disrupt the process of focusing attention on what the context has to offer. These concerns with regulation skills, and the corresponding experiences of agency that result, are particularly germane to OST contexts that are intentionally designed to help youth feel safe and interested so that attention can focus on the task at hand, motivation can emerge around task success, and skill learning and mastery can occur through repeated practice (Smith, Hillaker, & McGovern, 2014).

SEL Practice. We identified the critical components of context by asking experts who have been designing SEL contexts for many different youth over many years. These experts have developed practices that both focus on an identifiable set of SEL skills and help manage necessary individuation as staff respond to youth's different experiential histories in each program cycle. Our definition of SEL practices is broad and includes descriptions of (a) staff behaviors that occur in the moment (e.g., modeling appropriate use of emotion), (b) program structures that the staff put into place and that endure through time (e.g., curriculum), and (c) youth key experiences (e.g., taking on roles and obligations) that point to the developmental experiences of youth, the mental models that staff have for youth experience, and the practices and structures that staff use to reliably initiate and scaffold those types of experiences.

Although our choice to include *youth key experiences* in a set of standards for SEL practice adds additional complexity (because youth's mental content and processes are different from context), the descriptions of youth experiences from the perspectives of the staff include many cues about the practices implemented by the staff to initiate and scaffold SEL skill learning. This basic insight was reflected in many of the standards: Staff have to know their youth and their youth's communities so that SEL practices can be organized in such a way as to engage youth's store of prior experience.

We use the broader term *context* to refer to the physical, social, and informational features of an OST setting that are proximal to an individual youth. Perhaps most importantly, in this work, context refers to places where adults and youth come together and where SEL practices are implemented. The participating adults and other youth are also part of the context for a given youth. Descriptions of context are frequently referred to as setting, environment, situation, ecology, treatment, and many other terms.

Continuous improvement intervention. Because OST settings are uniquely positioned to build SEL skills with vulnerable youth, 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 cycles. If the opportunity to have a desired skill-building experience is dependent upon the power of the context to activate prior knowledge and skills (and deactivate mental contents that may block engagement), then the exact features of the context that optimize skill growth almost certainly vary within each group of individual youth and between groups of youth from different communities and with different histories. The continuous improvement approach

supports expert practitioners to review a repertoire of SEL practices in order to select practices that are most likely to activate previously-developed skills in the youth they serve so that skills can be extended, practiced, and eventually mastered. Continuous improvement intervention directly develops staff expertise to design curricula to reliably initiate and scaffold skill development for youth with different experiential histories and to allow the work to take novel turns that keep the youth highly engaged. Appendix A provides additional detail on the continuous improvement intervention elements, validation evidence, and scale of use.

Overview and Use of the Report

This report is organized in four sections, reflecting the four primary aspects of the work: (a) selection of expert practitioners and exemplary offerings, (b) developing the standards for SEL practice, (c) describing a suite of SEL performance measures, and (d) identifying performance benchmarks for exemplary SEL organizations and offerings. A final section discusses findings from the Challenge and how to take SEL practices to greater scale using the standards and performance measures.

This report has several purposes. First, the report presents detailed methods and findings for each of the four parts of the work so that readers can, for example, understand how we selected expert practitioners or how we decided which practices were important enough to call standards. Second, the report provides validity evidence for the standards for SEL practice. The exemplary offerings were of exceptionally high quality, and this lends credibility to the standards for SEL practice that were developed by studying them. Finally, the report assembles information about reliability, validity, and feasibility for a suite of SEL performance measures, including performance benchmarks for normative use in deciding what it takes to promote SEL skill growth.

For efficient use of this report and the other SEL Challenge materials, we provide the following guidance to users:

- For readers interested in the overall findings from the SEL Challenge, they are summarized in Chapter Five.
- For readers interested in locating expertise in their own communities, Chapter One describes the process we used to locate expert practitioners and summarizes the characteristics of organizations, offerings, staff, and youth who participated in the Challenge.
- For readers interested in the standards for SEL practice, the SEL Field Guide, *Preparing Youth to Thrive: Promising Practices for Social and Emotional Learning*, provides a detailed description of the standards, with supporting content. The qualitative methodology used to derive the standards and practice indicators is described in Chapter Two and Appendix B.

- For readers interested in the suite of SEL performance measures, Chapter Three provides a detailed description of the measures, procedures, and staff time necessary to replicate the Challenge data collection. Chapter Four presents performance results for the SEL Challenge organizations and offerings. Appendix E provides detail for reliability and validity of SEL skill measures. An aggregate performance report for all participating Challenge organizations at baseline and end-of-cycle is available at http://cypq.org/SELChallenge along with complete codebooks, permission forms, and other supporting documents for all measures used in the Challenge.

Chapter One. Study Design and Participants

The SEL Challenge study was designed to identify promising practices and support replication of those practices at greater scale.

Study Design

The study's design[ix] focus on describing SEL practices required that we select participants who brought expertise about practice based on years of trial-and-error experience and training in various disciplines and clinical approaches. We also needed those experts to bring an exemplary SEL offering that we could use as an object of investigation in both the qualitative and quantitative aspects of the study. We looked for expertise in OST programs because this field offers rich examples of experts who manage to find the resources and the freedom to mount the program designs they believe will build social and emotional skills and who have managed to continue refining and improving their curricula and learning designs over time.

Figure 1 describes the SEL Challenge study design as it was implemented over approximately 18 months. The study design consisted of several parts. An intensive application process identified expert practitioners and exemplary SEL offerings. The expert practitioners were recruited in teams of three: an organizational leader, an experienced instructional staff member, and an evaluator. We defined an OST program offering as including the same group of youth and adults meeting over multiple sessions for a planned learning purpose.

Three convenings of this learning community were designed to build relationships among participants, work on a common language for SEL practice, get feedback on the standards and skill measures, implement a performance study at each site, and review and synthesize interim findings.

To create the standards for SEL practice, we began with an existing evidence base about SEL practice, implemented qualitative data collection and methods with site visits and participant feedback loops, and produced a detailed case study for each exemplary offering.

We also conducted performance studies for each exemplary offering. This entailed development and implementation of a suite of SEL performance measures and, based on the performance data, development of several performance benchmarks for organizations doing SEL work.

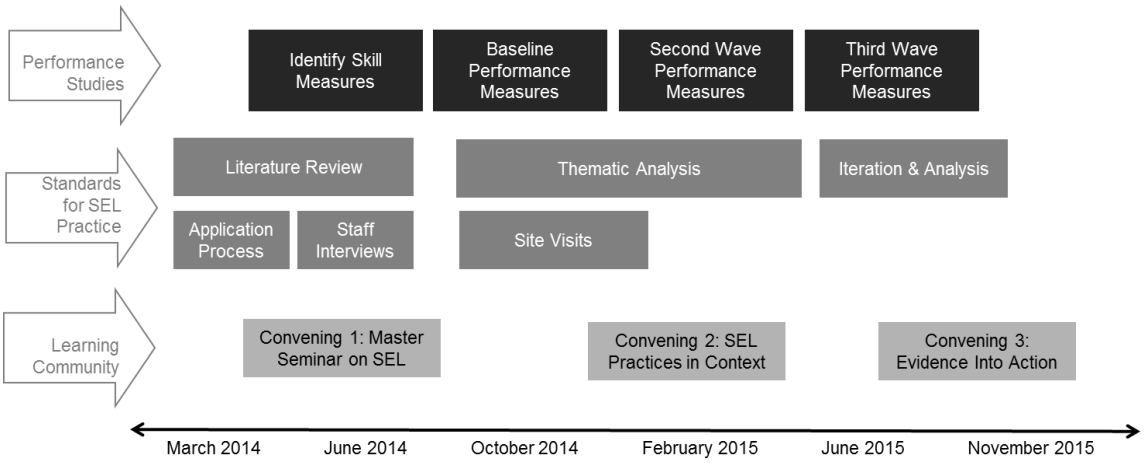

Figure 1. **SEL Challenge Implementation over 18 Months**

Expert Practitioners and Exemplary Offerings

The study design was focused on promising practices, so we recruited expert staff and eight exemplary offerings focused on SEL skills. The request for proposals resulted in 242 letters of intent. The final eight offerings were selected based on the following criteria: Focus on a targeted domain of social and emotional learning; clearly articulate a mature and/or innovative offering design and describe how youth learn in relation to the design; represent different offering types (e.g., community service, arts, apprenticeship); can provide evaluation evidence for the offering; serve vulnerable youth at sufficient intensity. In particular, it was the ability of the eight participating organizations to articulate the details for their theory of practice—in disciplinary, clinical, or other professional language—that set them apart from the other applicants. Appendix B contains the full set of questions used for the Letter of Intent and Application process.

Characteristics of Organizations, Offerings, Staff, and Youth

Tables 1-3 describe characteristics of the organizations, exemplary SEL offerings, staff, and youth who participated in the SEL Challenge. Table 1 shows that organizations in the Challenge ranged widely in revenue and capacity, operating in the cities of Santa Barbara, Milwaukee, Philadelphia, New York, St. Paul, St. Louis, and Boston. Boys & Girls Clubs of Greater Milwaukee (BGCGM) operated with the largest annual organizational budget, $24.8 million, whereas the Philadelphia Wooden Boat Factory (PWBF) operated with the smallest annual budget of $525,000.

Each organization contributed one exemplary offering to participate in the Challenge. The offerings ranged in size from 11 to 45 youth and two to five staff. The average ratio of adults-to-youth was one to six, and all of the offerings provided at least two fully engaged staff at all times. Dosage or

amount of participation required by each offering was intensive in all cases and varied from 39 contact hours to 370. Seven of the offerings had a multi-year evaluation history, and four of the offerings—Wyman's Teen Outreach Program (Allen, Philliber, Herrling, & Kuperminc, 1997), Voyageur Outward Bound School (VOBS) (Neil, 2003), The Possibility Project (TPP), and InIt at YWCA Boston (YW Boston)—were nationally disseminated, evidence-based program models.[x]

Each of the program offerings has project curricula to build skills in a wide range of skill sets, including community service, community organizing, theatre, boat building, anti-racism training, relationship wisdom, and community arts. These are highlighted in Table 2. We defined curriculum as a sequence of content and planned experiences fit to the developmental and learning needs of the learner and the supports necessary for the instructional staff to plan and implement that sequence. The standards identify two parallel curriculum sequences in each offering: (a) A *project curriculum* sequence of project-related content and skill building activities (e.g., wood working skills necessary to build a boat); and (b) an *SEL curriculum* indicating (1) where in the project curriculum youth are likely to experience challenge, stress, or strain that elicit responsive practices from staff, and (2) a sequence of regular check-ins that are known in advance.

Demographic and descriptive characteristics of staff and youth in the target offerings are included in Table 3. Almost all staff had completed a college degree, five of the eight organizations had staff with advanced degrees, and staff had extensive experience in the youth development field. Average tenure within the program ranged from one year to nearly eight years. More than half of the organizations have had minimal turnover in staff for the offering in recent years. Program managers and lead instructors labeled themselves as having relatively high rates of expertise, but not necessarily in SEL where only 50% of program managers and 11% of lead instructors rated themselves as experts.

Staff was ethnically diverse, and this mirrored the diversity of the youth in most cases. In the two programs that featured predominantly African American youth, nearly all staff was African American. One of the two programs with a higher percentage of Hispanic youth also had staff that was Hispanic.

Youth participation in the offerings was voluntary. Youth ranged in age from 12 to 19, and the average age was 15. Staff defined their programs as serving vulnerable youth, and they defined vulnerability in consistent ways: The youth presented low social and emotional skills during recruitment (e.g., introversion or few friends); lived in homes or neighborhoods where exposure to violence and toxic levels of stress were almost assured; were referred by a social service agency due to a history (e.g., foster care, juvenile offense) that was likely to include traumatic experience; and were exposed to systematic racism and exclusion.

We also asked the youth several questions about a few common risk indicators. For example, 29% of youth lived in a household where the highest educated adult did not go to college, 7% said they

received mostly Cs and Ds in school, and 3% were not currently in school. For a measure of risk related to attachments and relationships, about 15% of the youth were indicated. Together, these risk indicators represented 35% of the participating youth.

Table 1. Organization Characteristics

	Organization	*Location*	*Annual Operating Budget*	*Evidence-based model*
AHA! ATTITUDE • HARMONY • ACHIEVEMENT	Attitude, Harmony, Achievement! (AHA!)	Santa Barbara, CA	$1M	N
BOYS & GIRLS CLUBS OF GREATER MILWAUKEE	Boys & Girls Clubs Of Greater Milwaukee (BGCGM)	Milwaukee, WI	$24.8M	N
PHILADELPHIA WOODEN BOAT FACTORY	Philadelphia Wooden Boat Factory (PWBF)	Philadelphia, PA	$525,000	N
THE POSSIBILITY PROJECT EMPOWERING TEENAGERS TO CREATE A BETTER WORLD	The Possibility Project (TPP)	New York, NY	$700,000	Y
VOYAGEUR OUTWARD BOUND SCHOOL	Voyageur Outward Bound School (VOBS)	St. Paul, MN	$3.1M	Y
WYMAN Real Teens. Real Life. Real Results.	Wyman	St. Louis, MO	$5.9M	Y
YOUTH ON BOARD	Youth On Board (YOB)	Boston, MA	$374,000	N
yw boston	YWCA Boston (YW Boston)	Boston, MA	$2.2M	Y

Sources: Letters of Intent, Applications, Staff Survey.

Table 2. Target SEL Offering Content Descriptions

Organization	*SEL Offering*	*Content*	*Description of Target SEL Offering*	*Total contact hours*
AHA!	Girls' Relationship Wisdom Group	Group Self-Improvement	Empowers teens to create peaceful and connected communities.	39
BGCGM	Can You Hear Us Now?	Poetry and Community Service	Youth develop an artistic expression of an issue important to their lives and make a presentation to their community.	41
PWBF	Boat Build and Sail Club	Boat Building Apprenticeship	Teams of youth apprentices build boats in the shop during colder months and sail the completed projects during the summer.	351
TPP	Afterschool Program	Theater and Community Service	Youth work together to write, produce, and perform an original musical based on their lives and their ideas for change.	300
VOBS	Strive Forward	Outdoor Adventure	Young men learn wilderness and leadership skills through outdoor expeditions.	370
Wyman	Teen Outreach Program ®	Life Skills and Service Learning	TOP participants learn leadership skills and plan and deliver a number of community service learning experiences.	66
YOB	Boston Student Advisory Council Working Group	Organize youth policy in Boston	High school leaders from across the city work together to define issues that are most relevant to them and develop collective solutions.	134
YW Boston	Youth Leadership Initiative (Init)	Youth Organizing	Youth build their abilities to work across differences and ultimately engage other teens in a community action project that addresses inequity.	152

Sources: Letters of Intent, Applications, Staff Survey.

Table 3. Staff and Youth Characteristics for the Target Offering

		Education		***Experience***			
	Sample size	*College degree*	*Advanced degree*	*Years in youth development*	*Years in this program*	*Program content (% rating themselves "expert")*	*SEL content (% rating themselves "expert")*
Program Manager	(n=8)	37.5%	25%	17.88	9.08	62.5%	50%
Lead Instructor	(n=9)	66.7%	22.2%	11.67	6.44	44.4%	11.1%

Source: Staff Survey

	Ethnicity					
	Sample size	*African American/Black*	*Hispanic/ Latino (a)*	*Asian/Pacific Islander*	*White*	*Prefer not to disclose*
Staff	(n=23)	28%	12%	8%	40%	12%
Youth	(n=112)	48%	30%	8%	14%	4%

Source: Staff Survey Time 3; Youth Day-of-Observation Survey Time 1

Chapter Two. Standards for SEL Practice

Our goal in developing standards for SEL practice was to integrate descriptions of the multiple elements of SEL practice in plain language to help professionals talk to each other, and to youth, about the "how to" of SEL work. The focus was on description of the action and experiences that staff and youth see and understand to be happening in highly effective OST settings. Although the exemplary offerings differed amongst each other in the curricula they used, the way they went about initiating and scaffolding SEL skill learning had important similarities that are described by the standards we identified. In this section, we discuss theory, data, analyses, and results from the qualitative method used to produce the standards. The full presentation of the standards and supplemental content are provided in the guide *Preparing Youth to Thrive: Promising Practices for Social and Emotional Learning* (Smith, McGovern, et al., 2016).

It is important to note that our overall design concept was to use a small sample of intentionally selected offerings in order to identify practices that were common across the settings at a granular level. We were in search of best practices used in exemplary offerings and followed the general rule that prevalence of a practice across the eight exemplary offerings was an argument for its importance. We conducted the performance studies described in Chapter Four in part to validate our selection process, and these findings indeed suggest that the selected offerings were operating at an exceptionally high level across a suite of SEL performance measures. These were ideal offerings to study as exemplars of SEL practice.

Qualitative Method

The method used to produce the standards and supplemental content was conducted in four overlapping and iterative phases. In the first phase, we reviewed literature on SEL skill development and practice and produced a preliminary conceptual framework and set of "starter codes" for different elements of SEL practice. We also developed a theory that linked SEL skill learning to SEL practices so we could ask practitioners specifically about the learning cycles their SEL offerings provide. Second, we designed data collection instruments (based on the framework) and methods; and then collected the narrative and artifact data in several waves. In the third phase, we subjected the narrative data to an iterative qualitative method to evaluate and revise the primary codes and subcodes for specific SEL practices in within and across the six SEL domains. These codes eventually became the standards and curriculum features.[xi] In the fourth phase, supplemental content was assembled to support interpretation of the standards: Multiple vignettes from the expert practitioners were selected from the narrative data to

deepen practical illustration of each standard, and case studies describing curriculum and resources in each of the exemplary offerings were developed to provide additional context.

In the remainder of this section, we focus on the details for the qualitative method used to produce the SEL standards and curriculum features presented in Tables 4 and 5.

Literature Review and Theory for SEL Skill Learning

An initial literature review was conducted and produced a preliminary set of codes for SEL practices in each of the six domains. In particular, we reviewed the extensive Larson et al. literature based on hundreds of interviews with youth and instructional staff. The twenty articles reviewed represent a unique body of work using qualitative methodology to explicate youth experience in developmentally focused OST settings, with reference to how settings initiate and scaffold youth experience and which SEL skills youth learn through their experiences. This preliminary coding framework for youth key experiences, staff practices, and curriculum features is provided in Appendix B. The initial Larson et al. evidence base is identified with an asterisk in the References section.

Through the literature review process, we extracted preliminary codes describing key youth experiences, staff behaviors, and curriculum features within and across the six domains. These preliminary codes were used as "starter codes" to examine the first round of data. In an exploratory fashion, we read through the narrative data from letters of intent, applications, and artifacts to examine how well the codes fit the data and whether there were strong themes in the data that were missing from the codes. Based on these assessments, we made a first round of revisions to the language of the codes. In this first phase, a preliminary set of 31 primary codes and 72 subcodes was produced across the six domains. Several curriculum features were also identified through the literature review and exploratory coding.

Underlying theoretical work. We also reviewed theories that link context (SEL practice) to individual experiences of skill learning. We needed to ask questions of the study participants that would yield narrative data containing information relevant to description of SEL practices at the right level of granularity. Specifically, we needed our expert practitioners to respond to questions at a sufficient level of detail so that the lowest-level indicators of the standards could be illustrated with multiple vignettes of offering-specific variations on the same practice.

We started with the assumption that youth learn skills in each SEL domain much like other skills, through a cycle of experiential learning such as that represented in Figure 2. As youth encounter challenges embedded in the offering curricula, they go through a cycle of identification, planning, action, and evaluation that results in resolution, adaptation, or redirection in relation to the original challenge. As the approach to a novel challenge becomes routine through repetitive practice, mastery occurs, and the

skill becomes integrated as one part of a response to new, more complex challenges.[xii] Here we use the term *challenge* broadly to mean some form of cognitive or emotional dissonance, such as "Things don't add up," or "I can't learn this," or "I'm too stimulated to focus."

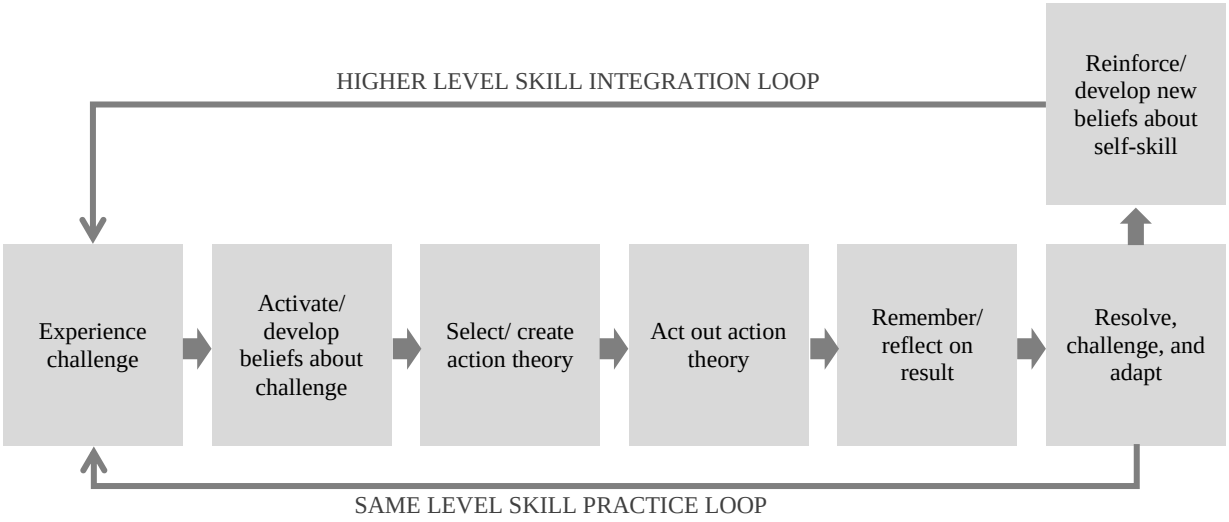

Figure 2. **Learning Cycle**

The implications of thinking about SEL learning in this way are, first, that the offering curricula are designed to intentionally produce the experience of a step in the learning cycle, and, second, that staff were professionally prepared to respond with supports during moments of challenge, to help move the youth forward. Because those moments were either a natural part of the activity or intentionally designed, we reasoned that these aspects of SEL practice were likely the "promising practices" that could be identified and replicated.

We worked with SEL Challenge participants, through interviews and in a workshop activity, to map the eight offering curricula in great detail. We repeatedly asked the expert practitioners about moments where the curriculum intentionally (or at least predictably) produced moments of challenge (e.g., requiring skills not yet learned, frustration with the process, interpersonal conflict, waning motivation), reflection, or demonstration of mastery. We wanted to know what was happening during those moments, what the youth were experiencing, and what the staff was doing in response to support those experiences. Because all youth experience challenge (first step of the learning cycle) differently, we asked staff about the types of practices they were prepared to deliver in advance, as the offering experience unfolded through time: Which kinds of practices – including the practice of restraint – did staff use in the moment when challenges emerge? How did staff get to know the youth and what their challenges are likely to be? Can staff anticipate youth's responses to challenges that may occur at different moments in the curriculum? What are the structures for check-ins that helped staff know when

youth are experiencing challenge? What does youth exposure to trauma mean for their experience of the curriculum and for staff response? When and how is a given practice applied?

Finally, we also reasoned that skill learning cycles vary in length for different SEL domains (this was clear from the Larson et al. evidence base). For example, in the emotion management domain, learning cycles might occur very rapidly (e.g., secondary appraisal of emotion may occur in seconds), whereas learning responsibility might occur over the entire nine-month arc of an offering. Our research design assumed that learning cycles for social and emotional skills occur on timescales relevant to the offering curriculum and staff practices that the curriculum was designed to deliver. That is, we are concerned with *moments* during the offering curriculum or a youth's struggles wherein a staff might provide a real-time response to youth behavior, or *sessions* or *sequences of sessions* during which, for example, specific parts of the curriculum are implemented and specific types of youth experiences occur (e.g., completed the trust building part of the curriculum). Extending further, the entire offering curriculum cycle may be the descriptive period over which skills grow (e.g., nine months), and in some cases over multiple years (e.g., adolescence) in the case of returning students who in some programs took on staff roles in subsequent years.

Data Collection and Coder Reliability

Several types of narrative and artifact data were collected. Appendix B provides questions from all data collection instruments: letter of intent, application, interviews, and focus groups. The letter of intent and written applications were submitted via online survey software. Next, two two-hour phone interviews were conducted with each organization leader and the lead instructional staff for the exemplary offering. Interviews were recorded and transcribed.

Each participating site also received a two-day visit that included an additional two-hour interview with the organization leader, an additional two-hour interview with lead instructional staff for the target offering, a one-hour interview with an expert youth who had participated in the program during a previous cycle, and a one-hour focus group with up to eight current youth participants. Interviews and focus groups were recorded and transcribed. Other data sources included copies of curriculum materials, evaluation reports, and detailed notes collected at each of three learning community meetings.

For the primary qualitative analyses employed to develop the SEL standards, a master narrative data file was created from the written applications and phone interviews by loading all responses into a data file using spreadsheet software. Each cell contained a text segment representing one response, or a portion of a response, and yielded a spreadsheet containing 2,261 text segments, averaging 452 characters per segment.

For the coding process, coders used a consensus method to improve reliability of coding data in the master narrative data file. Codes were defined and applied by multiple raters to ten segments of data at a time, percent perfect agreement was calculated, and a consensus discussion was held to examine differences. At three reliability checks conducted during the third phase, average percent perfect agreement was 78% at the primary code level and 68% at the subcode level before the consensus discussion.

Analytic Approach

In this section, we describe the qualitative analyses conducted using the *master narrative data file*. Our primary goal in these analyses was to identify standards for SEL practice in six SEL domains by using the master narrative data file to evaluate, revise, and refine codes at a relatively fine level of granularity.

Analysis methods of grounded theory and other qualitative approaches were employed. We used both thematic and content analysis techniques; that is, we used both theory and frequency of occurrence in the data as criteria for initially identifying practices that were important. We then used the following criteria for final inclusion of a staff behavior, a program structure, or a youth key experience as a standard or a practice indicator: (a) appears across the exemplary SEL offerings, (b) was described as having content and substantive validity by the expert practitioners, and (c) was supported by an evidence base[2] in the literature. By content and substantive validity, we mean that the practice was rated by the expert practitioners as being both important to their work (content validity) and frequently occurring (substantive validity) in their curriculum design.

The following two steps describe the iterative sequence of analyses in the third phase of the qualitative study:

1. *Coding all data into primary codes for the standards.* Using the preliminary framework of 31 primary codes and 72 subcodes, two raters independently coded text segments using codes from all six domains and came to consensus on definitions for the primary codes. All text segments in the master data file were then assigned to primary codes and were allowed to be assigned to multiple primary codes. All text segments were also identified by data source (e.g., interview, application, etc.) and question so that the data file could be queried by question and offering.

Subsequent iterative analyses then examined and revised these codes through a process of constant comparisons. We asked: Were all of the primary codes used frequently in all eight programs?

[2] It is worth pointing out that the invaluable Larson et al. evidence base for this work was almost exclusively qualitative, and it was useful in our formative work for precisely this reason.

Were text segments for each offering more frequently coded to the SEL domains targeted in that offering? Did emergent codes have support in the literature base? As a result of this process, several new codes were developed, and operational definitions for the evolving codes were revised. Results for the application of the primary codes in this step are presented in Appendix Table B-5.

2. *Coding data into subcodes for the practice indicators*. The next step in the qualitative analyses was focused on revising and applying subcodes (practice indicators) to represent more specific staff behaviors, program structures, and youth key experiences within their respective primary code (standard) and SEL domain. Our goals for this step in the analyses were to evaluate and improve the fit of the subcodes to the narrative data file while at the same time discovering new subcodes and making revisions. We asked: Could most of the text segments in each domain be assigned to a subcode? Were each of the subcodes supported by multiple text segments? Was the number of subcodes manageable in the sense that there were not too many (i.e., not a different set of subcodes for each offering)? Were the subcodes sufficiently granular to describe objective staff and youth behavior?

To code the text segments using the subcodes, we first used the master narrative data file to create within-domain data-sorts, including only text segments within a single domain. We then coded the text segments within each domain by any applicable subcode. The goal at this stage was to eliminate redundancies across skill domains and to identify codes without sufficient evidence from a broad cross-section of offerings. Results for the application of the subcodes in this step are presented in Appendix Table B-6.

In this final step, we also spent a lot of time clarifying language as the operational definitions of the primary codes and subcodes were finalized. Ultimately, all operational definitions for the codes and subcodes were reviewed by the entire SEL Challenge learning community, revised further, and presented as the final set of SEL standards and practices indicators.

Curriculum Features

Through the two steps of iterative analyses described above, 681 segments were coded into one or more primary codes (and subcodes) that were either (a) not a staff practice that occurred in the point-of-service setting of the offering or (b) were present in all of the domains. We referred to these cross-domain or organization-level features as curriculum features—aspects of the offering that describe the structure, sequence, and purpose of the target offering across multiple domains—or management practices that supported staff to implement the SEL practices named in the standards for all domains.

Vignettes and Case Narratives

Following the qualitative analyses and finalization of the language for the practice indicators within each standard, multiple text segments were selected as exemplary vignettes. These vignettes were selected to represent different disciplinary, clinical, or other professional languages used by the expert practitioners. We hoped that seeing the same practice described in different ways would support translation of the standards for application in many different kinds of contexts that make up the OST field.

To further support interpretation of the standards, practice indicators, and vignettes, case narratives for each offering were developed to add contextual depth regarding implementation of the SEL practices named in the standards. The case narratives are featured in Part Three of the SEL Field Guide, *Preparing Youth to Thrive: Promising Practices for Social and Emotional Learning* (Smith, McGovern, et al., 2016), and include a description of the history of each organization and offering, the background of staff and youth, and a detailed map of the curriculum sequence. Each case narrative was developed over several rounds of review and revision with the expert practitioners.

Results for Standards, Curriculum Features, OST-SEL Setting Intervention Design

Table 4 presents 34 primary standards and 60 practice indicators in six SEL domains. Each of the standards in Table 4 consists of a definition sentence and multiple practice indicators describing facets of the standard. It is possible to summarize across the practice indicators for some key characteristics of exemplary SEL settings. For example, 38% of the standards are responsive practices (modeling, scaffolding, coaching, facilitating) delivered by staff and explained further below; 22% of the practice indicators involve youth talking or communicating with others; 9% imply one-on-one communication between staff and youth; and 9% describe methods of staff shifting from leadership to support.

However, the expert practitioners also told us that most or all of the standards and practice indicators were both important and prevalent in their offerings. On a scale of one to five, where one was not important and five was very important, 60% of the practice indicators scored above four. In response to the question, "How likely are these practices and youth experiences to occur in your program?" all but four of the practice indicators were likely in 40% or more of the offering sessions. This feedback was a positive preliminary source of validation evidence for the standards.

Through implementation of the qualitative method described in this chapter, and from literature review to final analyses of data, it was clear that some important and prevalent staff practices were not appropriately coded into one of the six domains in the SEL standards. Some were management practices (e.g., youth recruitment policies), some were present in most of the domains (e.g., youth experience increasing agency), and some were articulated as the "non-negotiables" that were prerequisites for any effective implementation of any of the SEL standards (e.g., safe space). An additional 18 practice

indicators were developed to describe four SEL curriculum features and are presented in Table 5. These features resemble the list of promising practices outlined in a 2002 National Research Council publication that has widely influenced the OST field (Eccles & Gootman, 2002).

An OST Intervention Design for Vulnerable Adolescents

Although it is possible to either review all of the practice indicators individually or summarize discrete characteristics across all 78, there is also a more integrated story about an SEL intervention design that was present in all of the exemplary offerings. The SEL offerings were structured to build a wide variety of skill sets, including community service, elder care, community organizing, conflict resolution, theatre, boat building, anti-racism training, relationship wisdom, community arts, poetry, the local policy making process, and others. Each employed a project sequence with challenging content linked to unique skill sets—skills to build a boat are very different from those necessary to deliver a high-quality theatrical performance. However, a common intervention design for building SEL skills with vulnerable adolescents was also present across the offerings. In this intervention design, youth learned SEL skills while doing the challenging work projects, which we refer to as the *project curriculum*. It was through the demands of the project curriculum that opportunities arose for youth to practice SEL skills and for staff to implement an *SEL curriculum*. In contrast to the very different skill sets developed through the very different project curricula, the set of SEL skills required to build boats or mount a theatrical production are the same skill set.

The SEL curriculum is described in Table 5 (see SEL Field Guide, p. 18-31) and consists of several elements, including safe space, responsive practices, and a planned sequence of one-on-one or small group check-ins that occur during the offering cycle. The SEL curriculum also includes *responsive practices* that are delivered in the moment as youth experience the project curriculum, in particular during moments of challenge, and are defined as:

- Facilitating: Staff helps to foster or sustain youth-led group dynamics and successful collaboration.
- Scaffolding: Staff set up or adapts skill-building opportunities to suit youth's current skill level, interests, or connections, keeping the work challenging but possible.
- Modeling: Staff demonstrates or exhibits practices, characteristics, or skills they intend for the youth to emulate or develop.
- Coaching: Staff monitor, focus, and support youth's learning experience by providing perspective, feedback, or encouragement that is respectful of youth's autonomy (Smith, McGovern, et al., 2016, pp. 27-28).

One of these four responsive practices appear in 40% of staff behavior practice indicators in Table 4 and occur across all six domains. The SEL curriculum also included a calendar of planned meetings and check-ins, often one-on-one with adult staff, where socially- and emotionally-relevant content from inside or outside the program was allowed to surface and be discussed.

The two-tiered curricula—project curriculum and SEL curriculum—were also designed with an awareness of youth development and skill progression necessary for specific types of SEL skill learning to emerge. SEL practices were often delivered in a specific sequence, beginning with the empathy domain and progressing through teamwork to initiative and responsibility, whereas emotion management and problem solving were a consistent focus throughout. Further, most of the offerings followed a cycle-in, cycle-out model. This model first allowed youth to think about their place in various social structures (e.g., social class, race, neighborhood, etc.), then moved inward toward personal stories and histories, and finally cycled back out near the end of the project to take a place in their local communities through services, presentations, or political action (see the Smith, McGovern et al., 2016, SEL Field Guide, pp. 21-22).

Finally, in the SEL curricula, the practices described in the SEL standards were applied in a way that was broad and integrated. As reflected in Appendix Table B-5, SEL practices from all six domains were implemented in nearly all offerings and, as reflected in Table B-6, almost all of the practice indicators were present in multiple offerings.

This set of intervention elements—intensive participation in challenging project curricula; SEL curricula, including 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 practices in the six domains—constitutes a project-based learning with intensive co-regulation intervention design for OST offerings. This intervention design is focused on building SEL skills with vulnerable youth and is certainly not new.[xiii]

The term *co-regulation* refers to adult behavior designed to help children and youth successfully self-regulate; for example, to stay focused, keep moving, process emotion, and get the task at hand completed (Murray et al., 2015). Briefly, co-regulation draws on much prior research on the optimal range of self-regulation where moderate levels of stress (e.g., achievable expectations, deadlines, skill hierarchies) heighten performance, whereas negative effects on performance occur when stress levels get too high or too low. Typically-developing children and youth need less co-regulation from parents and adults as they move through the early life course. Youth with atypical patterns of development due to exposure to trauma or chronic stress may need higher levels of co-regulation from adults. Co-regulation is what happens when staff uses responsive practices to keep the stress and strain of a challenging project curriculum in the optimal range. The elements of the SEL curriculum, in particular the responsive

practices, are strategies for co-regulation of attention, motivation, emotion, and behavior—that is, adult supports for youth self-regulation leading to the experience of agency.

Finally, it is worth noting that the project curricula presented *challenge* in at least two ways. For some projects, the end goals of the work were more fully defined (e.g., a 14-foot boat, a play about your personal experiences), whereas for others the goals were more abstract (e.g., community service project, community arts project). Either way, high-level problem solving and repetitive practice were required in order to be successful with the project curriculum and thus created opportunities for implementation of the SEL curriculum during moments of stress or strain. Perhaps more importantly, these curricula all sought to create circumstances where youth (a) were producing novel solutions to some of the problems that emerged and (b) could be supported (e.g., given time, not judged) to take the work in new directions. In this sense, the concept of co-regulation should be extended to include the adult role in supporting youth choices about both goals and/or processes necessary to meet them. [xiv]

Table 4. Standards for SEL Practice

Emotion Management
Key Youth Experiences
Range of emotions. Youth experience a range of positive and challenging emotions in a safe context.
(EM1) Youth engage in program work and activities in which emotions occur, are expressed, and are recognized as an important and often valuable component of human experience. (EM2) Emotions are experienced within a shared program culture (e.g., rules, norms) structured to make emotional expression and reflection safe and supported.
Emotion awareness and skill. Youth practice and develop healthy and functional emotion skills.
Youth practice (EM3) being aware of, identifying, and naming emotions, (EM4) reasoning about causes and effects of emotion, (EM5) using strategies for healthy coping with strong emotions and for harnessing emotions to advance the program work.
Staff Practices
Structure. Staff creates and adjusts the structure of daily activities to accommodate youth's processing of emotion.
(EM6) Staff creates time, space, or rituals within program activities for youth to process and learn from emotion. (EM7) Staff adapts program activities to respond to youth's emotional readiness and needs.
Modeling. Staff model healthy strategies for dealing with emotion within the context of caring, mutually-respectful relationships with youth.
(EM8) Staff model healthy strategies for dealing with emotions such as: a) active listening, remaining calm during intense episodes, and using problem-solving methods; b) communicating effectively and honestly about emotions (including their own); c) respectfully acknowledging and validating emotions in others.
Coaching. Staff provides coaching to youth about handling and learning from their ongoing emotional experiences.
(EM9) Staff provide coaching that is respectful of youth's emotional autonomy, including: a) using deep understanding of youth and their emotional styles to monitor, appraise, and respond in the moment to youth's ongoing emotions;

b) fostering emotional awareness and reflection; helping youth frame the situation and emotion;
c) encouraging problem solving in response to challenging emotions and the situations creating them; suggesting strategies for dealing with them.

Empathy

Key Youth Experiences
Inequality and identity. Youth explore social structure and power in relation to themselves and others.
(E1) Youth explore effects of stereotypes, discrimination, and social structures (e.g., based on race, gender, class, sexuality, religion, ability, etc.). (E2) Youth own and articulate their identities, including in relation to these social structures.
Diverse perspectives. Youth share their stories and listen to the stories of others.
(E3) Youth develop and share personal stories, and (E4) provide attentive, empathic listening to the experiences, backgrounds, and perspectives of others.
Acceptance. Youth practice relating to others with acceptance and understanding.
(E5) Youth practice identifying, understanding, and managing judgments and (E6) experience empathy and demonstrate caring when others reveal or share emotional experiences.
Staff Practices
Structure. Staff provide programs with appropriate structure for sharing experience and promoting equity.
Staff cultivate a safe and caring space, including: (E7) Employing appropriate structure for sharing different cultural backgrounds, personal beliefs, and stories (particularly those that are emotionally charged) without judgment. (E8) Actively promoting inclusion and equity and demonstrating support for the principles that all are different, equal, and important. (E9) Cultivating a program culture in which people actively care for each other. (E10) Providing programs with ritual structures for multiple sessions that allow youth to first check in, then open up, and end with reflection.
Modeling. Staff model empathy skills with youth.
(E11) Staff model empathy skills, including: a) intentionally recognizing the influence of their own identities and how these may affect interpersonal interactions; b) active listening; c) serving as an ally for youth who are isolated by differences in culture, family background, privilege, or power; d) modeling boundary-setting, including sharing or withholding personal experiences as appropriate and as needed.

Teamwork

Key Youth Experiences
Trust and cohesion. Youth develop group cohesion and trust.
Youth participate in work teams that (T1) develop cohesion and trusting relationships and (T2) a sense of group identity and purpose.
Collaboration. Youth participate in successful collaboration.
(T3) Youth work together toward shared goals. (T4) Youth practice effective communications skills (e.g., turn-taking, active listening, respectful disagreement). (T5) Each group member's contribution is valued and affirmed.
Team challenge. Youth manage challenges to creating and maintaining effective working relationships.
(T6) Youth practice managing the challenges of group work, such as miscommunication, obstructive behavior, and conflict over goals and methods.

Staff Practices
Structure. Staff provides programs with norms and structure.
(T7) Staff helps youth cultivate norms and rituals for effective group work.
Modeling. Staff model teamwork skills with youth.
(T8) Staff model sensitive and high-level interpersonal functioning in staff-youth and staff-staff interactions.
Facilitating. Staff facilitates or intervenes as needed to foster or sustain youth-led group dynamics and successful collaboration.
(T9) Staff facilitates or intervenes as needed to foster or sustain youth-led group dynamics. This includes: a) cultivating mutual accountability (e.g., by communicating the importance of all youth's successful contributions to the group's work) (See also Responsibility); b) intervening only as needed, allowing youth to lead group processes; c) helping to manage individuals' personalities when warranted (e.g., through one-on-one conversations before, during, or after a group activity); d) diffusing unconstructive conflict, regrouping, reorganizing, getting group back on track and functioning well.

Responsibility

Key Youth Experiences
Roles. Youth take on roles and obligations within program activities.
(R1) Youth choose or accept roles and their obligations; in some cases they initiate the roles.
Demands. Youth encounter difficult demands.
(R2) As youth get into the roles, they encounter demands, requirements, and obligations; they understand that their actions in response to these demands will impact self, peers, or others.
Accomplishment. Youth draw on resources to fulfill challenging roles and internalize accomplishment.
(R3) Youth draw on resources to successfully fulfill roles and obligations. Resources include drawing on inner strength, commitment, or newfound resolve; a sense of obligation to their peers and the program goals; and/or leaders' support and encouragement. (R4) Youth succeed in their roles and internalize the experience of having fulfilled valued roles.
Staff Practices
Structure. Staff provide structured but open-ended roles for youth.
(R5) The program design and the staff help create a variety of roles for youth that: a) have clear expectations and requirements; and b) have sufficient flexibility to allow youth initiative and ownership and accommodate youth's growing skills. (R6) Staff helps fit individual youth to roles appropriate to their interests and capacities.
Modeling. Staff model and fulfill their own roles.
(R7) Staff model and fulfill their own roles in the program, defining and discussing them with youth.
Coaching. Staff promotes high expectations, respect youth's ownership of their roles, and provides help only as needed.
(R8) Staff articulate, encourage, and enforce high accountability for youth living up to roles and obligations. (R9) Staff vigorously supports youth's ownership, empowerment, and latitude for decision-making within their roles, providing assistance only as necessary.

Initiative

Key Youth Experiences
Set goals. Youth set ambitious but realistic goals.

(I1) Youth have experiences setting challenging but achievable short- and long-term goals.

Motivation. Youth develop and sustain motivation by doing work that matters to them.

(I2) Youth develop motivation as they:

a) form connections with collaborators;
b) build skills and confidence; and
c) see the value in the work for their futures (adult roles and career), their communities, and the world.

Perseverance. Youth have experiences persevering through the ups and downs of difficult work.

(I3) Youth have repeated experiences of persevering through strenuous tasks and challenging work.

(I4) Youth experience the satisfaction of accomplishment and social acknowledgment of their efforts and achievements.

Staff Practices

Scaffolding. Staff provides ongoing assistance to help youth develop motivation within the work.

(I5) Staff help youth develop motivation by having youth select or shape the program goals and project(s) according to what matters to them.

(I6) Staff support youth's discovery of personal motivation in the program work by kindling youth's experience of belonging, competence, and connection of the program work to personal goals or societal purpose.

Coaching. Staff encourages youth to persist through challenging work, making sure that the effort behind youth's achievements is recognized.

(I7) Staff gives youth opportunities to persevere through challenges, setbacks, tiredness/tedium/boredom and also provide encouragement as needed to keep youth's attention focused and their effort engaged in keeping the program work moving forward.

(I8) Staff help youth see the progress and successes that come from their effort and perseverance.

Problem Solving

Key Youth Experiences

Projects. Youth engage in projects that involve organizing actions over time.

(PS1) Youth build project-specific knowledge and skills (e.g., carpentry, leadership, public speaking).

(PS2) Youth conduct projects that require organizing multiple, cumulative steps of work (e.g., creating a work of art, planning an event or a service project).

Planning-action cycles. Youth learn through cycles of strategic planning, execution, responding to emergent problems, trial and error, and reflection on outcomes.

(PS3) Youth engage in planning, including:

a) brainstorming and generative planning;
b) thinking strategically about the purposes, methods, content, and outcomes of the project;
c) anticipatory thinking, if-then thinking (e.g., about how the work and various constraints interact), and contingency planning.

(PS4) Youth have multiple opportunities to practice implementing the same skills to achieve greater success (e.g., by trying and trying again).

(PS5) Youth grapple with adjusting short- and long-term goals and strategies to emerging challenges and changing circumstances in their work.

Outcomes verify skills. Youth reflect on how outcomes of their work provide information that helps build and verify youth skills.

(PS6) Youth reflect on the outcomes of their efforts at all stages of the work to identify mistakes and successes, note progress, and identify current challenges.

(PS7) Youth's sense of self-efficacy, accomplishment, or confidence grows as outcomes demonstrate their developing skills, and they critically evaluate how their actions influenced outcomes. (See also Initiative)

Staff Practices
Structure. Staff provides sufficient structure to youth-driven projects.
(PS8) Staff provides training experiences for youth to help them learn project-related skills. (PS9) Staff places a high priority on youth having latitude to make choices and learn from experimenting within their projects. (PS10) Staff set high expectations and structure projects that are achievable (e.g., by setting goals, setting timelines and deadlines, setting boundaries).
Modeling. Staff creates opportunities for youth to observe models of successful work.
(PS 11) Staff model skills youth need to learn for their projects (e.g., carpentry or speaking skills, skills for planning and problem solving) and expose youth to models of successful work that set high expectations (e.g., youth learn about projects from prior years, novices work with veteran youth or expert staff).
Scaffolding. Staff provides assistance, as needed, to help youth learn and solve problems on their own.
Staff scaffold youth progress on projects by balancing: (PS12) stepping in to provide assistance and input as needed to help youth solve problems and learn (e.g., helping youth develop strategies when stuck or unsuccessful). (PS13) stepping back to support youth's increasing independence in their work as their skill grows and to allow youth space to struggle with challenges.
Reflection. Staff offers youth opportunities for reflection on project outcomes.
(PS14) Staff ensures that youth have opportunities to reflect on the processes that led to the outcomes of their work and to evaluate the impact and meaning of completed projects for both the youth and other stakeholders.

Table 5. Standards for Curriculum Features

Curriculum Features
Project Content Sequence
1) Staff shapes the offering work with youth input, often requiring youth ownership. 2) Staff shapes the offering work with complex goals and/or a complex sequence of operations. 3) Staff shapes the offering work with repetitive skill practice in diverse contexts.
SEL Content Sequence
1) The offerings follow a progression through the SEL domains. 2) Offerings are structured for youth to engage their community. 3) Youth master social and emotional skills and experience increasing agency.
Safe Space
1) Staff cultivates ground rules for group processes (e.g. listening, turn-taking, decision-making) and sharing of emotions. 2) Staff cultivates a culture around the principles that all are different, equal, and important in which people actively care for, appreciate, and include each other. 3) Staff cultivates a culture where learning from mistakes and failures is highly valued. 4) Staff organizes consistent routines, activities, roles, or procedures to provide a structured and predictable experience.
Responsive Practices
1) Staff observes and interacts in order to know youth deeply. 2) Staff provides structure for check-ins to actively listen to and receive feedback from individual youth. 3) Staff coach, model, scaffold, and facilitate in real time as challenges occur.
Supports for Staff

1) The organization recruits youth who will benefit from the offering.
2) There is more than one staff member in every program session with the ability to implement responsive practices.
3) Staff works together before each program session to plan and collaborate on the session activities and regularly debrief following each session to discuss youth progress, staff response, and adjustments for future sessions.
4) Staff is supported to grow professionally and rejuvenate energy for the work.
5) Staff is supported by their organization to reflect on and improve their practices through a continuous improvement process.

Chapter Three. SEL Performance Measures

We conducted performance studies at each of the Challenge sites. The products of the performance studies are a suite of SEL performance measures and a set of performance benchmarks for OST organizations. Readers who are primarily interested in the benchmarks and other results from the performance studies should review the summary of measures in Table 6 and then skip ahead to Chapter 4. In this Chapter and related appendices, we provide technical details to support selection decisions and implementation planning for a suite of SEL performance measures configured for use in OST settings that are focused on improving SEL practices and skills.

Summary of SEL Measures

The SEL Challenge performance studies include 13 measures of setting-level features and 12 composite measures for individual youth skill. In Table 6, the setting measures describe characteristics of the environments where staff and youth carry out the work of OST and differentiate between the system, organization, and point-of-service levels of setting. Youth SEL skill measures describe beliefs and behaviors of youth and were implemented at three time points to describe skill change. Sample performance reports at baseline and end-of-cycle reports based on these measures are available at http://cypq.org/SELChallenge along with complete codebooks, permission forms, and other supporting documents.

Table 6. Summary of SEL Performance Measures – Construct Name and Description

System Level – Policy Quality
Accountability: Accountable for quality, shared quality standard, collaborates across sites.
Organization Level – Management Quality
School Day Content: Link with school day academics, participation in parent-teacher conferences.
Staff Capacity: Staff is trained, received program orientation, has adequate retention and staff/student ratios, are given time to plan, and have student goals in mind for program objective.
Horizontal Communication: Staff co-plan program policies or activities with other staff, discuss problems, and observe or are observed by other staff.
Vertical Communication: Supervisor provides feedback, is visible during program, knows what is being accomplished, challenges staff, and makes sure program goals and priorities are clear.
Job Satisfaction: Position is close to ideal, satisfied with job and would not change career.
Manageable Workload: The workload does not prevent staff from doing a good job

Point-of-Service Level – Instructional Quality
Youth Governance: Youth begin their own projects, select content, and design space. Youth involved in hiring, budgeting, return as leaders, develop partners.
Curriculum Planning: Sessions are planned in advance, targets specific learning goals, builds upon prior sessions, takes into account student feedback, and combines academic content with student interests.
Growth & Mastery: Students exposed to new experiences, responsibilities, and tasks that increase in complexity, long term group projects, acknowledge achievements, and identify personal strengths.
Instructional Quality: A structured environment with guidance and encouragement, opportunities for leadership and collaboration, and the capacity to promote planning and reflection.
Engagement: Youth find activities important, use skills, have to concentrate, experience moderate challenge.
Youth SEL Skills: Beliefs about Self and Behavior in Setting
Emotion Management - Beliefs: Optimism; Emotion Reappraisal; Identification of Emotions - Behavior: Identify positive and negative emotions (e.g., excitement, anger, joy)*;* Reason about causes and uses of emotion*;* Manages emotions for functional purpose
Empathy - Beliefs: Adolescent Empathy - Behavior: Values own/others perspectives and stories with sensitivity to context
Teamwork - Beliefs: Adolescent Social Competency - Behavior: Practices respectful and effective communication within a team*;* Coordinates and supports action toward team goals
Responsibility - Beliefs: Adolescent Diligence and Reliability - Behavior: Fulfils Roles and Commitments*;* Successfully defines, adjusts, and negotiates roles and commitments when required
Initiative - Beliefs: Adolescent Initiative Taking; Adolescent Purpose - Behavior: Develops and hones motivation for the OST task*;* Perseveres through internal and external circumstances that challenge the OST work
Problem Solving - Beliefs: Adolescent Goal Orientation; Problem Solving Strategies - Behavior: Intentionally learns OST-task related methods and tools (e.g., carpentry)*;* Uses problem-solving skills to develop, evaluate, and adapt a course of action; Successfully manages time; Connects with external stakeholders; Reflects on learning and significance of results

The remainder of this chapter provides technical information to support decisions about selection and/or implementation of SEL performance measures. The technical information includes: a theoretical argument for the suite of Challenge performance measures, evidence for reliability and validity of data produced by the measures, information about feasibility of implementation of the measures by organization staff, and a summary of the analytic approach used to produce the results in Chapter Four.

Theory for Continuous Improvement Intervention and Youth Skill Measurement

In this section, we extend prior research in the area of setting measurement to focus on intra-individual mental processes and behavior. In order for the continuous improvement intervention focused on SEL skill learning to be effective, we need a theory about what that skill is; how it can be measured, how it can be influenced, and how we will know if it changed. This requires a model of active ingredients and dynamics within and across multiple levels: from the policy-system level, to staff behavior at the organization and point-of-service levels, through youth mental processes, and on to youth behavior. In assembling theory to support our selection and development of a suite of measures for SEL offerings, our goals were:

- Integrate multilevel systems thinking into our theory of continuous improvement by, for example, differentiating among context, mental processes, and behavior;
- Develop theory that includes explicit reference to attention and emotion, given their central role in any definition of SEL;
- Use theory to develop SEL practice content and performance measures that can be integrated into the continuous improvement intervention.

Figure 3 summarizes the logic model we developed for the SEL Challenge and to which the 25 performance measures described in Table 6 were aligned. On the left-hand side of the model, a multilevel context system reflects multiple, *materially-nested* levels of setting where OST services are organized. At the *system level*, coordination of policy, funding, and regulation occurs. At the *organization level*, coordination of staffing, planning, and management of basic processes occurs. At the *point-of-service level*, adult-youth interaction and instruction occurs in wide variety of informal program spaces, workshops, classrooms, and field sites that house the OST service.

On the right-hand side, a multilevel person system reflects both (a) multiple levels of *functionally-nested* mental process (i.e., iconic, symbolic, and phenomenological) and (b) behavior that results when those mental processes engage with the context. Further details about the mental processes characterizing the multilevel person system, and the Basic Levels of Self model upon which they are based (Roeser et al., 2006), are described below in the section, "Theory for Youth SEL Skill Measurement," and in Appendix C. Here, consistent with our desire to focus on mental processes related to skill learning, attention, and emotion, Figure 3 highlights the extent to which (a) contexts, behaviors, and mental processes are differentiated, both theoretically and operationally, and (b) the effects of contexts on behavior are mediated by mental processes.

In the next two sections, we first summarize our *theory for continuous improvement intervention* and the attributes of performance data necessary to do the work of continuous improvement. We then

describe a *theory for youth SEL skill measurement* designed to produce performance information about youth SEL skills that can be used in the continuous improvement intervention.

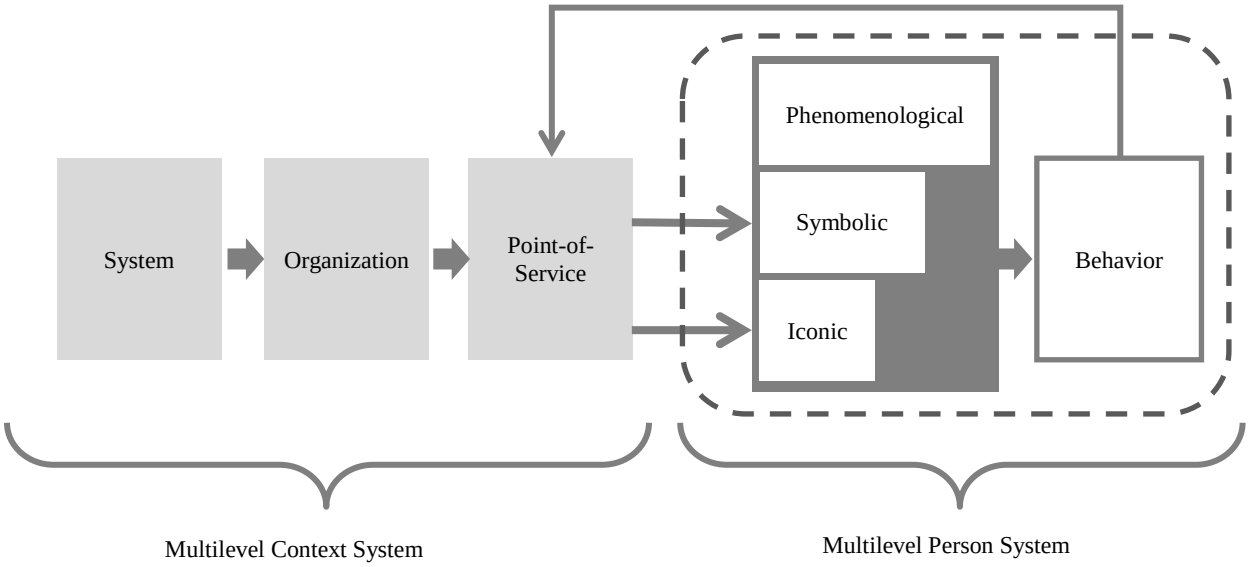

Figure 3. **Person-in-Context Model for the SEL Challenge Performance Measurement**

Theory for Continuous Improvement Intervention

Together, the two sides of Figure 3 provide a logic model to which a continuous improvement intervention can be aligned. A continuous improvement intervention consists of four elements: (a) a standard for good practices and aligned performance measures at multiple levels of setting, (b) data products that communicate about performance in actionable ways, (c) a continuous improvement cycle implemented over multiple program cycles, so that staff have opportunities to practice new skills and develop expertise, and (d) technical assistance, training, and technology necessary to implement the prior three elements (Smith, 2013).

The objective of the continuous improvement intervention is to produce a cascade of effects over the system, organization, and point-of-service levels of setting. The cascade of effects occurs as actors engage in skill learning at one level and then enact that learning as behavior at the next level. For example, managers are trained off-site to implement the continuous improvement intervention for SEL practices. Once trained, the managers return to their organizations to engage instructional staff in the continuous improvement intervention elements. Next, staff engages with the continuous improvement intervention elements and then enacts higher-quality SEL instructional practices with youth. The youth then engage more fully with the improved offering content and enact SEL skill in the OST setting. Finally, as youth experience mastery of SEL skills (i.e., skill growth), they enact new SEL skills in other

settings (i.e., skill transfer). The cascade of effects has been demonstrated in multiple studies, and the intervention is widely used in the OST field. Appendix A provides additional detail on the continuous improvement intervention elements, validation evidence, and scale of use.

The design for the continuous improvement intervention entails multiple components (e.g., service goals, service theory of change, performance measures, data products, incentives, improvement cycle; Smith, 2013), complete descriptions of which are beyond the scope of this report. However, the characteristics of the performance measures and data products are directly implicated in our efforts to develop a suite of SEL performance measures. Effective performance measures and data products that support action sequences focused on improvement require performance data that have eight characteristics:

- *Timely*. Data that are available in real time as events occur, or just after completion, are more likely to hold relevance for actors.
- *Objective***.** Objective data are focused on behaviors and conditions that can be identified through observation and easily named in relation to practice.
- *Reliable.* Data should be seen by all stakeholders as precise and factual due to standardization of measures/methods, clarity about the object and method of measurement, and repeated use of the instrument in field testing.
- *Sensitive.* Performance measures are focused on behaviors and conditions that are likely to change in response to intervention and can be used to describe change over a relevant performance period.
- *Valid.* Data are valid when they describe behaviors and conditions that are links in a causal chain of events desired by the actors involved (i.e., favoring insight about mechanism over prediction).
- *Feasible.* Data collection must be a feasible (i.e., the minimum data necessary are collected using typical organizational resources and from typical respondents).
- *Multilevel.* Ideally, measures should be designed to directly assess phenomena occurring at a specific level of the context or person. However, measures applied at lower-level units of analysis (e.g., staff) can be made useful at higher levels (e.g., organization) when aggregated across individual units to compose a higher-level rating. Rules for composition of information from lower-level units into representations of performance at higher levels require items that have an explicit theory for composition and rules for the necessary level of group agreement.
- *Multipurpose.* Performance data are multi-purpose when *both* data collection and data interpretation promote a shared language among actors and a framework to guide discussions about performance. In particular, observation-based data collection methods used by

organizational staff build shared understanding of the objects of measurement and typical performance levels.

The suite of SEL performance measures developed for the Challenge produce data that roughly meets these eight requirements, although substantial gains in the precision of measures will likely be made in subsequent iterations.

Theory for Youth SEL Skill Measurement

In this section, we present theory regarding person-in-context dynamics relevant to OST-SEL settings by describing three distinct kinds of mental content, or levels of representation: symbolic, iconic, and phenomenological. As described below (and, in more detail, in Appendix C), the distinctions and interrelations among these three kinds of *structured mental content* are useful for understanding and explaining basic processes of skill learning, particularly in relation to person-in-context dynamics (e.g., Type 1 agency) and intraindividual mediating processes, such as attention, reflection, and emotion (e.g., Type 2 agency). Behavior is described as an external manifestation of mental engagement with the context and is considered separately.

Although the specific elements of the multilevel person system shown in Figure 3 offer a gross simplification of mental contents and processes involved in skill learning, these elements and their interactive dynamics are of specific interest for OST programs because they are integral to mediating the cascading effects of context quality on individual behavior (e.g., point-of-service engagement and skill transfer). OST programs are designed to build youth skills by activating youth's prior knowledge (e.g., skills, memories, beliefs, schemas) and supporting youth's self-regulation of attentional and emotional processes. In contrast to other education and human services settings for youth, OST programs are designed to help youth feel safe, focused, and motivated[xv] around content that the setting has to offer, and the multilevel person system contains integral ingredients for making those experiences possible.

Symbolic Level of Mental Representation. The symbolic level depicted in Figure 3 includes basic beliefs and more complex belief systems (e.g., knowledge, attitudes, goals, and plans). These mental contents and processes are frequently equated with skill in the forms of declarative (i.e., names of things) and procedural (i.e., how things work) beliefs, as well as efficacy beliefs about skill potential (Bandura, 1977; Marzano, 1998; O'Neil, Perez, & Baker, 2014). Symbolic-level contents have also been referred to as theories about the self, the world, and how the self fits into the world (Epstein, 1985), indicating how the symbolic level carries important information about "me," the youth's specific configuration of beliefs about, for example, who they are, what they can do, where they are going, and where they belong. In these terms, skills are encoded in youth's beliefs and belief systems.

Referring to symbolic-level beliefs may imply the youth's conscious awareness of mental skills and conscious intentionality of behavior, but this is not necessarily the case. Skill activation occurs both within and outside of youth's conscious awareness, and, in many cases, mental engagement with the context is fairly automatic as, for example, when an interesting context activates unique stores of experience and memory. Consistent with the growing literature on dual-process models of psychological functioning (see Appendix C), mental contents are usually activated automatically (e.g., by other thoughts and context triggers) and, once activated, influence behavior despite whatever conscious decisions might appear to be occurring simultaneously.

Although the automatic activation of beliefs in response to contextual cues may sound relatively passive, it is not. Automatic activation occurs only when beliefs are both *available* (i.e., exist for a given person) and *accessible* (e.g., have been encoded in memory in a way that makes them sensitive to salient contextual cues) (Higgins & King, 1981). The probability of such automatic activation can be influenced further by how youth choose consciously to focus their awareness within the immediate context. In this sense, then, youth are active participants in their learning, self-organizing their unique phenomenological experiences from the activation of their particular symbolic beliefs (and iconic level schema and scripts; see Appendix C, Figure C-1).

Symbolic-level beliefs and belief systems are activated or deactivated, unconsciously or consciously, by youth's mental engagement with the context. Settings that use context to effectively activate youth's prior knowledge and goals are more likely to build youth motivation to engage with the content of the setting and extend repetitive skill practice. The project and SEL offering curricula described in Chapter Two were designed to build mental skills (e.g., declarative, procedural, efficacy beliefs) related to both the project curriculum (e.g., carpentry skills) and SEL curriculum (e.g., teamwork).

Because different youth bring different prior experiences to the OST setting, design of the setting in a way that activates students' prior skills—we might say playing to their strengths—depends on the staff's ability to know who the youth are and what is valued in their communities. The expert practitioners in the Challenge understood this point: Opportunities to know youth deeply was one of the elements that defined the Safe Space curriculum feature, and youth talking about their experience to both staff and peers is a prevalent activity across the standards.

Iconic Level of Mental Representation. The iconic level includes sensory-affective-motor schemas and more highly-developed sensory-affective-motor scripts that, along with symbolic level beliefs, make up the content of long term memory where mental skill is encoded. Understanding mental dynamics associated with these more holistic schemas and scripts is necessary to explain how emotional experience mediates youth's mental engagement with context.

In contrast to the symbolic representation system, which tends to store and process information cortically as abstract, analytic, logical, verbal forms of knowledge typical of belief systems, the iconic representation system tends to store and process information subcortically as concrete, holistic, associative, nonverbal forms of knowledge typical of sensory-affective-motor schemas. In these terms, iconic representations reflect emotional memories and experiential triggers, often related to prior experience of trauma and chronic stress, and can be barriers to both the successful activation of existing skills and the successful focus of attention on new skill content.

An important reason to distinguish iconic-level contents such as sensory-affective-motor-schemas from symbolic-level contents is to denote the immediacy of their operation: A contextual cue can lead to primary appraisal and behavioral response in milliseconds, long before a secondary-appraisal process brings this action sequence to conscious awareness and potential self-management. The application of an SEL practice to support youth's secondary appraisal of fast-moving, emotionally-charged mental and behavioral events also raises the potential that the effects of contextual cues on youth behavior can be mitigated as youth become aware of what triggers them and construct alternative responses so that these triggers are defused (Bryck & Fisher, 2012; Teper, Segal, & Inzlicht, 2013).

In particular, when youth who have suffered difficult SEL histories (e.g., through prior exposure to traumatic situations or chronic stressors) come to intentionally reflect on information in the context that triggers counter-productive emotional and attentional responses, in such a way as to better manage those emotional and attentional responses, secondary appraisal is occurring. For example, if youth arrive at program offerings with attachment-anxiety or social phobia, these more deep-seated issues may prevent youth from forming relationships with staff and youth and prevent the focusing of attention on program content. The SEL curriculum and responsive practices identified in Chapter Two—scaffolding, coaching, modeling, facilitating—are specifically designed to help staff respond quickly when disruptive emotions and loss of focus occur during moments of challenge presented by the project curriculum.

Phenomenological Level of Mental Representation. The phenomenological level includes currently activated iconic- and symbolic-level content. By focusing on attentional processes at the phenomenological level (i.e., attention to currently activated iconic- and symbolic-level content), we can distinguish between important aspects of how youth self-organize their own skill-learning. First, self-organization can be described as the process of belief differentiation and integration (i.e., building new belief structure) stimulated by both context triggers and secondary-appraisal processes. As youth encounter the context, relevant prior beliefs (i.e., skills) are quickly activated and reconstructed in response to the specifics of the immediate context, such that novel contexts can engender novel reconstructions that promote skill growth. In these terms, youth with more relevant background belief

content (also known as background knowledge[xvi]) can more flexibly think about, and behave toward, contextual opportunities and constraints.

However, the full potential of youth's self-organizing power occurs when individuals consciously focus their attention and awareness on (a) the immediate contextual cues and (b) the meaning of those cues for their learning goals. Individuals will gain the maximum benefit from well-supported contextual opportunities when they sustain the focus of their awareness, often during moments created intentionally for this purpose by the curriculum (e.g., planning and reflection), on relevant portions of the incoming stream of contextual information. Regardless of whatever benefits are achieved by bringing an already well-developed set of available and accessible beliefs (i.e., skills) into well-supported learning contexts, these benefits will likely increase when individuals focus consciously on relevant portions of the material being presented and use that focus to engage in secondary-appraisal processes designed to further elaborate or strengthen existing beliefs or construct and encode new beliefs. Similarly, benefits will likely decrease when individuals allow their focus of awareness to shift to either (a) irrelevant portions of the incoming stream of contextual information, or (b) irrelevant thoughts that may have become activated as the result of either internal or external cues.

Describing SEL skill-learning processes in terms of self-organization and self-regulation, and the result of these two processes as a youth's experience of agency, we are referring to agency in two senses (see Peck, 2007; Endnote i; and Appendix C). In the first sense, Type 1 agency, we refer to the more automatic, primary-appraisal processes characterizing self-organization that have been described as "less effortful processes associated with stress physiology, emotional arousal, and attention focusing"... and further as "reactive, highly automatic, and phylogenetically older styles of response" (Blair & Raver, 2012, p. 648). In the second sense, Type 2 agency, we refer to the more intentional, secondary-appraisal processes, characterizing self-regulation (or executive function), that have been described as "working memory, inhibitory control, and the flexible volitional shifting of the focus of attention" (Blair, 2016, p. 1) and that entail conscious reflection on the meaning of information presented by the context and deliberate decision-making processes (e.g., planning and goal setting).

Optimal skill learning results from actively paying attention to the information presented, and reflecting on the meaning of that information.[xvii] This dynamic raises the potential for SEL programs to teach active awareness of contextual cues and their meaning as a way to build new belief systems that serve purposes reinforced by the OST curriculum and allow other, less useful belief systems (e.g., I can't learn, even if I work hard) that are not reinforced by the OST context, to fade. Further, and of particular importance for youth exposed to prior traumatic experience and chronic stressors, the application of the conscious focus of awareness to triggering contextual cues, behavioral events, and affective states creates the potential for SEL skill growth where it is potentially most valuable: in the potential alleviation of

suffering for individual youth and the improved contribution that these youth are able to make in communities where they learn and live.[xviii]

Behavior in context. Behaviors are important indicators of the application of mental skill because they are the ultimately-desired output of skill development processes. For individuals who internalize important aspects of the context over time (e.g., norms, values, and goals), behavior and its causes are likely to be related to setting processes (cf., behavior setting theory, Barker, 1968), meaning that behaviors called for by the context can be useful indicators of mental skill development. However, behaviors must be differentiated from mental processes because of the equifinal and multifinal relations between mental engagement with context and the resulting behavior; that is, people do similar things for different reasons and may react to the same situation in different ways because of how the context activates unique prior experience (Cicchetti & Rogosch, 1996; Richters, 1997; Roeser et al., 2006).

Implications of Theory for Selection of Performance Measures. The discussion of the right-hand side of Figure 3, the multilevel person system, raises several implications for the selection of performance measures for individual youth skill learning: Because beliefs are both influenced by contexts and influence behavior, we selected both belief and behavior measures for the SEL Challenge. Youth beliefs about their capacity to enact a skill (e.g., I can set goals) represent a behaviorally-oriented aspect of *efficacy* in that skill domain (Bandura, 1977). These beliefs about the capacity to enact behavior are predictors of actual behavior (Bandura, Caprara, Barbaranelli, Gerbino, & Pastorelli, 2003). Because SEL Challenge programs emphasize methods known to build domain-specific efficacy,[xix] beliefs about the capacity to enact context-appropriate behaviors are important outcomes of SEL programs and reflect the presence of declarative and procedural skill.

Behavior data is an important complement to belief data. Behavioral data makes good performance data because behavioral indicators reflect objective conditions that can be observed and provide direction for improvements to a setting. Further, to the extent that average SEL skill behavior in a setting is low or high, it is possible to treat mean levels of a behavior as a setting measure, reflecting the press of the context on the behavioral application of mental skill. Finally, given our earlier discussion of mental processes related to attention and emotion, we note that both belief and behavior measures of SEL skills include specific measures focused on emotional processes of primary and secondary appraisal of emotional events (emotion management domain) as well as youth's capacity to enact skills related to shifting and focusing awareness (problem solving domain).

The SEL skill measures used in the Challenge can be thought of as a measurement net to detect the presence of beliefs and behaviors that indicate skill levels and, over time, can be used to provide evidence of skill learning in either beliefs, behaviors, or both. Importantly, and following implications of the net metaphor, our proposed belief and behavior measures also reflect different levels of abstraction in

that the items for the belief measures refer to the youth's ability in general, irrespective of any specific context (e.g., I can set goals), whereas the behavioral indicators are designed to reflect the youth's direct proximity to the program context, with staff rating youth's frequency of specific behavior in the program in the past month. We realize that belief and behavior measures, with referents framed at different levels of abstraction, are unlikely to be highly correlated, and this was in fact the case in our data.[xx] We reasoned that the SEL domain-specific belief measures cast at a high level of abstraction might be a useful complement to behavior measures that were more closely focused on the OST setting. More specifically, because knowledge about how SEL skills change is the critical issue for performance measurement, we wanted to examine the extent to which beliefs about the self and world, and behaviors, relate differentially to skill growth as predicted by our theory of change. We expect that context-specific behaviors will increase more steeply during the project period than general beliefs about the capacity to enact a behavior in any setting.

When thinking about the nature of, and interrelations among, contexts, skills and behaviors related to SEL, the discussion of a multilevel person-in-context framework, and in particular, the Basic Levels of Self (BLoS) (see Figure 3 and Appendix C) model suggests that mental disengagement with the context may be a barrier to either activation of prior skill or intentional focus of awareness on practicing new skills. For this reason, we included several measures of both setting quality and youth engagement. The BLoS model further suggests that challenging SEL histories may lead to anxiety or avoidance issues that cause disengagement with the context and also reduce the rate of potential SEL skill learning. For this reason, we included measures of attachment-related risk, indicating the presence of iconic-level schemas and scripts that may carry difficult emotional content into the OST setting.

Reliability and Validity Information for Suite of SEL Measures

In this section, we provide sources as well as reliability and validity information for the setting and individual SEL skill measures. We also describe how we constructed the risk index to assess attachment-related anxiety, avoidance, and social phobia. Finally, we provide feasibility information for implementation of the suite of SEL measures and describe the analytic approach used to produce the results presented in the next chapter. Basic descriptive information at the scale and item level for all measures are provided in Appendix D. Appendix E provides additional detail for reliability, validity, and growth trajectories for the SEL belief and behavior measures.

Setting Measures

The performance measures used for the system, organization, and point-of-service levels of setting have a long history of use in the OST field and considerable evidence of reliability and validity. Most recently, data from these measures were the subject of a multi-year study of growth trajectories for a state 21st Century Community Learning Centers system (Smith, Moxley, McGovern, Helegda, & Roy, 2016). All measures demonstrated acceptable levels of reliability as indicators of individual beliefs about settings (internal consistency accounting for between-site variance) and as within-setting group means that can be used to differentiate between settings (using the intra-class correlation). Further, construct validity was assessed for all measures using latent variable models that accounted for between-site variance. In all cases, model fit statistics demonstrated mixed evidence of construct validity (i.e., at least one of the five model fit statistics was in the acceptable range). Finally, growth trajectories for each measure provided evidence of sensitivity to changes hypothesized to be driven by the continuous improvement intervention. A full discussion is provided in Smith, Moxley, et al. (2016).

Observation-based measures of staff instructional practice were generated using the Youth Program Quality Assessment (PQA) Form A (Smith & Hohmann, 2005). The Youth PQA Form A produces scores for 63 items comprising 18 scales in four domains: safety, supportive environment, interaction, and engagement. A composite rating, the instructional total score, was calculated for each offering by averaging the domain scores for supportive environment, interaction, and engagement at each time point and then taking the average over the three time points. Evidence suggests that the PQA instrument produces data of acceptable reliability using data collection procedures and rules for score composition like those employed in the Challenge (Naftzger, 2012; Smith, 2013; Smith, Akiva, Sugar, et al., 2012). Higher-quality instruction, as assessed by the PQA, has been associated with increased levels of youth engagement (Akiva, Cortina, Eccles, & Smith, 2013; Naftzger et al., 2013) and gains in academic skills demonstrated in both OST settings (Smith et al., 2015) and the school day (Naftzger, Devaney, & Foley, 2014; Naftzger, Hallberg, & Yang, 2014; Naftzger et al., 2013).

The PQA items were crosswalked against (i.e., compared to and aligned with) the practice indicators in the standards for SEL practice. Six SEL domain scores were produced to describe the quality of SEL practice in the offerings. Appendix E provides item-level descriptive information for all PQA items, results from the crosswalk, and descriptive information for the new SEL PQA domain scores.

Engagement of youth in the SEL offering content was also assessed by observers at each time point. At the end of the offering session, observers administered a brief survey to youth focused on their level of interest, challenge, and sense of belonging during the offering session for that day. These scores were averaged across all youth in the setting and then averaged again across the three time points to produce an overall youth engagement rating for each offering.

Individual Measures for SEL Skill

We measured individual SEL skills in six domains using youth survey measures of SEL beliefs and staff ratings of youth SEL behaviors. Belief measures included 10 measurement scales across the six domains. Behavior measures included 15 measurement scales across the six domains. Although in Chapter 4 we present findings using composite scores—that is, all measurement scales in a domain averaged together to produce a single belief or behavior score for each of the six domains—all reliability, validity, and growth trajectory analyses were conducted at the scale level for the 25 belief and behavior measurement scales.

Most of the youth belief measures were developed by the Child Trends organization as prosocial and protective indicators for adolescent well-being (Lippman, Moore, et al., 2014; Lippman, Ryberg, et al., 2014). Developers conducted cognitive interviews focused on item content, response scales, and scale anchors with adolescent participants drawn from low-income households that were one third European American, one third Hispanic, and one third Africa -American. In a larger, low-income youth sample, each measure has also demonstrated evidence of reliability and predictive validity with academic success (grades) and risk behavior (e.g., fighting, depression, and smoking). The factor structure for each measure was also tested with subgroups by gender, age (older or younger adolescents), and household income. As an earlier stage of this measure development project, Child Trends conducted a review of 80 studies using measures in these domains to produce evidence of criterion validity: studies in which an initial level of an SEL skill (often produced through a randomized intervention to change that skill level) was associated with a wide range of outcomes at least one year later.

The Child Trends measures did not cover the emotion management domain. For beliefs in this domain, we selected constructs for optimism (Scheier, Carver, & Bridges, 1994), emotional reappraisal and emotional suppression (Gross & John, 2003), and four items focused on identification of emotions from a widely used *emotional intelligence* scale (Schutte et al., 1998).

Behavioral measures were developed for the SEL Challenge through an iterative process, including literature review and input from expert practitioners. The items were scaled around expert descriptions of behavioral demonstration of the SEL skills targeted for youth in each domain. More specifically, we asked the experts to describe skill behaviors that youth did not demonstrate consistently at the beginning of the offering but often did demonstrate consistently by the end of the offering. The measurement scale was frequency of skill demonstration, where the lowest scale point was almost never demonstrated in the past month and the highest scale point was almost always demonstrated in the past month. Our primary goal was to create SEL measures that were highly sensitive to the effects of OST offerings so that skill change over the short timeframe of the offering could be assessed.

Following the Weikart Center's approach to building reliable and valid observational measures (Smith, Hallman, et al., 2012), we first worked with local-community experts to describe what skill demonstration looked like in their OST settings. We presented the preliminary items to the experts and revised them further by identifying conceptually-related groupings of items (measurement scales) and modifying items within each scale to assure that each scale included both easier and harder items.

Next, we worked with the experts to assure wording clarity, content validity, and substantive validity. Content validity refers to experts indicating that the behavior was important in their program setting. Substantive validity refers to experts indicating that the behavior was prevalent enough to reliably observe.

All 25 of the SEL belief and behavior measures demonstrated acceptable levels of reliability and construct validity. To test for construct validity, we conducted confirmatory factor analyses using multilevel structural equation models. Reliability and construct validity results are provided in Appendix E, including results for inter-rater reliability for the staff ratings of youth behavior.

Data Collection and Feasibility

The measures, permission forms, and protocols for the study were submitted to Chesapeake Institutional Review Board (Protocol Id Pro00010177) for review of protections for human subjects and approved on September 3, 2014. Data collection included youth and staff surveys, staff ratings of youth behavior, and observation-based ratings of the exemplary offerings. All data collection was conducted by the team evaluator who was a professional with evaluation experience, and also, not the program manager or lead staff for the target offering. All data were uploaded to a data portal at the Weikart Center website. Table 7 provides detail on instruments, items, time per complete, and time points at which the various instruments were completed.

Table 7. Schedule for Implementation of Measures

	T1 Oct 2014	*T2 Jan 2015*	*T3 Apr 2015*
Youth survey, 81 items, 15 minutes per complete	X	X	X
Youth day-of-observation engagement survey, 17 items, 3 minutes per complete	X	X	X
Staff rating of youth behavior, 51 items, 5 minutes per complete	X	X	X
Staff survey, 52 items, 15 minutes per complete			X
Observer PQA Form A, 63 items, 105 minutes per complete	X	X	X

The local evaluator was also trained on the use of all measures and developed a schedule for data collection. For staff surveys, youth surveys, and staff ratings of youth behavior, a two-hour webinar was conducted at baseline to explain written protocols and answer questions. For the Youth PQA Form A, (Smith, Akiva, Jones, et al., 2012; Smith & Hohmann, 2005), an evaluator training (Blazevski & Smith, 2007) is required to achieve at least 80% perfect agreement at the item level with a set of "gold standard" scores. An additional check-in webinar was also held at the second and third time points to troubleshoot and assure familiarity with protocols. Estimated staff time required to implement the performance measurement design in Table 7 was 53 hours for the local evaluator associated with each offering and a total of 32 hours for all instructional staff in a given offering. Averaged across a three-person team, implementation of the performance study required roughly 28 hours per person and for example, at $50/hour, a staff-time cost of $4,250. The training and other supports provided to the study participants was typical for many other implementations of similar packages of measures supported by the Weikart Center.

Analytic Approach

Analyses of setting data for the system, organization, and point-of-service measures included calculation of means, standard deviations, intra-class coefficients (ICC) for reliability (including alpha coefficients where appropriate), skewness, and other basic statistics to describe average performance on multiple dimensions of performance at each site and across sites at each time point.

The analytic approach for the youth skill measures included three steps: (a) reliability and validity analyses, (b) growth analyses, and (c) composite scores and multivariate profiles.

Reliability and Validity Analyses. We conducted reliability and validity analyses for all of the belief and behavior measures in the same sequence. All analyses were conducted using the 10 SEL belief and 15 SEL behavior scales. First, we examined internal consistency of the reporter for each scale, reporting Cronbach's alpha for group-centered data to remove the effect of between-offering variance. Next, we calculated intra-class correlations to assess the degree of similarity across reporters in the same offering. For the staff rating of youth behavior, we also calculated inter-rater agreement estimates at all three time points using Kappa and ICC statistics. Finally, we conducted confirmatory factor analyses, using models adjusted for the nesting of the data, and examined factor loadings and fit statistics.

Growth analyses. We conducted a range of univariate growth curve analyses to understand patterns of change over time. First, we used unadjusted scale scores to estimate three-level hierarchical linear models with time at level 1, youth at level 2, and offerings at level 3. We examined fixed effects for time (slope), covariance of the random effects for time and youth baseline-skill level (indicating the extent to which youth who start out low have steeper slopes), and random effects for offerings (indicating

the extent to which there are differences in slopes across offerings). Next, we estimated all 25 of the growth models again using a latent-variable approach (i.e., scale scores adjusted for measurement error) to compare the pattern of findings based on the unadjusted and adjusted scale scores.

Composite scores and skill profiles. This final analytic steps involved two parts. First, we created composite scores for each of the six SEL domains to produce 12 composite scores for each individual, one for SEL beliefs and one for SEL behavior in each of the six domains. The composite scores were created by averaging across scales within each domain to produce a domain-level mean for each participating youth. Literature in the field of organizational psychology describes predictive and criterion validity for composite scores created from the unweighted combination of multiple performance measures (Bobko, Roth, & Buster, 2007; Fralicx & Raju, 1982).

Second, we used pattern-centered methods (e.g., cluster analysis) to define multivariate profiles of youth skill across the six domains. Variable-centered methods (e.g., looking for relationships one variable at the time) provide less insight into our theorized structure of skill capacity and growth, which is explicitly pattern- or person-centered, meaning that multiple skills are enacted and grow together and in different combinations for different individuals in different contexts. Pattern-centered theory and methods are ideally suited to address this kind of complexity (Bergman, Magnusson, & El-Khouri, 2003; Peck, 2007; von Eye & Bogat, 2006). For example, pattern-centered theory suggests that individual skill sets function as integrated wholes, and pattern-centered methods allow us to treat patterns of values on youth skill variables as integrated wholes (i.e., profiles). Treating individual youth (instead of variables) as the unit of analysis allowed us to identify a small number of relatively-homogeneous subgroups characterized by distinct profiles of skill.

By using composite scores to describe performance in each of the six domains of youth social and emotional skill, and then pattern-centered methods to identify relatively-homogenous subgroups, we moved the unit of analysis from single unidimensional measures (variables) to qualitatively distinct individual patterns of skill enactment and capacity that simultaneously encompass multiple variables. In particular, we were interested in locating low-skill, high-skill, and mixed profiles so that the distribution of the low-skill profile across offerings could be examined. OST settings serving high concentrations of lower-SEL skill youth will likely require greater resources to achieve effects similar to settings serving higher-skilled youth.

To create the skill profiles, we subjected the composite scores to cluster analyses using the ROPstat (version 2.0) statistical package for pattern-oriented analyses (Vargha, Torma, & Bergman, 2015). After using ROPstat modules for addressing missing data and multivariate outliers, we used Ward's method, with squared Euclidian distances as the dissimilarity measure, to identify clusters followed by *k*-means relocation analyses. For each successively more complex cluster solution, an index

of the decrease in the error sums of squares was produced (ESS) and entered into a scree-type plot to determine the statistically-justifiable upper and lower number of cluster groups that provided unique information (Bergman et al., 2003).

Chapter Four. Performance Study Results

The performance studies were implemented at each of the eight SEL Challenge organizations with a focus on the exemplary offering at each site. One important purpose for the performance studies was to provide validation evidence for exemplary status of the Challenge offerings. If the offerings were indeed of exceptionally high quality and effectiveness, there was more reason to describe the standards as promising practices (see Endnote i). Second, we wanted to take an opportunity to assess the feasibility, reliability, and validity of the suite of SEL performance measures to better inform and assure future users. Finally, we wanted to develop empirically grounded performance benchmarks so that readers could decide how much the OST intervention delivered through the eight SEL Challenge offerings was similar to, or different from, their own work. The performance measures used to produce these results are described in Table 6 and Appendix D.

Comparison Group and Benchmarks

Because the quality and effectiveness of services can only be defined in comparison to other conditions and samples, we assembled a comparison group from other high school and middle school OST offerings.[xxi] The comparison group sample for this project included a wide mix of OST programs representing community-based organizations, national youth-serving organizations, and OST systems organized by state departments of education and state departments of human services. Because this reference sample is drawn from programs that are implementing the continuous improvement intervention over multiple cycles, it is possible that they are already operating at relatively higher levels of quality.

We refer to eleven of the performance measures used in the Challenge as benchmark measures because they are highly aligned with important standards, curriculum features, and youth SEL skills identified in the Challenge study. The benchmark describes the type and level of performance in SEL offerings for others to use as a goal or an evaluative comparison. The benchmarks are defined by the range of the measurement scale and performance levels demonstrated by Challenge offerings on each benchmark measure. The Challenge offerings were consistent in performance across the setting measures and there were few statistically or substantively significant differences, so the bottom cut point for the benchmark range was set at one standard deviation below the mean.

Of the 25 performance measures described in this Chapter, eleven were selected as benchmark measures: horizontal communication and vertical communication are two measures of organizational culture that reflect collaborative practices between management and staff in the organization, including the shared knowledge of the offering curriculum and presence in the offering setting. Growth and mastery skills reflect the complexity and degree of challenge available through the project curriculum. The

instructional total score describes adherence to the standards for SEL practice, including the responsive practices. Youth engagement describes mental engagement or disengagement with the context. Six youth behavior composite scores indicate the presence or absence of skill growth. Finally, as described in Chapter One, several structural features are good candidates for benchmarks as well. These include the intensity of the offering in terms of the number of sessions and contact hours, the diversity of youth participants, and the ratio of staff to youth.

For the setting level benchmark measures, we combined the information from the comparison group and the benchmarks by reporting the proportion of sites in the comparison sample that are performing within the range of performance for the Challenge offerings, i.e., the proportion of comparison sites performing in the benchmarked range.

Results for Setting Measures

There are 12 performance measures at the organization and point-of-service levels of setting. For each set of measures, we present a bar chart to describe the mean for the eight Challenge offerings and the mean for the comparison sample. Following the presentation of means for the point-of-service level measures, we also present means for staff practices aligned to the six domains in the standards for SEL practice using 68 items from the instructional total score. These data were produced using an observation of staff performance on three occasions. Finally, we present the proportion of comparison group sites with a mean score falling within the benchmark range for the exemplary offerings.

Organization Setting

The organizational setting is where management practices are implemented as managers and staff interact to make decisions about how to produce the SEL service—at staff meetings, during professional development, in one-on-one meetings and conversations. The organizational setting measures include accountability,[3] staffing model, horizontal communication, vertical communication, job satisfaction, workload, and school day content. Together, horizontal and vertical communication represent culture (how collaborative your work practices are), and job satisfaction and workload represent climate (how you feel about work). Figure 4 compares the average scores for each construct against the de-identified reference sample.

SEL organizations showed similar patterns of results with the comparison group on accountability and staffing capacity, and an inverted pattern for school day content, as many of them were not directly

[3] Accountability is actually a system-level measure, but we include it with the organization level measures to save space.

targeting participants' academic performance. Challenge offerings were slightly higher on the organizational climate measures job satisfaction and workload. Vertical communication and horizontal communication demonstrated substantially higher performance than the comparison group.

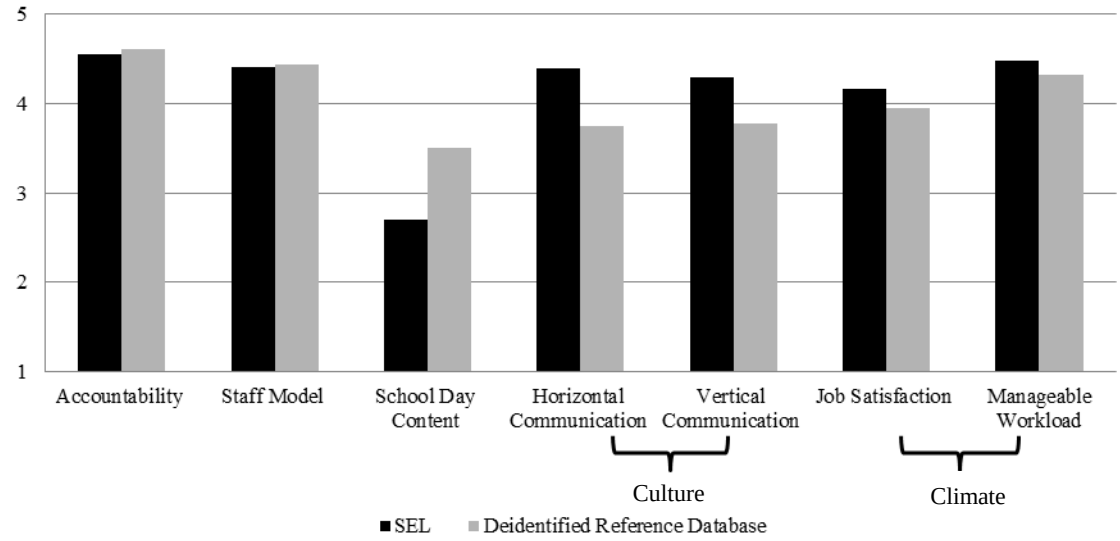

Figure 4. **Organization Setting Performance Profile Point-of-Service Level Setting**

Point-of-service setting is wherever staff delivers curriculum and instruction with youth. This is where the quality of instructional practice facilitates youth engagement. The point-of-service level measures include curriculum planning, youth governance, growth and mastery skills, instructional total score, and youth engagement. Figure 5 compares the average scores for each construct against the de-identified reference sample.

Curriculum planning scores for the exemplary offerings were similar to the comparison sample, while youth governance scores were substantially higher. Growth and mastery skills, instructional total score, and youth engagement demonstrated substantially higher performance compared to the comparison sample.

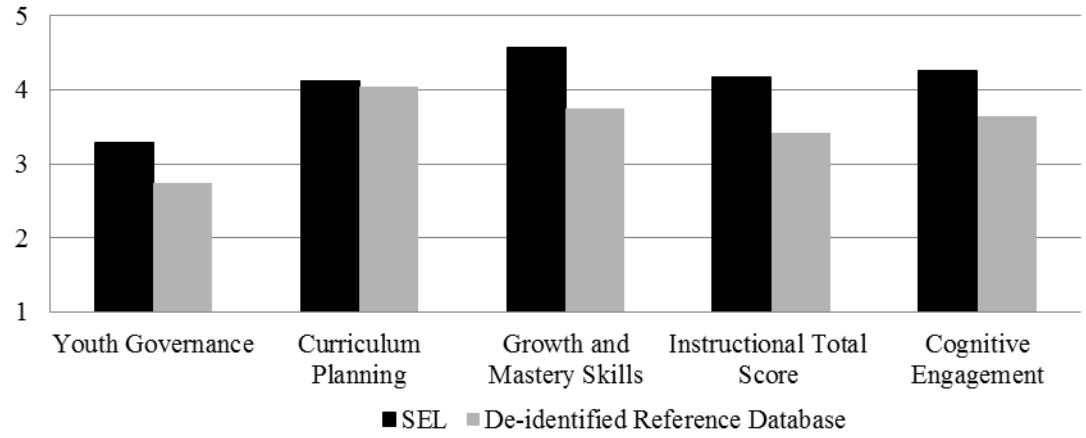

Figure 5. **Point-of-Service Setting Performance Profile**

To provide performance information about staff SEL practices in each domain, the items from the observational measure of staff behavior—instructional total score—were aligned with the practice indicators in the standards. The alignment of observational items to the six SEL domains is not perfect, with a few significant gaps in content in the emotion management and empathy domains. The alignment is presented in Appendix E. Figure 6 presents the scores for the exemplary programs and the comparison sample. The exemplary offerings substantially outperformed the comparison sample in all domains of SEL practice.

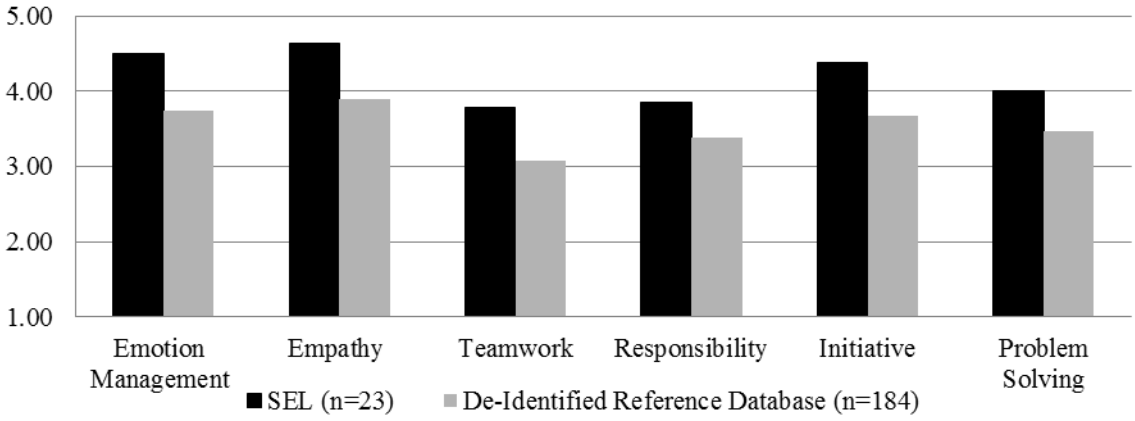

Figure 6. **Social and Emotional Program Quality Assessment by SEL Domain**

Benchmarks

There were five benchmark indicators for the organization and point-of-service levels of setting, indicators that were especially well aligned with the SEL standards and curriculum features. Table 8 presents the proportion of the comparison group that is performing in the range of the exemplary offerings. On average, 44% of the comparison offerings are performing within the benchmarked range, that is, had a mean score greater than the one standard deviation below the mean of the Challenge offerings. However, only 17% of comparison group sites provided a similarly intensive project curriculum as evidenced by the growth and mastery skills measure, and only 33% of comparison group sites implemented staff SEL practices at a similar level to the Challenge offerings.

Table 8. Proportion of Comparison Sites in Exemplary Performance Range

Horizontal Communication	*Vertical Communication*	*Growth and Mastery Skills*	*Instructional Total Score*	*Youth Engagement*
50%	69%	17%	33%	49%

Results for Youth SEL Skills

SEL skills were measured using belief and behavior measures in each of the six domains at three time points. This section presents results for overall SEL skill change, SEL skill change for higher risk youth, and SEL skill profiles.

Overall SEL Skill Change

The domain composite scores (see discussion of composite scores in Chapter Three) for each SEL domain at each time point are presented in Figures 7 and 8. Results for multilevel latent construct growth models summarized in Appendix E confirm the general pattern of results. SEL skills demonstrate positive change for both types of measures in all domains across the three time points, although not all changes were statistically significant for the belief measures.

Importantly, growth trajectories for mental processes related to self-regulation of attention (problem solving strategies) and emotion (emotional reappraisal) were positive and statistically significant, indicating growth in skills that youth use to optimize skill learning in all domains.

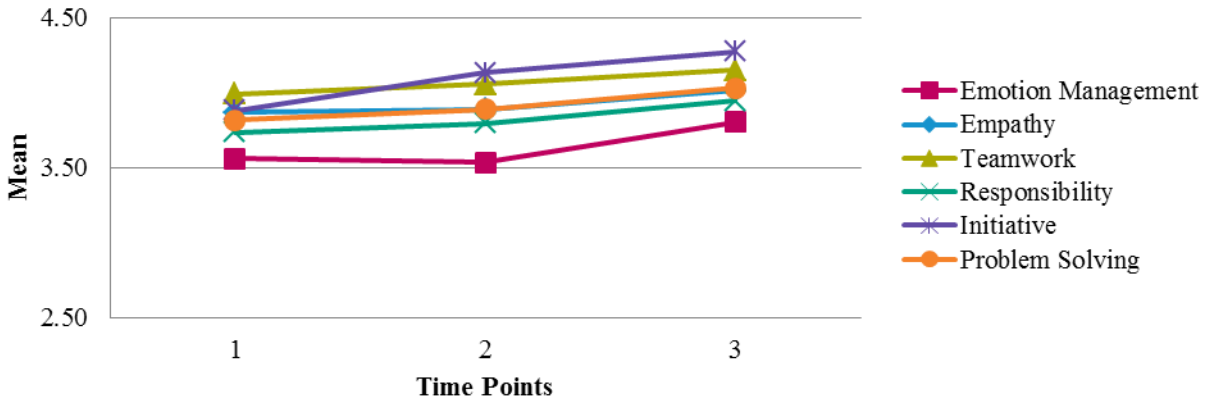

Figure 7. **SEL Belief Change**

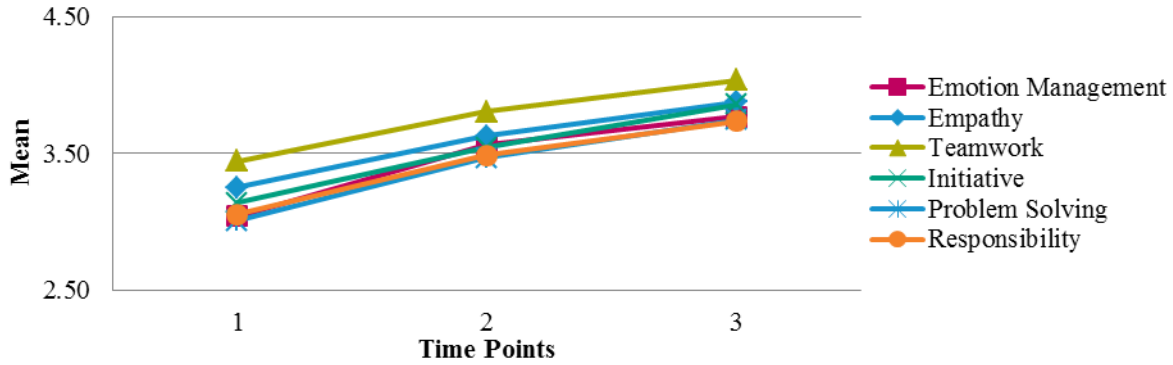

Figure 8. **SEL Behavior Change**

To provide a shorthand comparison of the amount of change in beliefs and behaviors, respectively, Figure 9 presents Cohen's d-type effects sizes—calculated by subtracting the Time 1 mean from the Time 3 mean and dividing by the standard deviation of the Time 1 mean—for the belief and behavior measures in each domain. In keeping with the theory for SEL skill measures outlined in Chapter Three, the OST intervention had greater effects on SEL behaviors demonstrated during the offering than on youth's broadly construed beliefs about the efficacy of their own SEL skills.

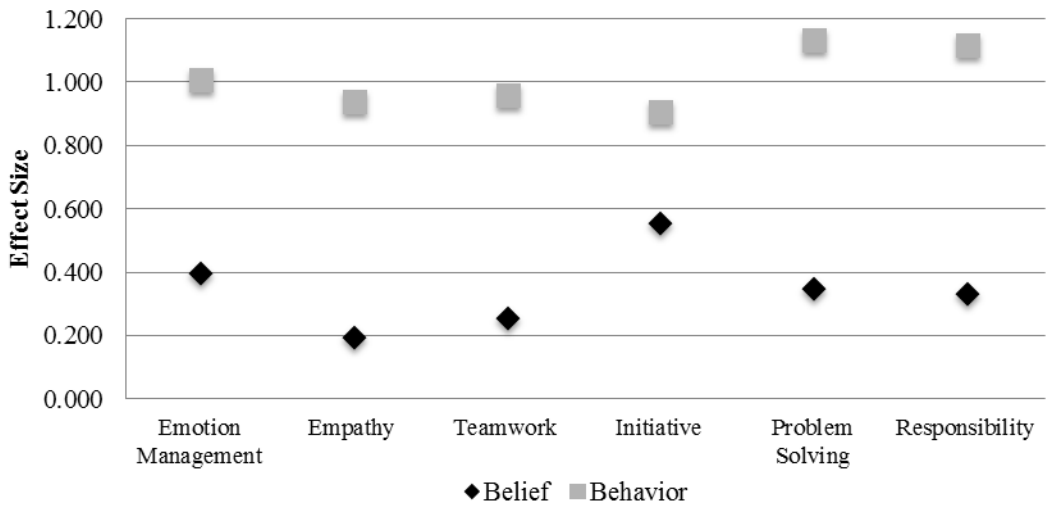

Figure 9. **Effect Sizes for Changes in SEL Belief and Behavior Measures from Time 1 to Time 3**

Higher Risk Youth

Risk was assessed at baseline, and we used models to test for differences in skill growth, comparing trajectories for higher risk groups to the rest of the youth. Risk was indicated in two ways: (a) membership in a low-skill belief or behavior subgroup at baseline, and (b) self-reported attachment anxiety, attachment avoidance, or social phobia.

The low-skill subgroups were defined by membership in the low belief or behavior profiles shown in Figures 10 and 11. As presented in Appendix D, Figures D-1 through D-12, youth in the low-skill subgroups demonstrated as much or more change over time than youth in the higher-skill subgroups. Results from growth models also showed that the intercept-slope correlation was negative in all cases, indicating that lower starting points were associated with greater gains. These findings do not rule out improvements due to regression to the mean (i.e., the higher odds that low-scoring youth at baseline will score higher at subsequent time points due to misclassification as low at baseline).

The second approach was based on a risk index constructed from measures of attachment-related anxiety, avoidance, and social phobia. All youth with mean scores in the negative part of the response scale (i.e., disagree) were scored as having the risk indicator, and these indicators were summed to create an index ranging between 0 (low risk) and 2 (high risk). In a variety of models, youth with higher baseline levels of attachment gained just as much, and occasionally more, than their lower-risk peers.

SEL Skill Profiles

Finally, in Figures 10 and 11, we present results from cluster analyses for the belief and behavior measures. Several facets of these profiles are of interest. First, these figures describe SEL skill subgroups of youth at baseline. Approximately 17% of youth began the offerings in the low-beliefs profile (Figure 10, Cluster 1) and another 16% were in a profile at the same low level on four of the six domains (Figure 10, Cluster 4). These students held beliefs that they were generally unsuccessful with the many SEL skills named in the youth survey items. For the staff ratings of youth behavior in the offering, approximately 25% of youth began the offering in the low profile (Figure 11, Cluster 3). According to teachers, these students generally failed to demonstrate an SEL skill when given the opportunity.

Additional information from the examination of the skill profiles at Time 3 (not presented) suggests that (a) the low profiles disappeared between the first and third time points; (b) most students moved to a next higher skill profile, indicating incremental growth and stability in skills (e.g., youth who started out in the middle-skill group ended up in the higher-middle-skill group); and, finally, (c) the shapes of the skill profiles (i.e., the relative levels of the six domain scores within each cluster) were maintained from baseline to the final time point. This last point indicates that the skills changed in amount but not in type during the SEL offering.

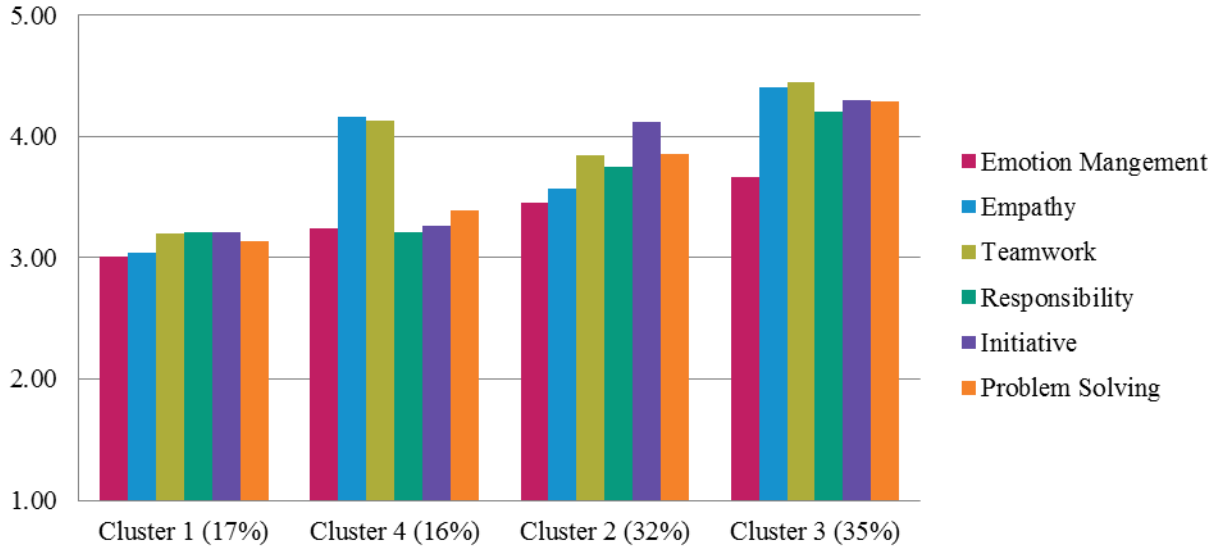

Figure 10. **Belief skill profiles Time 1**

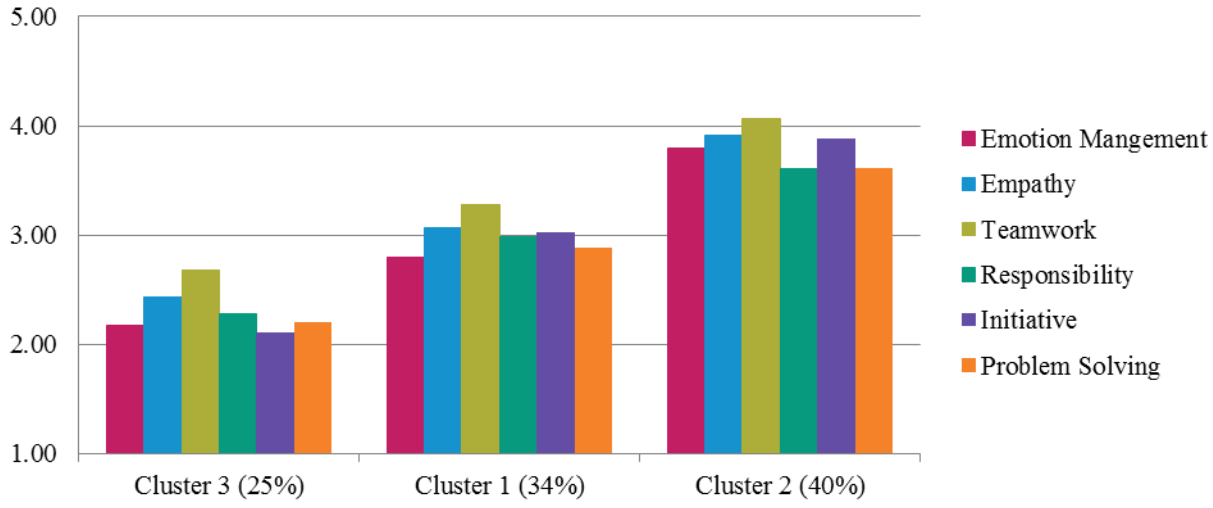

Figure 11. **Behavior skill profiles at Time 1**

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 settings[4]. 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.

[4] 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.

References

Ajzen, I. (2005). Laws of human behavior: Symmetry, compatibility, and attitude-behavior correspondence. In A. Beauducel, B. Biehl, M. Bosniak, W. Conrad, G. Schönberger, & D. Wagener (Eds.), *Multivariate research strategies* (pp. 3-19). Aachen, Germany: Shaker Verlag.

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84*(5), 888.

Akiva, T., Cortina, K. S., Eccles, J. S., & Smith, C. (2013). Youth belonging and cognitive engagement in organized activities: A large-scale field study. *Journal of Applied Developmental Psychology, 34*(5), 208-218.

Allen, J. P., & Philliber, S. (2001). Who benefits most from a broadly targeted prevention program? Differential efficacy across populations in the Teen Outreach Program. *Journal of Community Psychology*, *29*(6), 637-655.

Allen, J., Philliber, S., Herrling, S., & Kuperminc, G. P. (1997). Preventing teen pregnancy and academic failure: Experimental evaluation of a developmentally based approach. *Child Development, 68*(4), 729-742.

American Psychological Association. Coalition for Psychology in Schools and Education.(2015). *Top 20 principles from psychology for preK–12 teaching and learning*.

Anderson, N. B. (1998). Levels of analysis in health science: A framework for integrating sociobehavioral and biomedical research. *Annals of the New York Academy of Sciences, 840*(1), 563-576.

Apsler, R. (2009). After-school programs for adolescents: a review of evaluation research. *Adolescence, 44*(173), 1-19.

Arbeau, S. R. (2002). *Strategies of adolescent problem solving.* (MQ76366 M.N.), The University of New Brunswick (Canada), Canada. Retrieved from http://ezproxy.emich.edu/login?url=http://search.proquest.com/docview/305442060?accountid=10650 ProQuest Dissertations & Theses Full Text; ProQuest Nursing & Allied Health Source database.

Ayduk, Ö., Zayas, V., Downey, G., Cole, A. B., Shoda, Y., & Mischel, W. (2008). Rejection sensitivity and executive control: Joint predictors of borderline personality features. *Journal of Research in Personality, 42*(1), 151-168.

Bailey, D., Duncan, G., Odgers, C., & Yu, W. (2015). *Persistence and fadeout in the impacts of child and adolescent interventions* (No. 2015-27).

Baker, E. A. (2014). Learning and assessment: Twenty-first century skills and cognitive assessment. In H. F. O'Neil, R. S. Perez, & E. A. Baker (Eds.), *Teaching and measuring cognitive readiness*: Springer.

Baltes, P. B. (1997). On the incomplete architecture of human ontogeny: Selection, optimization, and compensation as foundations of developmental theory. *American Psychologist, 52,* 366-380.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191.

Bandura, A. (1992). Exercise of personal agency through the self-efficacy mechanisms. In R. Schwarzer (Ed.), *Self-efficacy: Thought control of action*. Washintgon, DC: Hemisphere.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. P. & T. U. Urdan (Eds.), *Self-Efficacy Beliefs of Adolescents.* (pp. 307-337): Information Age Publishing.

Bandura, A., Caprara, G. V., Barbaranelli, C., Gerbino, M., & Pastorelli, C. (2003). Role of affective self-regulatory efficacy in diverse spheres of psychosocial functioning. *Child Development, 74*(3), 769-782.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54*(7), 462.

Barker, R. G. (1968). *Ecological psychology: Concepts and methods for studying the environment of human behavior*. Standford, CA: Stanford University Press.

Bergman, L. R., Magnusson, D., & El-Khouri, B. M. (2003). *Studying individual development in an interindividual context: A person-oriented approach*. Mahwah, NJ: Erlbaum.

Berntson, G. G., Boysen, S. T., & Cacioppo, J. T. (1993). Neurobehavioral Organization and the Cardinal Principle of Evaluative Bivalencea. *Annals of the New York Academy of Sciences, 702*(1), 75-102.

Berntson, G. G., & Cacioppo, J. T. (2003). A contemporary perspective on multilevel analyses and social neuroscience. In F. Kessel, P. L. Rosenfeld, & N. B. Anderson (Eds.), *Expanding the boundaries of health and social science: Case studies in interdisciplinary innovation* New York, NY: Oxford University Press.

Berntson, G. G., & Cacioppo, J. T. (2008). The neuroevolution of motivation. In J. Shah & W. Gardner (Eds.), *Handbook of motvation science* (pp. 188-200). New York, NY: Guilford.

Blair, C. (2016). Developmental science and executive function. *Current Directions in Psychological Science, 25*(1), 3-7.

Blair, C., & Raver, C. C. (2012). Individual development and evolution: Experiential canalization of self-regulation. *Developmental Psychology, 48*(3), 647.

Blazevski, J., & Smith, C. (2007). Inter-rater reliability on the youth program quality assessment. Ypsilanti, MI: High/Scope Educational Research Foundation.

Blumberg, S. J., Carle, A. C., O'Connor, K. S., Moore, K. A., & Lippman, L. H. (2007). Social Competence: Development of an Indicator for Children and Adolescents. *Child Indicators Research, 1*(2), 176-197.

Bobko, P., Roth, P. L., & Buster, M. A. (2007). The usefulness of unit weights in creating composite scores: A literature review, application to content validity, and meta-analysis. *Organizational Research Methods, 10*(4), 689-709.

Bowden, D. M., & Martin, R. F. (1995). NeuroNames brain hierarchy. *Neuroimage, 2*(1), 63-83.

Boysen, S. T., & Berntson, G. G. (1995). Responses to quantity: perceptual versus cognitive mechanisms in chimpanzees (Pan troglodytes). *Journal of Experimental Psychology: Animal Behavior Processes, 21*(1), 82.

Bronson, G. (1965). The hierarchical organization of the central nervous system: Implications for learning processes and critical periods in early development. *Behavioral Science, 10*(1), 7-25.

Brown, A. L. (1992). Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. *Journal of the Learning Sciences, 2*, 141-178.

Bryck, R. L., & Fisher, P. A. (2012). Training the brain: practical applications of neural plasticity from the intersection of cognitive neuroscience, developmental psychology, and prevention science. *American Psychologist, 67*(2), 87.

Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, NH: Earlbaum.

Cacioppo, J. T., & Berntson, G. G. (1992). Social psychological contributions to the decade of the brain: Doctrine of multilevel analysis. *American Psychologist, 47*(8), 1019-1028.

Campbell, D. T. (1990). Levels of organization, downward causation, and the selection-theory approach to evolutionary epistemology. In G. Greenberg & E. Tobach (Eds.), *Theories of the evolution of knowing* (pp. 1-17). Hillsdale, NJ: Erlbaum.

Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., & Hawkins, J. D. (2004). Positive youth development in the united states: Research findings on evaluation of positive youth development programs. *Annals of the American Academy, 591*, 98-124.

Chappell, S., Nunnery, J., Pribesh, S., & Hager, J. (2011). A Meta-Analysis of Supplemental Educational Services (SES) Provider Effects on Student Achievement. *Journal of Education for Students Placed at Risk, 16*(1), 1-23.

Cicchetti, D., & Rogosch, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and psychology, 8*, 597-600.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284-290.

Collaborative for Academic Social and Emotional Learning. (2015). CASEL Guide: Effective Social and Emotional Learning Programs. Chicago, IL.

Corcoran, R. P., & Slavin, R. E. (2016). Effective Programs for Social and Emotional Learning (SEL): A Systematic Review. Oslo, Norway: The Campbell Collaboration.

Côté, J. E. (2000). *Arrested adulthood: The changing nature of maturity and identity*: NYU Press.

Côté, J. E., & Levine, C. G. (2002). *Identity formation, agency, and culture: A social psychological synthesis*. Mahwah, NJ: Erlbaum.

Cunha, F., & Heckman, J. J. (2006). *Investing in our young people*. University of Chicago, Department of Economics.

Curtis, W. J., & Cicchetti, D. (2007). Emotion and resilience: A multilevel investigation of hemispheric electroencephalogram asymmetry and emotion regulation in maltreated and nonmaltreated children. *Development and Psychopathology, 19*(3), 811-840.

J Davidson, R., Dunne, J., Eccles, J. S., Engle, A., Greenberg, M., Jennings, P., ... & Roeser, R. W. (2012). Contemplative practices and mental training: Prospects for American education. *Child Development Perspectives*, *6*(2), 146-153.

Davis, M. H. (1994). *Empathy: A social psychological approach*: Westview Press.

*Dawes, N. P., & Larson, R. (2011). How youth get engaged: grounded-theory research on motivational development in organized youth programs. *Developmental Psychology, 47*(1), 259-269.

Derryberry, D., & Tucker, D. M. (1991). *The adaptive base of the neural hierarchy: Elementary motivational controls on network function.* Paper presented at the Nebraska symposium on motivation.

Deutsch, R., & Strack, F. (2006). Duality models in social psychology: From dual processes to interacting systems. *Psychological Inquiry*, *17*(3), 166-172.

Dewey, J. (1938). *Education and Experience*. New York: Simon and Schuster.

Dodge, K. A. (2011). Context matters in child and family policy. *Child Development, 82*(1), 433-442.

Duckworth, A. L., & Quinn, P. D. (2009). Development and Validation of the Short Grit Scale (Grit–S). *Journal of personality assessment, 91*(2), 166-174.

Durlak, J. (2015). *Handbook of social and emotional learning: Research and practice*. New York, NY: The Guilford Press.

Durlak, J., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implemetation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327-350.

Durlak, J. A., & Weissberg, R. P. (2007). The Impact of After-School Programs that Promote Personal and Social Skills. *Collaborative for Academic, Social, and Emotional Learning (NJ1)*.

Durlak, J., & Weissberg, R. (2010). A Meta-Analysis of After-School Programs That Seek to Promote Personal and Social Skills in Children and Adolescents. *American Journal of Community Psychology, 45*(3-4), 294-309.

Durlak, J., 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*(1), 405-432.

Durlak, J., & Wells, A. M. (1997). Primary Prevention Mental Health Programs for Children and Adolescents: A Meta-Analytic Review. *American Journal of Community Psychology, 25*(2), 115-151.

*Dworkin, J., Larson, R. W., & Hansen, D. (2003). Adolescents' accounts of growth experiences in youth activities. *Journal of Youth and Adolescents, 32*(1), 17-26.

Eccles, J., & Gootman, J. (Eds.). (2002). *Community programs to promote youth development*. Washington, DC: National Academy Press.

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2007). Prosocial Development *Handbook of Child Psychology*: John Wiley & Sons, Inc.

Epstein, M. H., Mooney, P., Ryser, G., & Pierce, C. D. (2004). Validity and Reliability of the Behavioral and Emotional Rating Scale (2nd Edition): Youth Rating Scale. *Research on Social Work Practice, 14*(5), 358-367.

Epstein, S. (1985). The Implications of Cognitive-experiential Self-theory for Research in Social Psychology and Personality1. *Journal for the Theory of Social Behaviour, 15*(3), 283-310.

Epstein, S. (1990). Cognitive-experiential self-theory. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research*. Ne York, NY: Guilford.

Evans, G. W., & Fuller-Rowell, T. E. (2013). Childhood poverty, chronic stress, and young adult working memory: the protective role of self-regulatory capacity. *Developmental science, 16*(5), 688-696.

Farrington, C. A., Roderick, M., Johnson, D. W., Keyes, T. S., Allensworth, E., Nagaoka, J., & Beechum, N. O. (2012). Teaching Adolescents to Become Learners: The Role of Noncognitive Factors in Shaping School Performance. *The University of Chicago Consortium on Chicago School Research*.

Feldman Farb, A., & Matjasko, J. L. (2012). Recent advances in research on school-based extracurricular activities and adolescent development. *Developmental Review, 32*, 1-48.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *American Psychological Association, 87*(6), 477-531.

Fischer, K. W., & Bidell, T. R. (2006). Dynamic development of action and thought. In W. Damon & R. M. Lerner (Eds.), *Theoretical models of human development. Handbook of child psychology* (6 ed., Vol. 1, pp. 313-399). New York: Wiley.

Fischer, K. W., & Van Geert, P. (2014). Dynamic development of brain and behavior. In P. Molenaar, R. Lerner, & K. Newell (Eds.), *Handbook of developmental systems theory and methodology* (pp. 287-315). New York, NY: The Guilford Press.

Flannelly, K. J., Koenig, H. G., Galek, K., & Ellison, C. G. (2007). Beliefs, mental health, and evolutionary threat assessment systems in the brain. *The Journal of nervous and mental disease, 195*(12), 996-1003.

Ford, D. H., & Lerner, R. M. (1992). *Developmental systems theory: An integrative approach*. Newbury Park, CA: Sage Publications, Inc.

Fralicx, R. D., & Raju, N. S. (1982). A comparison of five methods for combining multiple criteria into a single composite. *Educational and Psychological Measurement, 42*, 823-827.

Gideonse, S. (1988). *Determinants of Youth's Successful Entry Into Adulthood*. William T. Grant Foundation Commission on Work, Family and Citizenship.

Gottfredson, D. C., & Wilson, D. B. (2003). Characteristics of Effective School-Based Substance Abuse Prevention. *Prevention Science, 4*(1), 27-38.

Gratz, K. L., & Roemer, L. (2004). Multidimensional Assessment of Emotion Regulation and Dysregulation: Development, Factor Structure, and Initial Validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment, 26*(1), 41-54.

Grene, M. (1988). Hierarchies and behavior. In G. Greenberg & E. Tobach (Eds.), *Evolution of social behavior and integrative levels* (Vol. 3-17). Hillsdale, NJ: Erlbaum.

Gresham, F. M., & Elliot, S. N. (1990). Manual for the social skills rating system. *Circle Pines, MN: American Guidance Service*.

Gross, J. J. (2013). *Handbook of Emotion Regulation* (2nd ed.). New York, NY: The Guilford Press.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology Bulletin, 85*(2), 348-362.

Halpern-Felsher, B. (2009). Adolescent decision making: an overview. *The Prevention Researcher, 16*(2), 3+.

*Hansen, D. M., Larson, R., & Dworkin, J. B. (2003). *What adolescents learn in organized youth activities: A survey of self-reported developmental experiences*. Journal of Research on Adolescence. University of Illinois at Urbana-Champaign.

Hattie, J. (2009). *Visible learning : a synthesis of over 800 meta-analyses relating to achievement*. London, NY: Routledge.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of educational research, 77*(1), 81-112.

Heckman, J. J. (2008). *Schools, Skills, and Synapses*. National Institute of Health: Public Access.

Heckman, J. J. (2013). *Giving kids a fair chance*. Cambridge, MA: Massachusetts Institute of Technology Press.

Herrick, C. J. (1949). A biological survey of integrative levels. In W. T. Parry, R. W. Sellars, V. J. McGill, & M. Farber (Eds.), *Philosophy for the Future: The Quest of Modern Materialism* (pp. 222-242). New York, NY: MacMillan.

Hiebert, J., Gallimore, R., & Stigler, J. W. (2002). A knowledge base for the teaching profession: What would it look like and how can we get one? *Educational Researcher, 31*(5), 3.

Higgins, E. T., & King, G. A. (1981). Accessibility of social constructs: Information-processing consequences of individual and contextual variability. In N. Cantor & J. F. Kihlstrom (Eds.), *Personality, cognition and social interaction* (pp. 69-121). New York: Erlbaum.

Hoffman, D. M. (2009). Reflecting on Social Emotional Learning: A Critical Perspective on Trends in the United States. *Review of educational research, 79*(2), 533-556.

Hogan, R., Johnson, J. A., & Briggs, S. R. (1997). *Handbook of personality psychology*: Elsevier.

Holland, J. H. (1995). *Hidden order: How adaptation builds complexity*. New York, NY: Addison-Wesley.

Holopainen, L., Lappalainen, K., Junttila, N., & Savolainen, H. (2011). The Role of Social Competence in the Psychological Well-being of Adolescents in Secondary Education. *Scandinavian Journal of Educational Research, 56*(2), 199-212.

Jackson, J. H. (1884). Croonian lectures on the evolution and dissolution of the nervous system: Lecture II. . *The Lancet*, 649-652.

Jaffee, S. R., & Christian, C. W. (2014). The biological embedding of child abuse and neglect. *Social Policy Report, 28*(1), 1-36.

James, W. (1890). *The principles of psychology*. New York, NY: Holt and Company.

Jennings, P. A., Frank, J. L., Snowberg, K. E., Coccia, M. A., & Greenberg, M. T. (2013). Improving classroom learning environments by Cultivating Awareness and Resilience in Education (CARE): Results of a randomized controlled trial. *School Psychology Quarterly, 28*(4), 374.

Jones, S. M., & Bouffard, S. M. (2012). Social and Emotional Learning in Schools: From Programs to Strategies. Social Policy Report. Volume 26, Number 4. *Society for Research in Child Development*.

Jung, C. G. (1923). *Psychological types: or the psychology of individuation* (H. G. Baynes, Trans.). Oxford, England: Harcourt Brace.

Kahneman, D., & Frederick, S. (2007). Frames and brains: Elicitation and control of response tendencies. *Trends in Cognitive Sciences, 11*(2), 45-46.

Kochanska, G., & Knaack, A. (2003). Effortful control as a personality characteristic of young children: Antecedents, correlates, and consequences. *Journal of Personality, 71*(6), 1087-1112.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

*Larson, R. W. (2000). Toward a psychology of positive youth development. *American Psychologist, 55*(5), 170-183.

*Larson, R. W. (2011). Positive development in a disorderly world. *Journal of Research on Adolescence, 21*(2), 317-334.

*Larson, R. W., & Angus, R. M. (2011). Adolescents' development of skills for agency in youth programs: Learning to think strategically. *Child Development, 82*, 277-294.

*Larson, R. W., & Brown, J. R. (2007). Emotional development in adolescence: What can be learned from a high school theater program? *Child Development, 78*(4), 1083-1099.

*Larson, R. W., & Hansen, D. (2005). The development of strategic thinking: Learning to impact human systems in a youth activism program. *Human Development, 48*, 327-349.

*Larson, R. W., Hansen, D., & Moneta, G. (2006). Differing profiles of developmental experiences across types of organized youth activities. *Developmental Psychology, 42*, 849-863.

Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-School-Time Programs: A Meta-Analysis of Effects for At-Risk Students. *Review of educational research, 76*(2), 275-313.

Lavelli, M., Pantoja, A. P. F., Hsu, H., Messinger, D., & Fogel, A. (2008). *Using Microgenetic Designs to Study Change Process*.

Lazarus, R. S. (1991). *Emotion and Adaptation*. New York, NY: Oxford University Press.

Leech, N. L., & Onwuegbuzie, A. J. (2009). A typology of mixed methods research designs. *Quality & Quantity, 43*(2), 265-275.

Lewis, M. D., & Todd, R. M. (2007). The self-regulating brain: Cortical-subcortical feedback and the development of intelligent action. *Cognitive Development, 22*(4), 406-430.

Li, J., & Julian, M. M. (2012). Developmental relationship as the active ingredient: A unifying working hypothesis of "what works" across intervention settings. *American Journal of Orthopsychiatry, 82*(2), 157-166.

Lieberman, M. D., Gaunt, R., Gilbert, D. T., & Trope, Y. (2002). Reflection and reflexion: A social cognitive neuroscience approach to attributional inference. In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 199-249). New York, NY: Academic Press.

Liem, G., & Martin, A. J. (2012). The Motivation and Engagement Scale: Theoretical Framework, Psychometric Properties, and Applied Yields. *Australian Psychologist, 47*(1), 3-13.

Lippman, L. H., Moore, K. A., Guzman, L., Ryberg, R., McIntosh, H., Ramos, M. F., . . . Kuhfeld, M. (2014). *Flourishing children: Defining and testing indicators of positive development*. New York, NY: Springer.

Lippman, L. H., Moore, K. A., & McIntosh, H. (2011). Positive indicators of child well-being: A conceptual framework, measures, and methodological issues. *Applied Research in Quality of Life*, *6*(4), 425-449.

Lippman, L. H., Ryberg, R., Terzian, M., Moore, K. A., Humble, J., & McIntosh, H. (2014). Positive and protective factors in adolescent well-being *Handbook of Child Well-Being: Theories, Methods and Policies in Global Perspective* (pp. 2823-2866). Netherlands: Springer Netherlands.

Lipsey, M. W., Howell, J. C., Kelly, M. R., Chapman, G., & Carver, D. (2010). Improving the effectiveness of juvenile justice programs. *Washington, DC: Georgetown Center for Juvenile Justice Reform*.

Lovelace, V., Eilenfeld, A., & Francese, M. (2014). Resilience Reimagined: Mindfulness Practices in the Classroom. *Educator's Voice*, *7*, 2-11.

MacLean, P. D. (1990). *The triune brain in evolution: Role in paleocerebral functions*. Springer Science & Business Media.

Martin, J. E. (2016). Social and Emotional Learning and Assessment: The Demand is Clear. Retrieved from http://gettingsmart.com/2016/05/social-and-emotional-learning-and-assessment-the-demand-is-clear/

Marzano, R. J. (1998). A theory-based meta-analysis of research on instruction. Aurora, CO: Mid-continent Regional Educational Laboratory.

Marzano, R. J. (2004). *Building background knowledge for academic achievement: Research on what works in schools*. Alexandria, VA: Association for Supervision and Curriculum Development.

Maxwell, J. P. (1989). Mediation in the schools: Self-regulation, self-esteem, and self-discipline. *Conflict Resolution Quarterly, 7*, 149-155.

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications* (pp. 3-31). New York: Basic Books.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2004). Emotional Intelligence: Theory, Findings, and Implications. *Psychological Inquiry, 15*(3), 197-215.

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review, 96*(4), 690-702.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*(3), 419-457.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews: Neuroscience, 4*(4), 310-322.

McEvers, K. (Writer). (2016). MacArthur 'Genius' Angela Duckworth Responds to a New Critique of Grit, *All Things Considered*. Washington, DC: National Public Radio.

Merton, R. K. (1949). *On sociological theories of the middle range [1949]*: na.

Messick, S. (1995). Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741.

Mesulam, M. (2000). *Principles of behavioral and cognitive neurology*. Oxford, England: Oxford University Press.

Miettinen, R. (2000). The concept of experiential learning and John Dewey's theory of reflective thought and action. *International Journal of Lifelong Education, 19*(1), 54-72.

Miller, B., & Surr, W. (2007). Survey of After-School Youth Outcomes (SAYO): Wellesley Centers for Women, Wellesley College, Massachusetts Department of Education.

Murphey, D., Stratford, B., Gooze, R., Bringwald, E., Cooper, P. M., Carney, R., & Rojas, A. (2014). Are the Children Well? A Model and Recommendations for Promoting the Mental Wellness of the Nation's Young People. Bethesda, MD: Child Trends.

Murray, D. W., Rosanbalm, K., Christopoulos, C., & Hamoudi, A. (2015). Self-Regulation and Toxic Stress: Foundations for Understanding Self-Regulation from an Applied Developmental Perspective. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Naftzger, N. (2012). Using Rasch Modeling Techniques to Address Issues of Reliability and Validity in Observational Measures: A Case Study Employing the School-Age Youth Program Quality Assessment. William T. Grant Foundation.

Naftzger, N. (2014). *A Summary of Three Studies Exploring the Relationship between Afterschool Program Quality and Youth Outcomes*. Paper presented at the Ready by 21 National Meeting, Covington, KY.

Naftzger, N., Devaney, E., & Foley, K. (2014). Summary of Analyses Related to Nashville After Zone Alliance Program Outcomes. Chicago, IL: American Institutes for Research.

Naftzger, N., Hallberg, K., & Tang, Y. (2014). Exploring the relationships between afterschool program quality and youth outcomes: Summary of findings from the Palm Beach County quality improvement system.*Naperville, IL: American Institutes for Research*.

Naftzger, N., Manzeske, D., Nistler, M., Swanlund, A., Rapaport, A., Shields, J., . . . Sugar, S. (2013). Texas 21st Century Community Learning Centers: Final Evaluation Report. *Naperville, IL: American Institutes for Research.*

Nagaoka, J., Farrington, C. A., Ehrlich, S. B., & Heath, R. D. (2015). Foundations for Young Adult Success: A Developmental Framework. Concept Paper for Research and Practice. *University of Chicago Consortium on Chicago School Research*.

Neil, J. T. (2003). Reviewing and benchmarking adventure therapy outcomes: Applications of Meta - Analysis. *The Journal of Experiential Education, 25*(3), 316-321.

O'Connell, M. E., Boat, T., & Warner, K. E. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*: National Academies Press.

O'Neil, H. F., Perez, R. S., & Baker, E. L. (2014). *Teaching and measuring cognitive readiness*: Springer.

Oden, S., Kelly, M. A., Ma, Z., & Weikart, D. P. (1992). *Challenging the potential: Programs for talented disadvantaged youth*. Ypsilanti, MI: High/Scope Press.

Oosterhof, A., Rohani, F., Sanfilippo, C., Stillwell, P., & Hawkins, K. (2008). The capabilities-complexity model. Florida State University.

Pattee, H. H. (1973a). The physical basis and origin of hierarchical control. In H. H. Pattee (Ed.), *Hierarchy Theory: The Challenge of Complex Systems*. New York, NY: Braziller.

Pattee, H. H. (1973b). Unsolved problems and potential applictions of hierarchy theory. In H. Simon & H. H. Pattee (Eds.), *Hierarchy Theory: The Challenge of Complex Systems* (pp. 129-156). New York, NY: Braziller.

*Pearce, N. J., & Larson, R. W. (2010). How teens become engaged in youth development programs: The process of motivational change in a civic activism organization. *Applied Developmental Science, 10*, 121-131.

Peck, S. C. (2005). *TEMPEST in a gallimaufry: Applying hierarchical systems theory to person-in-context research.* Paper presented at the Symposium conducted at the Society for Personality and Social Psychology Conference, New Orleans, LA.

Peck, S. C. (2006). *TEMPEST in a gallimaufry II: Using multilevel systems theory to integrate "self" and "identity" theories.* Paper presented at the Symposium conducted at the 4th International Biennial SELF Research Conference, Ann Arbor, MI.

Peck, S. C. (2007). TEMPEST in Gallimaufry: Applying Multilevel Systems Theory to Person-in-Context Research. *Journal of Personality*, 75, 1127-1156 [Special Issue: B. W. Roberts (Ed.), Contextualized identities: Integrating self-in-context to traditional issues in personality psychology].

Peck, S. C. (2009). Using multilevel systems theory to integrate dynamic person-in-context systems. *Journal of Research in Personality, 43*, 262-263. [Invited essay; Special Issue: M. B. Donnellan (Ed.), On the legacy of the person-situation debate and the future of person-situation integration in personality psychology].

Piaget, J. (1977). *The development of thought:equilibration of cognitive structures*. New York: Viking Press.

Platt, J. J., Spivack, G., Altman, N., Altman, D., & Peizer, S. B. (1974). Adolescent problem-solving thinking. *Journal of Consulting and Clinical Psychology, 42*(6), 787-793.

Porath-Waller, A. J., Beasley, E., & Beirness, D. (2010). A Meta-Analytic Review of School-Based Prevention for Cannabis Use. *HEALTH EDUCATION & BEHAVIOR, 37*(5), 709-723.

Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology, 12*(03), 427-441.

Putnam, R. D. (2015). *Our kids: The American dream in crisis*: Simon and Schuster.

Richters, J. E. (1997). The Hubble hypothesis and the developmentalist's dilemma *Development and Psychopathology* (Vol. 9, pp. 193-229): Cambridge University Press.

Rodriguez, M. L., Mischel, W., & Shoda, Y. (1989). Cognitive person variables in the delay of gratification of older children at risk. *Journal of Personality and Social Psychology, 57*(2), 358.

Roeser, R. W. (2016). Processes of teaching, learning, and transfer in mindfulness-based interventions (MBIs) for teachers: A contemplative educational perspective. In K. A. Schonert-Reichl & R. W. Roeser (Eds.), *Handbook of mindfulness in education: Integrating theory and research into practice* (pp. 149-170). New York, NY: Springer.

Roeser, R. W., & Peck, S. C. (2009). An education in awareness: Self, motivation, and self-regulated learning in contemplative perspective. *Educational Psychologist, 44*(2), 119-136.

Roeser, R. W., Peck, S. C., & Nasir, N. S. (2006). Self and identity processes in school motivation, learning, and achievement. In P. A. Alexander & P. H. Winne (Eds.), *HANDBOOK OF EDUCATIONAL PSYCHOLOGY* (pp. 391-424). Mahwah, NJ: Erlbaum.

Roeser, R. W., Peck, S. C., Nasir, N. S., Alexander, P. A., & Winne, P. H. (2006). Self and identity processes in school motivation, learning, and achievement *HANDBOOK OF EDUCATIONAL PSYCHOLOGY* (pp. 391-424). Mahwah, N.J.: Lawrence Erlbaum.

Roeser, R. W., Schonert-Reichl, K. A., Jha, A., Cullen, M., Wallace, L., Wilensky, R., . . . Harrison, J. (2013). Mindfulness training and reductions in teacher stress and burnout: Results from two randomized, waitlist-control field trials. *Journal of Educational Psychology, 105*(3), 787.

Rothbart, M. K. (2007). Temperament, development, and personality. *Current Directions in Psychological Science, 16*(4), 207-212.

*Rusk, N., Larson, R. W., Raffaelli, M., Walker, K., Washington, L., Gutierrez, V., . . . Cole Perry, S. (2013). Positive Youth Development in Organized Programs: How Teens Learn to Manage Emotions. In C. Proctor & P. A. Linley (Eds.), *Research, Applications, and Interventions for Children and Adolescents: A Positive Psychology Perspective* (pp. 247-261): Springer.

Salovey, P., & Mayer, J. D. (1990). Emotional Intelligence. *Imagination, Cognition and Personality, 9*(3), 185-210.

Salthe, S. N. (1985). *Evolving hierarchical systems: their structure and representation*. New York, NY: Columbia University Press.

Salthe, S. N. (1993). *Development and evolution: complexity and change in biology*. Cambridge, MA: Massachusetts Institute of Technology Press.

*Salusky, I., Larson, R. W., Griffith, A., Wu, J., Raffaelli, M., Sugimur, N., & Guzman, M. (2014). How Adolescents Develop Responsibility: What Can Be Learned From Youth Programs. *Journal of Research on Adolescence, 24*(3), 417-430.

Sambrano, S., Springer, J. F., Sale, E., Kasim, R., & Hermann, J. (2005). Understanding prevention effectiveness in real-world settings: the National Cross-Site Evaluation of high risk youth programs. *The American journal of drug and alcohol abuse, 31*(3), 491-513.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal for Experimental Psychology: Learning, Memory, & Cognition, 13*, 501-518.

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the life orientation test. *Journal of Personality and Social Psychology, 67*(6), 1063-1078.

Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 92-120). New York, NY: Oxford University Press.

Schneirla, T. C. (1949). Levels in the psychological capacities of animals. In R. W. Sellars, V. McGill, & M. Farber (Eds.), *The quest of modern materialism* (pp. 243-286). New York, NY: MacMillan.

Schultheiss, O. C. (2001). An information processing account of implicit motive arousal. In M. Maehr & P. Pintrich (Eds.), *Advances in motivation and achievement: New directions in measures and methods* (Vol. 12). Greenwich, CT: JAI Press.

Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences, 25*(2), 167-177.

Seligman, M. E. (2004). *Character strengths and virtues: A handbook and classification*: Oxford University Press.

Shamay-Tsoory, S. (2009). Empathic processing: Its cognitive and affective dimensions and neuroanatomical basis. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 215-232). Cambridge, MA: MIT Press.

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review, 115*(2), 314.

Sherry, D. F. (2006). Neuroecology. *Annual Review of Psychology, 57*, 167-197.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*(2), 127-190.

Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). Cambridge, MA: Massachusetts Institute of Technology Press.

Slavin, R. E. (1996). Research on Cooperative Learning and Achievement: What We Know, What We Need to Know. *Contemporary Educational Psychology, 21*(1), 43-69.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*(1), 3-22.

Smith, C. (2013). Moving the needle on "moving the needle": Next stage technical guidance for performance based accountability systems in the expanded learning field with a focus on performance levels for the quality of instructional services. In *Weikart Center for Youth Program Quality, a division of the Forum for Youth Investment.*

Smith, C., Akiva, T., Jones, M., Sutter, A., Hillaker, B., Wallace, L., & McGovern, G. (2012). *Program Quality Assessment Handbook: Youth Version*: David P. Weikart Center for Youth Program Quality at the Forum for Youth Investment.

Smith, C., Akiva, T., Sugar, S., Lo, Y. J., Frank, K. A., Peck, S. C., & Cortina, K. S. (2012). Continuous quality improvement in afterschool settings: Impact findings from the *Youth Program Quality Intervention study*: David P. Weikart Center for Youth Program Quality at the Forum for Youth Investment.

Smith, C., Hallman, S., Hillaker, B., Sugar, S., McGovern, G., & Devaney, E. (2012). Development and early validation evidence for an observational measure of high quality instructional practice for

science, technology, engineering and mathematics in out-ofschool time settings: The STEM supplement to the Youth Program Quality Assessment. In *Weikart Center for Youth Program Quality, a division of the Forum for Youth Investment and Providence Afterschool Alliance*

Smith, C., Helegda, K., Ramaswamy, R., Hillaker, B., McGovern, G., & Roy, L. Quality-Outcomes Study for Seattle Public Schools Summer Programs Summer 2015 Program Cycle.

Smith, C., Hillaker, B., & McGovern, G. (2014). *Program Quality Assessment Handbook: Academic Skill-Building Supplement to the Youth PQA.* Ypsilanti, MI: David P. Weikart Center for Youth Program Quality at the Forum for Youth Investment.

Smith, C., & Hohmann, C. (2005). Full findings from the Youth PQA validation study. Ypsilanti, MI: *High/Scope Educational Research Foundation.*

Smith, C., McGovern, G., Larson, R., Hillaker, B., & Peck, S. C. (2016). *Preparing Youth to Thrive: Promising Practices for Social and Emotional Learning.* Ypsilanti, MI: David P. Weikart Center for Youth Program Quality at the Forum for Youth Investment.

Smith, C., Moxley, K., McGovern, G., Helegda, K., & Roy, L. (2016). *Evaluation of Quality Improvement System Performance for Oklahoma 21st Century Community Learning Centers.* Ypsilanti, MI: David P. Weikart Center for Youth Program Quality at the Forum for Youth Investment,.

Smith, C. A., & Kirby, L. (2000). Consequences require antecedents: Toward a process model of emotion elicitation. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 83-106). New York, NY: Cambridge University Press.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and social psychology review, 4*(2), 108-131.

Springer, J. F., Sale, E., Hermann, J., Sambrano, S., Kasim, R., & Nistler, M. (2004). Characteristics of Effective Substance Abuse Prevention Programs for High-Risk Youth. *The Journal of Primary Prevention, 25*(2), 171-194.

Stein, S. (2000). Equipped for the Future Content Standards. What Adults Need To Know and Be Able To Do in the 21st Century.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*(3), 220-247.

Surr, W., & Tracy, A. (2009). SAYO-Y: Youth Survey (for grades 9-12).

Teper, R., Segal, Z. V., & Inzlicht, M. (2013). Inside the Mindful Mind How Mindfulness Enhances Emotion Regulation Through Improvements in Executive Control. *Current Directions in Psychological Science, 22*(6), 449-454.

Thayer, R. E. (1989). *The biopsychology of mood and arousal*. New York, NY: Oxford University Press.

Vargha, A., Torma, B., & Bergman, L. R. (2015). ROPstat: A general statistical package useful for conducting person-oriented analyses. *Journal for Person-Oriented Research, 1*, 87-98.

Von Eye, A., & Bogat, G. A. (2006). Person-oriented and variable-oriented research: Concepts, results, and development. *Merrill-Palmer Quarterly, 52*, 390-420.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: Harvard University Press.

*Watkins, N. D., Larson, R. W., & Sullivan, P. J. (2007). Bridging intergroup diffrence in a community youth program. *American Behavioral Scientist, 51*(3), 24.

Wentzel, K. R. (1991). Social Competence at School: Relation Between Social Responsibility and Academic Achievement. *Review of educational research, 61*(1), 1.

Wheeler, M. E., Keller, T. E., & DuBois, D. L. (2010). Review of Three Recent Randomized Trials of School-Based Mentoring: Making Sense of Mixed Findings. Social Policy Report. Volume 24, Number 3. *Society for Research in Child Development*.

Wilson-Ahlstrom, A., Yohalem, N., DuBois, D., Ji, P., Hillaker, B., & Weikart, D. P. (2014, January). From soft skills to hard data: Measuring youth program outcomes. In *Forum for Youth Investment*. Forum for Youth Investment. The Cady-Lee House, 7064 Eastern Avenue NW, Washington, DC 20012-2031.

Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine, 33*(2), S130-S143.

Winter, D. G. (1992). Responsibility. In C. P. Smith & J. W. Atkinson (Eds.), *Motivation and personality: Handbook of thematic content analysis*. New York: Cambridge University Press.

*Wood, D., Larson, R. W., & Brown, J. R. (2009). How adolescents come to see themselves as more responsible through participation in youth programs. *Child Development, 50*(1), 295-309.

Wray-Lake, L., & Syvertsen, A. (2011). The Developmental Roots of Social Responsibility in Childhood and Adolescence. In C. A. Flanagan & B. D. Christens (Eds.), *New Directions for Child and Adolescent Development. Youth civic development: Work at the cutting edge* (Vol. 134, pp. 11-25).

Yeager, D., Bryk, A., Muhich, J., Hausman, H., & Morales, L. (2013). Practical measurement. *Palo Alto, CA: Carnegie Foundation for the Advancement of Teaching*.

Yoder, N. (2013). Teaching the whole child: Instructional practices that support socialemotional learning in three teacher evaluation frameworks. American Institutes for Research.

*indicates initial Larson et al. evidence base

Appendix A - Continuous Improvement Intervention Summary and Evidence

The continuous improvement intervention consists of four elements: (1) A standard for good practices and aligned performance measures, (2) data products that communicate about performance in actionable ways, (3) a continuous improvement cycle implemented over multiple program cycles, so that staff have opportunities to practice new skills and develop expertise, and (4) technical assistance, training, and technology necessary to implement the prior three (Smith, 2013). The products of the SEL Challenge include performance standards, performance measures, data products, and supplemental content and training as supports for the adult skill learning and behavior change process.

System-Level and Organization-Level Continuous Improvement Cycles

The Youth Program Quality Intervention (YPQI) is an evidence based continuous improvement intervention (YPQI;Smith, 2013; Smith, Akiva, Sugar, et al., 2012) and is currently deployed in over 105 systems and over 4,500 program sites by an estimated 34,000 staff. Figure A-1 describes the System and Organization-level cycles implemented as part of the YPQI. A key is provided in Figure A-1 to identify several key elements of the continuous improvement intervention:

1. *Design – Implement – Evaluate Cycle*. This system-level continuous improvement cycle is typically the responsibility of agency or system leaders and includes the design of the quality improvement system, implementation of the assess-plan-improve sequence at each program site, and evaluation of aggregate performance information during each annual cycle, and over multiple cycles.
2. *Assess – Plan – Improve Cycle*. This site-level continuous improvement cycle is typically the responsibility of the site manager and includes assessment of site performance, review of performance data, and implementation of an improvement plan for the site. For Weikart Center clients, this cycle is the YPQI.
3. *Training and Technical Assistance*. These supports include a wide range of design consulting and evaluation services for the system-level cycle and the YPQI package of supports for program self-assessment of performance, planning with performance data, and implementation of the performance improvement plan.
4. *Performance Data Collection*. Performance data can include a wide range of measures and indicators, including observational ratings of teacher practices as well as measures of youth SEL skills. These data are the core of the planning element of the assess-plan-improve cycle.

5. *Performance Reporting*. These data products are designed to support decision making at the site and system levels. Site-level performance reports are typically produced for each program site early in the annual cycle so that improvement planning can occur. The aggregate report carrying information about all program sites is typically produced after the end of an annual cycle so that design adjustments can occur in the system and site-level cycles.

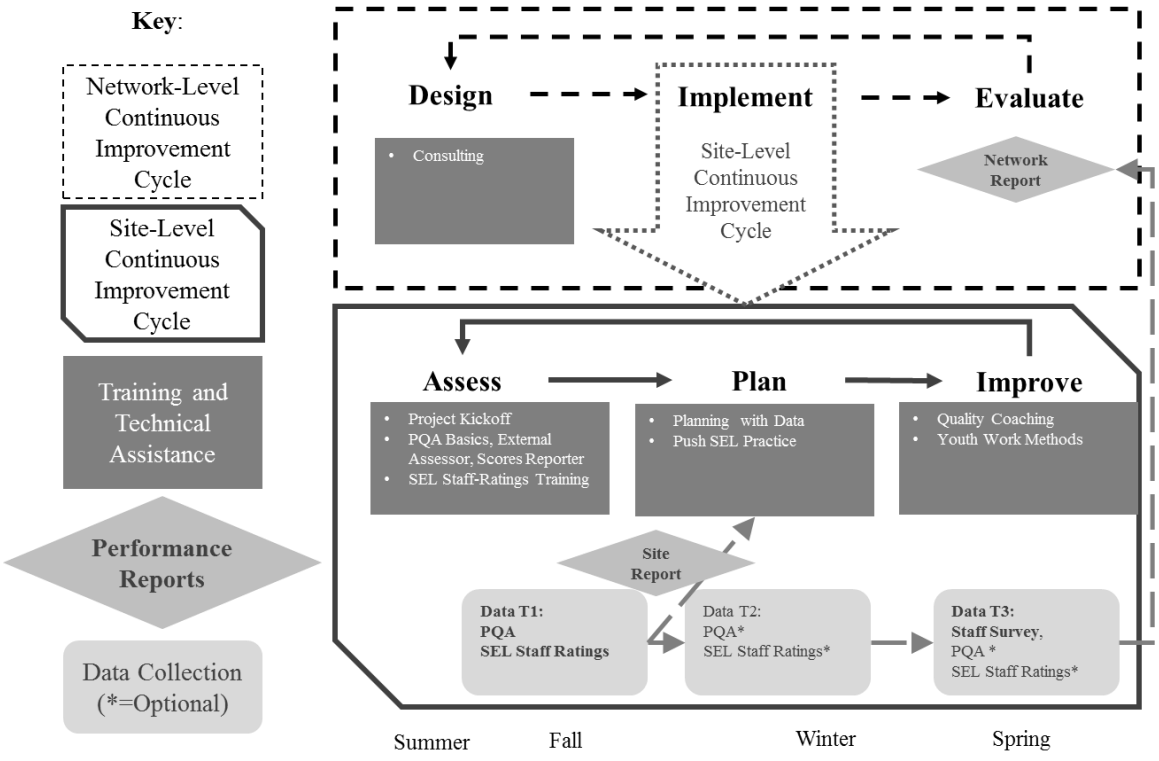

Figure A-1. **Logic Model for the Continuous Improvement Intervention**

Summary of Validation Evidence

The YPQI improves the quality of afterschool program offerings. YPQI is a multilevel intervention that produces a cascade of effects from the network, or system, level, where it is introduced, through to managers of afterschool organizations, through to improved quality of teaching practices and curriculum at the point of service, and culminating in youth skill growth and transfer. In a prior, randomized controlled trial for the YPQI, an intent-to-treat cascade of effects was demonstrated with positive and statistically significant effects at the network (d=1.15), organization (d=.75), and point of service (d=.55) levels (Smith, Akiva, Sugar, et al., 2012). The final point-of-service level effect was for the quality of instruction (e.g., SEL practices) during afterschool offerings as assessed using the Youth Program Quality Assessment (Youth PQA) (Youth PQA;Smith & Hohmann, 2005).

Subsequently, several studies have linked high-quality instruction (e.g., SEL practices), as assessed by the Youth PQA, to changes in student skills and outcomes (Naftzger, 2014), completing the hypothesized cascade of effects. In these studies, participation in higher-quality afterschool settings was associated with (a) heightened levels of mental engagement (Akiva et al., 2013), (b) mastery of specific SEL and academic skills (Smith et al., 2015), and (c) improved school success outcomes (e.g., grades, achievement, behavior, on-time promotion) for both elementary and secondary students (Naftzger, Devaney, et al., 2014; Naftzger, Hallberg, et al., 2014; Naftzger et al., 2013).

Appendix B -Technical Supplement for the Qualitative Study

This appendix presents supplemental information for the qualitative study used to produce the SEL standards and curriculum features. The literature review was developed to inform the interview and focus group questions and to set a framework for the corresponding narrative data that were the answers to those questions. We also selectively reviewed existing SEL measures in each of the six SEL domains. The initial 20 Larson et al. articles that were reviewed to develop the framework are listed with asterisks in the reference section.

Next, the questions used to generate the narrative data are presented, by instrument. Finally, results from supplemental content analyses are presented to demonstrate the distribution of SEL practices across the eight exemplary offerings.

Literature Review for Preliminary Coding Framework

This literature review provided a preliminary coding framework for the qualitative analysis and supported the identification SEL practices as well as an early set of curriculum features. The review is organized by the six social and emotional skill domains, although empathy and teamwork were originally treated as a single SEL domain.

The review for each of the skill domains has three parts. First, we describe domain-specific staff practices and key youth experiences drawn from the Larson et al. corpus of research. Second, we summarize how the skills have been measured in previous research. Finally, we provide a table capturing the preliminary coding framework and skill descriptors for the domain. In addition, we provide a brief description of the curriculum features underlying the social and emotional learning across the domains.

Emotion Management

Review of the Larson et al. Literature on Emotion Management

Youth programs that successfully teach emotion management foster an environment where teens repeatedly encounter a matrix of positive and challenging emotions and their causes and provide a space for the youth to manage their emotions through discussion, suppression, motivation, or attention (Rusk et al., 2013). Larson and Brown (2007, p. 1091) found in a study of a youth theater program that "positive emotions were encouraged and often spread throughout the group. Negative emotions were discussed openly and often elicited supportive responses that helped dispel them."

Youth also learn to manage the emotions of others by gaining knowledge of personality differences and experiencing empathy-building occasions with peers. Through this learning cycle, youth

are able to begin to detect the causal roles of certain settings and factors, as well as how to best respond. A youth respondent in the theater program observed a learning cycle through his participation in the program, saying, "One thing drama has definitely taught me is that when you're tired, you are emotional. If I've had a long day or the rehearsal has gone on a little bit too long, you really realize that you're a lot angrier, you're a lot [more] short-tempered, or a lot more emotional in pretty much every way than you would normally be" (Larson & Brown, 2007, p. 1092).

Youth also learn to manage emotions by channeling what they are feeling into motivation or attention. A young athlete describes his experience by saying, "I was mad because I knew we could get to the playoffs if we do better. And I was scared because I didn't think we would get better. [So] I started practicing. I would practice by myself so I could get better and release frustration. It made me practice harder" (Rusk et al., 2013, p. 253).

Adult leaders in high quality programs normalize strong emotions and model appropriate and positive emotional management strategies. Leaders make themselves available to youth to coach problem solving and reflection during and after emotional moments. When youth look upset, a teacher in a youth program asks, "What happened? You were down low. Talk to me. What's going on? Because I've seen that you just weren't in it today" (Rusk et al., 2013, p. 255). Leaders in these settings work to cultivate an environment of emotional support, consistent with productive adult work groups that are mission-driven and relationship-centered (Larson & Brown, 2007).

Measurement of Emotion Management Skills

Researchers over the past decades have accumulated increasing evidence that emotion-related skills are important for personal well-being, social competence, personal relationships, and even academic learning. We define *emotion management* as the ability to be aware of, and constructively handle, both positive and challenging emotions. Through our work with expert practitioners, we have identified the following skill sets that comprise the broader emotion management domain: (1) Identify positive and negative emotions; (2) Reason about causes and uses of emotion; (3) Manage emotions for functional purpose.

The skill sets for the emotional management domain tend to be identified in the field as aspects of emotional intelligence or of emotion regulation. In general, emotional intelligence is the ability to use emotional information to reason abstractly (Lippman, Moore, et al., 2014). Mayer and Salovey's (1997; 2004; 1990) theory of emotional intelligence explains the existence of four related branches of emotional intelligence arranged in a hierarchical order from the least to the most psychologically complex. Three of the four branches of Mayer and Salovey's model align well to the three subdomains: Perceiving and Identifying Emotions is the ability to recognize how you and those around you are feeling and aligns to subdomain 1; Using Emotions to Facilitate Thought is the capacity to use emotions to assist thinking and

aligns to subdomain 2; Understanding Emotions is the ability to analyze emotions, notice their trends over time, and understand their outcomes, and is contained within subdomain 2; and Managing Emotions is the ability to manage emotions in yourself and in others and aligns to subdomain 3.

Related constructs in the Behavioral and Emotion Rating Scale (BERS-2) include Interpersonal Strength (e.g., reacts to disappointment in a calm manner), which measures the ability to control emotions and behaviors in social situations, and Affective Strength (e.g., acknowledges painful feelings), which focuses on the ability to give and receive affect (Epstein, Mooney, Ryser, & Pierce, 2004).

Other measurement concepts with some overlap in meaning with emotion management include self-regulation (Maxwell, 1989), self-management, optimism, emotional stability (Hogan, Johnson, & Briggs, 1997), impulse control (Gratz & Roemer, 2004), and hope. Processes for emotion regulation may include cognitive reappraisal (i.e., altering the emotional response by reinterpreting the meaning of the stimulus) and expressive suppression (i.e., where behavioral expressions of emotions are inhibited) (Gross, 2013).

Other existing measures widely used in the field are focused on the lack of emotional management. For instance, the Difficulties in Emotion Regulation Scale measures the absence of several of our constructs, including non-acceptance of emotional responses, lack of emotional awareness, limited access to emotion regulation strategies, and lack of emotional clarity.

Preliminary Coding Framework and Skill Descriptors for Emotion Management

Key Youth Experiences. Repeatedly encounter matrix of positive/negative emotions—stress, anger, frustration, anxiety, pride, elation; Repeatedly recognize emotions and their causes; Use if-then thinking about causes and outcomes; Repeatedly manage own emotions to optimize work and progress through challenges through discussion, suppression, motivation, attention, and flow; Repeatedly manage others' emotions to optimize work and progress through challenges.

Staff Practices. Fostering awareness of and reflection on emotion; Openness to discuss emotions; Suggesting strategies for managing emotions in self and others; Encourage problem solving.

Skill and Belief Descriptors. Recognize emotions and their causes; Anticipate emotional dynamics for performance; Manage own emotions to optimize performance; Manage others' emotions to optimize performance; Belief that self is capable of skills named above

Teamwork/Empathy

Review of the Larson et al. Literature on Teamwork/Empathy

Researchers found that when students are exposed to others' personal stories, they are able to form connections between their own experience and the experiences of others, and this connection transforms their motivation and the tone of their actions (Larson, 2011). A student explained, "In the

streets you don't really go up to a person and talk. Right here you get the opportunity to actually talk to a person, you know, see what's different about cultures" (Watkins, Larson, & Sullivan, 2007, p. 390). Building empathy in youth occurs in three stages: (1) interacting with groups different from their own; (2) gaining a better understanding of the groups; and (3) acting in ways that are informed by a greater awareness of diversity. According to the researchers "the program serves as a facilitative environment that provides conditions for youth to become empowered to change" (Watkins et al., 2007, p. 396).

Learning empathy, teamwork, and leadership also provides youth with an opportunity to exercise emotional management and responsibility skills. A teen reflects that "by being captain [of the team], you have to represent your whole team. If you do something stupid, it looks like your whole team is doing it…a lot of people look up to you and they follow your example. Like if you started clapping because somebody got hurt and they walk off the field, everyone will do it. You just have to be the good example" (Dworkin, Larson, & Hansen, 2003, p. 23). Leadership is an important component of this domain, and also includes such skills as learning how to delegate tasks effectively, as well as learning to take and give feedback from others.

Measurement of Empathy/Teamwork Skills

Empathy and teamwork are necessary skills for navigating social interaction with peers, colleagues, and others. For this project, we define teamwork/empathy as abilities to collaborate and coordinate action with others and relating to others with acceptance, understanding, and sensitivity to their diverse perspectives and experiences. This definition focuses on social skills and empathy as applied in a group setting where the group needs to work together to accomplish a specific task, goal, or purpose. Therefore we selected the label teamwork instead of a more generic label such as prosocial skills or communication skills. Empathy/teamwork also captures the element of cultural sensitivity, of particular importance in a multicultural group, as well as personal sensitivity to others' feelings.

Through our work with expert practitioners, we have identified the following skill sets that comprise the empathy/teamwork domain defined above, and which are especially of interest in out-of-school time programs: (1) Values own/others perspectives and stories with sensitivity to context; (2) Practices respectful and effective communication within a team; (3) Coordinates and supports action toward team goals. The teamwork/empathy skill sets correspond loosely to objects of measurement identified in other standard measures in the field.

The widely used Social Skills Rating Scale (SSRS; Gresham & Elliott, 1990) delineates Empathy, Cooperation, Self-Control and Assertion as subscales. Our definition of the empathy/teamwork domain overlaps with three of the four SSRS subscales. SSRS Empathy corresponds to subdomain 1, "Valuing others perspectives and stories." This is the empathetic element of our empathy/teamwork domain, which captures what other researchers call *cognitive empathy* (Davis, 1994). Subdomain 3, "Coordinate and

support action toward team goals" captures the essence of SSRS Cooperation. SSRS Assertion is one aspect of subdomain 2, "practices respectful and effective communication"—as evidenced in the behavioral indicator 2a "Clearly articulates point of view and/or asks follow-up questions for clarification" and 2b "Communicates without dominating, interrupting, or showing disrespect for others' ideas." Additionally, our constructs align with primary dimensions/items of the Social Skills Scale of the National Survey of Children's Health Social Competence Scale (Blumberg, Carle, O'Connor, Moore, & Lippman, 2007): shows respect; gets along well with other(s); tries to understand other people's feelings, tries to resolve conflicts.

Other researchers have also identified similar elements. Shamay-Tsoory (2009) included a conceptual, perspective-taking process as empathy. Stein (2001) included the following in cooperation with others: demonstrates respect for others' ideas, opinions, and contributions; seek input from others in order to understand their actions and reactions; offer clear input on own interests and attitudes so others can understand one's actions and reactions; try to adjust one's actions to take into account the needs of others and/or the task to be accomplished. In the literature, these elements are variously captured in measures labeled prosocial skills, prosocial behavior (Eisenberg, Fabes, & Spinrad, 2007), teamwork, social competence or social competencies (Holopainen, Lappalainen, Junttila, & Savolainen, 2011), social skills, empathy, cooperation, communication skills, social awareness, and relationship skills (For a review of measures, see: Wilson-Ahlstrom, Yohalem, Dubois, Ji, & Hillaker, 2014).

Preliminary Coding Framework and Skill Descriptors for Teamwork/Empathy

Key Youth Experiences. Engage in collaborative work with peers outside of normal networks; Increase understanding of others' stories and own; Build relationships with loyalty and intimacy; Reflect on culture, stereotypes, injustices, and discover the humanity in others; Practice skills such as respect leaders, resolve conflict, receive and give feedback, communicate, listen; Integrate responsibility and emotion management skills.

Staff Practices. Speak up on issues of diversity; Cultivate youth collaboration; Balance supporting the work of teams with letting teams make decisions; Structure opportunities for youth to learn to give and receive peer feedback.

Skill and Belief Descriptors. Pro-social behaviors; Act with awareness of cultural diversity; Demonstrate domain behaviors (e.g., resolve conflict, performance feedback, communication & listening); Belief that self is capable of skills named above.

Responsibility

Review of the Larson et al. Literature on Responsibility

Research suggests that youth develop responsibility in a four-part cycle: (1) youth voluntarily enter into obligations, often in the form of a structured role; (2) youth experience challenge and strain as a part of their obligation; (3) youth are motivated to persevere despite the challenges and effectively draw upon contextual resources and/or their pre-existing dispositional reserves to follow through; and (4) youth experience success in fulfilling obligations and transfer responsibility to other contexts. By moving through this cycle, youth internalize the motivations behind their responsible behavior, which helps them view themselves as more mature and self-confident (Salusky et al., 2014).

According to youth in research studies, successfully prioritizing their program obligations over social activities contributed to youth seeing themselves as more responsible and capable of meeting demands. A Future Farmers of America member reflected on the honor and responsibility of earning an elected position: "If you want to be a chapter officer, then you've got to treat others with respect and you've got to prove yourself to be responsible and mature and things like that" (Wood, Larson, & Brown, 2009, p. 301). Additionally, each experience of fulfilling new and more demanding expectations further deepens their perception of themselves as responsible.

Adult leaders support the development of youth responsibility by balancing high expectations while judiciously adjusting expectations and providing support. Additionally, leaders in exceptional programs provide structured opportunities for youth to take on obligations while also leaving room for the youth to exercise agency and decision-making skills. Adult staff members made clear that there were tangible consequences that accompanied failing to follow through with an obligation, simultaneously communicating confidence that youth are capable of success (Salusky et al., 2014).

Measurement of Responsibility Skills

The more adult roles adolescents play in continuing education, the workplace, and their families require young people to be able to fulfill obligations and commitments. This is the essence of what we are labeling responsibility: the dispositions and abilities to reliably meet commitments and fulfill obligations of challenging roles. While many different skills and traits go into being able to carry out the tasks for which one is responsible, for this project we are primarily interested in those skills or constructs specifically related to responsibility or commitment to a specific role in an OST context. This is aligned with the definition of diligence and reliability from Lippman et al. (2014): "the performance of tasks with thoroughness and effort from start to finish where one can be counted on to follow through on commitments and responsibilities" (p. 13).

Through our work with expert practitioners, we have identified the following skill sets that comprise responsibility: (1) fulfills roles and commitments, and (2) successfully defines, adjusts, and negotiates roles and commitments when required.

The term responsibility is used in a number of literatures to mean a number of different things. It is, along with diligence and reliability, one of the facets of Conscientiousness, one of the big five personality factors (Seligman, 2004). It is typically defined as the quality of being someone who can be counted on to fulfill obligations (Winter, 1992). A broader term, social responsibility, is defined as adherence to social rules and role expectations (Wentzel, 1991). Related terms include diligence and reliability—"the performance of tasks with thoroughness and effort from start to finish where one can be counted on to follow through on commitments and responsibilities" (Lippman, Ryberg, et al., 2014, p. 13). Our definition of responsibility is intentionally narrower and focuses on the development of responsibility within the context of taking on roles within an OST program.

Preliminary Coding Framework and Skill Descriptors for Responsibility

Key Youth Experiences. Voluntary entry into obligations for work with high expectations and real consequences where goals are known and means are owned by youth; Experience role stress (and wavering commitment) and potentially adjust expectations; Draw upon resources (self-concept, leaders' expectations and guidance, solidarity with and obligation to peers) in fulfillment of role/obligation; Successfully meet obligations.

Staff Practices. Create structured but open-ended roles; Balance high expectations with support; Cultivate peer cohesiveness and teamwork.

Skill and Belief Descriptors. Successfully meet obligations associated with role and adjust expectations or role where necessary; More organized, confident, and able to manage emotions associated with stress; Learn domain-specific role (e.g., meeting organizer); Belief that self is capable of skills named above.

Initiative

Review of the Larson et al. Literature on Initiative

Youth in programs that effectively teach initiative acknowledge that the program activities acquire greater significance when they speak to youth's personal values, ambitions, or identity. Research shows that youth attributed increased engagement to changes within themselves, and their motivation became more self-determined. Such programs encourage psychological engagement, which develops through youth reflecting on their identity and their personal or transcendent goals (Dawes & Larson,

2011). Initiative is realized when youth have learned to set realistic goals, manage their time, and take the responsibility for themselves to persevere toward the actualization of their goals (Dworkin et al., 2003).

Connections to the program and mission may be particularly strong when the mission and activities are collective and the youth share a cause. A youth in a leadership development program said that she just "wanted to get the [service] hours and then quit, [but I became] interested in all the subjects [they were] talking about. I realized that a lot of kids have been dying because of gangs, and I want to stop that" (Dawes & Larson, 2011, p. 264). Over time in these programs, the youth experience the tasks to be fun, exciting, and enjoyable—in other words, they find the OST programs intrinsically motivating (Pearce & Larson, 2010).

Staffs in programs that successfully instill initiative in the youth provide learning opportunities that help youth form personal connections to the program's mission and activities in a deliberate fashion. Staff also encourage critical thinking among the youth and construct challenges for the teens to take on (Pearce & Larson, 2010).

Measurement of Initiative Skills

Youth and adults face a variety of challenges as they navigate the increasingly complex, technical, and multicultural world. Success in school, work, and life depends on the ability to persevere through these challenges and continue to strive for long-term and higher-order goals. For this project, we define initiative as the capacities to take action, sustain motivation, and persevere through challenge toward an identified goal.

Through our work with expert practitioners, we have identified the following skill sets for initiative: (1) Develops and hones motivation for the task, and (2) Perseveres through internal and external circumstances that challenge the work.

The skill sets for initiative comprises a set of concepts referred to in the literature using a variety of terms, including grit, purpose, initiative, and persistence. Duckworth and colleagues have championed the work of identifying grit as a particular key to youth success in achieving long range goals. Although, in actuality, a wide range of intertwined skills support long-range goal achievement, for this domain we focus on the elements in Duckworth et al.'s concise definition of grit as "perseverance and passion for long-term goals" (Duckworth & Quinn, 2009, p. 166). The two components of Duckworth et al.'s definition align well with our subdomains.

Lippman and colleagues' concept of a "purpose" aligns well with our initiative domain. Purpose contains three components: (1) a sense of directedness that stimulates one's goals, manages one's behaviors, and provides one a sense of meaning; (2) a broad and sustained intention to accomplish something one finds meaningful; and (3) guidance for the use of attention and energy (Lippman, Ryberg, et al., 2014, p. 17).

Initiative is another term that in the literature sometimes includes aspects of perseverance and the motivation necessary to persist. For this project, we have distinguished persistence and motivation from other elements of initiative for inclusion in our domain of initiative. Lippman and colleagues (2014) defined "initiative taking" as the practice of initiating and manifesting activity toward a specific goal by adopting the following characteristics: "(a) reasonable risk taking and openness to new experiences, (b) drive for achievement, (c) innovativeness, and (d) willingness to lead" (p. 14). Measures of initiative have included the idea of persisting despite challenges (SAYO; Miller & Surr, 2007) or identified "initiative experiences" as including the constructs of goal setting, effort, problem solving, and time management (Youth Experiences Survey; Hansen, Larson, & Dworkin, 2003).

Persistence is a construct in the Adaptive Behavioral Dimension of the Motivation and Engagement Scale defined as the extent to which students sustain their engagement (Liem & Martin, 2012). Lippman et al. (2009) included in intrinsic motivation the tendency to persist with a difficult task to achieve mastery. Other concepts related to our domain of grit/initiative include self-discipline, persistence, self-control, task persistence, and goal-directed behavior.

Preliminary Coding Framework and Skill Descriptors for Responsibility

Key Youth Experiences. Set realistic goals, demonstrate effort and perseverance, manage time, demonstrate responsibility; Develop personal connection to the work: competency, future/ career fit, broader social purpose; Experience ownership of the work.

Staff Practices. Engage with youth's personal values and goals; Explore multiple opportunities to find and develop authentic personal connections to work, especially as competence, future/career fit, and broader social purpose.

Skill and Belief Descriptors. Persistence; Domain skills (e.g., set goals, manage time); Belief that self as capable of skills named above.

Problem Solving

Review of the Larson et al. Literature on Problem Solving

Support for the problem solving domain is found in studies of youth in a variety of program settings. Programs that successfully build problem solving skills in youth provide opportunities for teens to practice strategic thinking while they acquire context-specific skills. By applying strategic thinking to concrete tasks, youth are also gaining the capacity to set goals and become future-oriented.

Researchers have described how teens are developmentally poised to be future-oriented. Larson observed that "for these youth, their new motivation was quite internal, authentic, and powerful…Planning and preparing for the future is a task of adolescence in our culture. Forming

connection from the program to future goals really got their fires lit" (Larson, 2011, p. 325). Additionally, teens are beginning to be able to engage in reasoned anticipation and forecast possible scenarios (Larson, 2011). Programs that allow youth to set goals and face real consequences are best suited for this area of skill development.

A student in a community organizing program demonstrated the development of strategic thinking by saying, "If we have an idea, we'll make plan A, and just in case plan A fails, we have plan B. We've never had to use plan B, so we always went through with plan A. But I think this specific case with the [new citywide exam], I think we might have to go to plan B" (Larson & Hansen, 2005, p. 339). Such programs allow youth to understand interactions within complex systems and create contingency plans to meet their goals.

By applying strategic thinking to concrete tasks, youth are also gaining the capacity to set goals and become future-oriented. A student in an art program said, "At first it was just more for fun, just to go and do some artwork and stuff. But now I really want to go do art a lot more, like [be] an art major at school" (Larson & Angus, 2011). Researchers also observed that these opportunities often build teamwork, because effective programs are likely to encourage collaboration with domain-specific tasks (Larson & Angus, 2011).

Research suggests that adult leaders are less interested in teaching teens and more interested in helping them teach themselves by being responsive and providing appropriate structure, challenge, and support. Youth at a recording studio described the training provided by the staff as, "They tell us the basics of how the sound gets in there and how to put the sound into the machines. They let us tinker with it on our own until we find a good pattern and know what sound it is, then we go on from there. Basically you get to do it by yourself" (Larson & Angus, 2011, p. 288). Effective adult leaders of agency development provided assistance to youth on a conditional basis—*when and if* youth requested or needed assistance—so that the youth maintained control of their goals and actions (Larson & Angus, 2011). Adult leaders should see themselves as experienced collaborators who contribute to the youth's projects.

Measuring Problem Solving Skills

To take an idea, put it into action, and successfully navigate the process to completion requires a set of high level and complex skills. We have called this problem solving, which we have defined as the abilities to plan, strategize, and implement complex tasks. Problem solving skills involve both learning specific task-related skills and applying a broader set of skills, including anticipating consequences and making adjustments to support achievement of the larger task. One must be able to set priorities, assess options, adapt plans as circumstances warrant, and manage time. At more advanced levels, the action and adaptation required may involve connecting and communicating with external stakeholders.

Through our work with expert practitioners, we have identified the following skill sets for problem solving: (1) Intentionally learns out-of-school time task-related methods and tools (e.g., carpentry, sexual health, theatre, organizing), (2) Uses problem-solving and strategic thinking skills to develop, evaluate, and adapt a course of action, (3) Successfully manages time, (4) Connects to external stakeholders and (5) Reflects on learning and significance of results.

Problem solving skills have been widely measured independently and as part of a number of variously labeled broader constructs. Our first subdomain is context-specific; the skills learned have immediate applicability to a "real" task and therefore are both important and necessary. This is connected to constructs frequently measured as sense of competence (Miller & Surr, 2007; Surr & Tracy, 2009) or self-efficacy (Bandura, 2006).

Our second and third subdomains are skill sets with broad applications across many contexts. Skills for action are similar to Lippman and colleagues' definition of goal orientation: "the ability to make viable plans and take action toward desired goals" (Lippman, Ryberg, et al., 2014, p. 2832). The ability to make viable plans includes the problem-solving and strategic thinking skills necessary to assess whether a plan is feasible. The action part of goal orientation, particularly action that pushes toward goal completion, is included under our initiative domain.

These skills are frequently described as problem-solving skills. (SAYO; Arbeau, 2002; Miller & Surr, 2007; Stein, 2001). Problem solving is also often called critical thinking, strategic thinking, systems thinking, or decision making. Problem solving includes optional thinking, causal thinking, creating step-by-step solutions, and consequential thinking, among other things (Platt, Spivack, Altman, Altman, & Peizer, 1974). Competent decision-making includes the process of weighing and considering all of the options, risks, benefits, and other key components involved in the decision-making process (Halpern-Felsher, 2009). Goal setting, problem solving, and time management have been included as sub-constructs under initiative (YES 2.0, Hansen et al., 2003). Prioritizing tasks is another aspect of skills for action that has been included under initiative (SAYO-T, Miller & Surr, 2007). Metacognitive strategies—planning how to approach a given learning task and evaluating and monitoring progress—is another term used (Hattie, 2009). The elements of metacognition that involve evaluating and monitoring progress also relate to subdomain five, reflecting on learning and significance. Construct four is an advanced indicator of problem solving; connecting with external stakeholders takes problem solving into a broader sphere and moves closer to the concept of social responsibility (Wray-Lake & Syvertsen, 2011).

Preliminary Coding Framework and Skill Descriptors for Problem Solving

Key Youth Experiences. Practice setting, pursuing, adapting, and reaching difficult goals; Pursue difficult goals that entail challenges (e.g., puzzles, obstacles, problems, and situational demands);

Experience successes and failures; Acquire knowledge of domain-specific content (e.g., school discipline policies, theater production); Practice strategic thinking (e.g., active anticipation, using knowledge of how people think and act, flexible planning adapted to anticipated scenarios, transfer of learning); Use models and abstractions; Develop guidelines; Evaluate trial and error evidence; Integrate initiative, responsibility, emotion management, and teamwork/ empathy skills; Connect work to own future and broader purposes.

Staff Practices. Facilitative advising model assuring youth freedom to experiment; Initial training for problem solving skills; Contributing to youth-driven planning; Providing back-up assistance; Cultivating youth personal connection to the work through growing sense of competency, future career alignment, internalization of broader social purpose.

Skill and Belief Descriptors. Set goals and mobilize effort; Demonstrate domain-general skills and knowledge of domain-specific content.

Curriculum Features

Our initial literature review revealed five curriculum features that effectively facilitate social and emotional development among teens: a defined offering sequence, real-world context, elaborated structure and high expectations, cultivation of youth ownership and personal connection, and facilitative advising (Rusk et al., 2013). We define these as follows:

Offering sequence and learning cycles. An offering sequence is the sequentially organized description of the content available during a program offering.

Real-world characteristics of work. Real-world characteristics of program work refers to youth encountering real choice about the means to achieve goals for the work, tangible connections between the work and future adult roles, and apparent consequences of under-performance. Importantly, we use the term *characteristics* because the purpose of exemplary youth programs is not to let youth fail but to create youth experiences of success (including successful adaptation through mistakes and failures) as they encounter challenges.

Elaborated structure and high expectations. Exemplary youth programs provide sufficient structure and goals to allow youth choice about means in pursuit of difficult goals. Programs with elaborated structure have staff with expertise in the explicit work of the offering and are able to guide youth as they "rediscover" effective means to achieve difficult goals.

Youth ownership through personal connection. Youth experience ownership of projects when they have a sense of personal connection to the work, defined as a growing sense of competency, future career alignment, or internalization of broader social purpose.

Adult youth interaction model includes high quality instruction and facilitative advising. This feature describes how adults and youth interact in programs. High quality instruction is defined by staff

practices that promote a warm and caring climate, active learning, and higher order cognitive engagement. Facilitative advising describes the staff purpose to see that youth stay in control of their work by providing freedom to make mistakes, offering guidance at strategic moments, and providing help when asked. It does not include "rescuing" youth.

Training in domain-specific content. Program provides training in skills necessary to complete the work of the offering (e.g., community organizing offerings provide training in aspects of community organizing).

Qualitative Study Questions

Tables B-1 through B-4 present questions used during the SEL Challenge to generate narrative data about promising SEL practices as part of the qualitative method described in Chapter Two.

Table B-1. Questions from the Letter of Intent

Participant: Program Manager or Grant Writer
Duration: Completed online – had one month to respond

Basics
1. Please include organization name, contact name, website, mission, a brief summary of the organization and program, participant demographics and how your program fits with SCE's Social and Emotional Learning Program priorities.
Background
2. Provide a brief history of your work as it relates to social and emotional learning. How have you explicitly integrated social and emotional learning into your program design?
Focus
3. Provide a clear explanation of how the program helps develop social and emotional skills for youth ages 14-18. Indicate what steps you took to implement social and emotional learning strategies, what practices you have used, and how youth have benefited from these practices. Provide details on the specific skills that the program promotes. If you have measured outcomes, please describe them.
Budget
4. Include your program budget, total youth services budget (required for multiservice organizations) and annual organizational budget including sources of support.
Funding and Partnerships
5. Indicate primary sources of major support and key community partners.
Evaluations and Training Manuals
6. Please include any samples of previous internal and/or external evaluation reports and tools. Please include as an attachment any curriculum or samples of staff training materials relevant to social and emotional learning practices.
Other
7. Indicate how your organization could contribute to and participate in the learning community.

Table B-2. Questions from the Application

Participant: Program Manager or Grant Writer
Duration: Completed online – had one month to respond

Target Offering Description

1. Brief description of offering content.
2. Number of offering sessions and sequence of major activities with approximate dates.
3. Number and characteristics of youth participants.
4. What features does the selected offering(s) have at its core? This could include the structure of activities, program roles for youth, program culture, the nature of the relationships between youth.

Target Skill Domains (2 selected from 5 possible)

5. Describe how youth learn [Skill Domain #1] in your program. Describe the process as you conceive it. Describe the role program staff play in facilitating this learning / process. Provide a specific example.
6. What challenges or obstacles do frontline staff encounter in trying to facilitate these kinds of learning episodes for youth? What strategies do staff use for addressing these challenges?

Questions 5 & 6 are repeated for Skill Domain #2

Youth Backgrounds

7. For the SEL Challenge work, we want to be attentive to how programs are adapted to the background and prior experiences of youth they serve. How are youth's backgrounds' important to understanding these learning episodes and how you facilitate them?

Staff Competencies

8. Describe the competencies staff (and/or volunteers) need to facilitate these learning experiences.
9. Describe the training you provide to staff, especially as it relates to social emotional learning.
10. For each of the roles below [Program Manager, Instruction Staff Member, Local Evaluator or third staff member], please provide a biography of the team member that will be involved. Please include in the biography both a description of the role of the individual within your organization and how the individual will make a unique contribution to the Challenge Learning Community.

Table B-3. Questions from the Telephone Interviews

Participant: Program Manager and Lead Instructional Staff at each site, individually
Duration: 2 hours each

Goals

1. Now I want to ask about your goals for the offering.
2. What do you want youth to get from the offering? What do you hope they will take away?

Offering Sequence

3. Let's dig a little bit deeper into how the offering is designed and what happens over the course of the year to help youth build these skills.
4. In the application, you described the timeline and major activities of the offering. Is my representation in the map accurate?
5. How is the offering structured over time to achieve social-emotional learning goals?
6. Describe the theory behind this progression and how it helps youth build social and emotional skills. Why do you do A before B? What do you want youth to experience at each stage?
7. What is the youth's experience over the progression towards the learning? What are the key experiences/ milestones for youth? What do youth learn at each stage – or how are they changed? I want to understand HOW youth learn. Could I get you to use the voice of a youth in the offering and give their description of how they learned it or why they were changed?

Staff's Role in the Offering Sequence

8. I want to understand how you (or staff) make sure the youth's experience happens and capitalize on it? How do staff monitor, guide, and coach youth through these experiences to help facilitate these experiences?
9. Can we walk through this and you explain the key points in this progression of activities and youth's experiences where you typically play an important role in structuring, guiding, redirecting, or helping make sure youth learn? This could also include: giving feedback, supporting, and helping youth reflect.
10. What do you do to set up or structure the experience at the start?
11. What cues do you look for that youth are on track and learning or are headed in the wrong direction? Can you give me an example?
12. What kinds of things do you do or say to help facilitate the ongoing learning experience? Can you give me an example?
13. Are there things you do or say at the end to help make sure youth learn from the experience? Can you give me an example?
14. What would you share with a novice as the key to do or watch for to make this activity work?
15. How does your role change over time in accordance with these learning cycles?
16. Is the offering based in a set of professional methods or a framework or theory?

Importance of Challenging Work

17. Youths' experiences of setbacks, mistakes, wrong turns, and conflicts can sometimes be key to their social-emotional learning. Let's talk some about opportunities like this in this offering.
18. When do opportunities for setbacks and mistakes happen in the offering sequence?
19. We are interested in examples. Were there any good examples of this this year? Could you tell me about it? What happened? How did you respond? Did you have a role helping address the situation – or helping youth learn from it? What did you say? How did it work out?

Social and emotional skill-building

20. Let's talk about what youth are learning as they go through the offering, both implicit and explicit. Your application identified [Skill Domain #1] and [Skill Domain #2] as two major areas of focus for youth skill development. As I ask you this next set of questions, think about this past year that you are finishing now or have just finished. I'd like to now focus in on just one event or project from the year at a time, really zooming in and getting a good sense of what happened.
21. Can you think of an example of an activity, event or project that went well this year that contributed to one of these domains of youths' social and emotional skills? It could be short or long, it could be a day or a week or a month.
22. Where in the offering sequence did this experience fall?
23. Which skills did youth build as a result of the experience? What insights did youth learn or gain from the experience?
24. In what ways did this event or project help youth develop social emotional skills?
25. What made it a success? Start at the beginning and tell me what happened that led to learning the named skill.
26. What was your role in making this happen at each step?
27. When you started the year, did you expect this event or project to develop in this way?
28. What adjustments did you make to your plan – or to the direction of youth's work? Why did you make these adjustments?
29. If you were mentoring a novice leader, how would you explain your thought process for deciding when and how to adjust your plans?
30. Can you think of an example of an activity, event or project this year that was designed to support youth's SEL learning that didn't go as well as you'd have liked? What happened?

Table B-4. Questions from the On-site Visit

The on-site visit consisted of the following elements and participants. Interview and focus group questions are below.

Element	*Duration*	*Attendees*
Program session observation	1-2 hours (full program session)	Instructional staff, youth, evaluator
Scoring meeting	2 hours	Evaluator
Staff interview	Up to 2 hours	Instructional staff
Manager interview	Up to 2 hours	Program Manager
Expert youth participant interview	Up to 2 hours	Youth from a past program cycle who can reflect on and share their experience
Youth participant focus group	Up to 2 hours	6-10 current youth participants

Questions for Program Manager Interview:

1. How do you recruit youth into the program? How do you know who is the best fit? Which youth is your program designed to help the most?
2. Please tell us how your program approaches traumatic experiences youth have had in the past or may currently be having? What are the kinds of traumatic experience youth in the program have had and how prevalent are these experiences?
3. What kind of social and emotional supports do you provide for your staff? Are your efforts at supporting staff intentionally designed to address their exposure to youth who have experienced trauma?
4. How do you intentionally help staff to be effective SEL skill builders and why is this effort important?
5. What are your guiding principles for working with youth like those in the target offering?
6. What are the strengths of a good leader for this offering? What are the skills you would look for if you were hiring for your position?
7. How did you get good at this work?

Questions for Lead Instructional Staff Interview:

1. What are your guiding principles for working with youth like those in the target offering?
2. How prevalent is traumatic experience in the lives of the youth you serve? How does the experience of working with youth who have experienced trauma affect you?
3. How does your program effectively support your social and emotional development? Is there a parallel process here?
4. If you have one, tell us a good story about a youth you have had in this program in the past and how the program helped them?
5. What are the strengths of a good leader for this offering? What are the skills you would look for if you were hiring for your position?
6. How did you get good at this work?

Questions for Expert Youth Participant:

1. Please tell us how your experience in this program (target offering) last year has affected your life, if at all? How did it affect your life and why was it important to you?
2. We are interested in learning how skills group in five areas: Emotion management, responsibility, etc. Have your skills grown in any of these areas in particular? How do you know?
3. Thinking about the learning that you just mentioned, where else outside of the program do these skills come in handy? Examples?
4. What was most important about how the program worked, about the process? What did the staff do that was most important for you? What opportunities to were most engaging?

Questions for Youth Participant Focus Group:

1. How did you learn about this program? How did you come to participate? What made you want to join?
2. What were the most important things you've learned so far this year in this program?
3. We are interested in learning how skills group in five areas: Emotion management, responsibility, etc. Have your skills grown in any of these areas in particular? How do you know?
4. Thinking about the learning that you just mentioned, where else outside of the program do these skills come in handy? Examples?
5. What were some of the most meaningful moments in the program so far this year? Were there moments when a light bulb went on and you realized something important about yourself?

Supplemental Content Analysis Results

This section presents supplemental results from content analyses of the text segments and reveals patterns in the distribution of SEL practices across the exemplary offerings.

Coding all data into primary codes for the standards

Data from the master data file were coded using 15 primary codes for key youth experience and 16 primary codes for staff practices derived from the preliminary framework. This process resulted in 55% of the 2,261 text segments coded to one or more of the primary codes. Thirty five percent of these segments were coded to multiple codes. Of the 637 segments drawn from interview questions specifically focused on youth experience and staff practices, 83% of the segments were coded to one or more of the 31 primary codes. This indicates that when we asked our expert practitioners about what they did to grow youth SEL skills, most of their answers fit into one of the primary codes, suggesting that the primary codes were capturing large and important areas of professional practice.

We further expected that the two focal SEL domains identified by each program would each have a greater number of coded segments than the remaining three SEL domains that were not selected by each of the eight programs. When we asked programs about staff practices and youth experiences in their two target SEL domains, a preponderance of the text segments were coded into those same two domains. As reflected in Table B-5, in nearly all cases the highest numbers of segments were coded in both of the two focal domains, which are highlighted for each of the eight target offerings. There were only three instances out of 16 where the highest numbers of coded segments for a target offering were not in the two focal domains. Table B-5 also indicates that almost all of the primary codes were used frequently in all programs. This pattern provides empirical support for a message that was often reiterated by the expert practitioners, that all of the domains are important for effective SEL practice.

Table B-5. Frequencies for text segment assignment highlighting two focal SEL domains for offering

Program Number	*Teamwork/ Empathy**	*Emotion Management*	*Problem Solving*	*Initiative*	*Responsibility*
1	35	15	5	2	5
2	74	34	42	21	21
3	27	17	56	21	22
4	58	22	44	24	34
5	47	13	58	36	23
6	65	19	67	22	22
7	43	2	49	20	20
8	59	14	52	23	15
Total Segments	408	136	373	169	162

**At the time of this coding, teamwork and empathy comprised a single domain.*

Coding data into subcodes for the practice indicators

Table B-6 provides the total number of segments coded within each SEL domain (column 1), the total number of those segments coded to each primary code (column 2), and the total number of segments coded to each subcode (column 3). Table B-6 indicates that much of the data was fit to the subcodes and that most subcodes had multiple text segments supporting each practice.

Table B-6. Frequencies for sub-codes within each of five within-domain data sorts

Total within-domain data segments	*Number of segments coded to each primary code*		*Number of segments coded to each subcode*	
Emotion Management segments, N=179	Range of emotions	29	E1	0
	Manage emotions	31	E2	7
	Safe space	18	E3	0
	Modeling	18	E4	3
	Coaching	29	E5	2
	Structure	54	E6	17
			E7	2
			E8	12
			E9	16
			E10	3
			E11	26
Empathy-Teamwork segments, N=575	Diverse perspectives	112	T1	4
	Team challenge	52	T2	46
	Trust and collaboration	93	T3	13
	Safe space	120	T4	31
	Modeling	57	T5	16
	Reflection	54	T6	27
	Structure	27	T7	26
	Coaching	60	T8	39
			T9	44
			T10	45
			T11	81
			T12	3
			T13	43
			T14	26
			T15	29
			T16	0
			T17	34
			T18	6

Table B-6. Frequencies for sub-codes within each of five within-domain data sorts (continued)

Total within-domain data segments	*Number of segments coded to each primary code*		*Number of segments coded to each subcode*	
Responsibility segments, N=229	Roles	80	R1	7
	Strain and adjustment	26	R2	25
	Accomplishment	30	R3	0
	Structure	59	R4	19
	Coaching	34	R5	0
			R6	13
			R7	13
			R8	17
			R9	8
			R10	13
			R11	17
			R12	17
Imitative segments, N=208	Set goals	23	G1	5
	Motivation	54	G2	7
	Perseverance	39	G3	38
	Scaffolding	45	G4	17
	Coaching	47	G5	8
			G6	17
			G7	7
			G8	7
			G9	50
			G10	38
Problem Solving segments, N=544	Authentic work	146	A1	6
	Trial and error	40	A2	9
	Strategic thinking	42	A3	86
	Outcomes verify skills	56	A4	26
	Structure	101	A5	11
	Modeling	25	A6	16
	Scaffolding	69	A7	21
	Reflection	65	A8	19
			A9	2
			A10	27
			A11	4
			A12	7
			A13	8
			A14	46
			A15	18
			A16	15
			A17	7
			A18	17
			A19	13
			A20	58
			A21	46

Appendix C - Basic Levels of Self (BLoS) Model Summary

Multilevel systems theories highlight the multilevel structure of both person-in-context systems as a whole as well as persons and contexts considered as separate but interdependent multilevel systems. When considering the potential effects of context quality on SEL skill development, as well as the extent to which youth participate in their own learning, the Basic Levels of Self (BLoS) model (Roeser, Peck, & Nasir, 2006) provides a framework for understanding how different parts of the self-system play distinct roles in processing information about, and formulating goal-directed behavioral responses to, contextual opportunities and constraints. Before describing in more detail the key components of BLoS and their methodological implications, we first describe the larger scope of multilevel systems theories in which BLoS are inextricably embedded.

Healthy development in general, as well as positive youth development in particular, implicates at least three broad disciplinary foci: biological, psychological, and sociological. These disciplinary foci highlight what appear, at first glance, to be three different "levels of analysis." However, more detailed examination reveals a far more complicated picture. For example, it is popular to consider such "levels of analysis" as if the objects of study (similar to the disciplinary perspectives used to study these objects) can be neatly ordered along a single continuum constituting a unidimensional multilevel system (e.g., biological factors are nested within psychological factors are nested within social factors). Unfortunately, the "reality" of biological, psychological, and social systems is far more complicated than implied by such unidimensional multilevel system approaches (Peck, 2005, 2006, 2007, 2009). For example, among the wide variety of multilevel systems schemes found in the literature, the concept of *levels of organization* (LoOrg) is often used to frame theory and research that implicates units of analysis hypothesized to exist and function at different "levels" of the overall system being studied (cf. Anderson, 1998; Cacioppo & Berntson, 1992; Campbell, 1990; Ford & Lerner, 1992; Pattee, 1973b; Salthe, 1993).

Depending on the discipline of origin, such LoOrg are typically described in terms of units of analysis ranging from the molecular to the sociocultural (e.g., cells, tissues, and organs; or individuals, groups, and organizations). However, the full range of phenomena related to healthy development do not appear to conform to such unidimensional LoOrg schemes (for reasons discussed below), hence our ability to understand and facilitate human and social development should benefit from the development and application of theories and analytic models that better match the observed complexities. For example, biological variables such as cortisol levels, behavioral variables such as physical exercise, and social variables such as discrimination are typically ordered along a single dimension of "levels of organization" or "levels of analysis" (cf. Anderson, 1998; Cacioppo & Berntson, 1992), but close inspection of the relation between the units and levels described by these LoOrg schemes reveals internal inconsistencies

that have yet to be resolved by reference to a unidimensional levels system (Peck, 2005, 2007). For example, most of the psychological concepts developed over the past century (e.g., cognition, emotion, motivation, identity) have yet to be mapped clearly on to any version of LoOrg.

From the Multidimensional Multilevel Systems Theory (MMST) perspective used here, of which BLoS is a integral part, the problem with applying a unidimensional LoOrg scheme to human development is that (a) the range of phenomena implicated by studying healthy development do not conform to a single unidimensional multilevel system and (b) attempts to use such unidimensional systems to understand this range of phenomena imposes conceptual and statistical constraints that make it difficult or impossible to understand the complex nonlinear dynamics characterizing healthy development. For example, contemporary "multilevel" statistical models are based on the LoOrg concept of materially-nested hierarchical systems, but the full range of health-related phenomena implicated by most interdisciplinary research projects may be characterized better as functionally-nested heterarchical systems (described below) than materially-nested hierarchical systems. Understanding heterarchically-organized systems and their implications (e.g., for causal theories and statistical models) requires attending to the nature and principles of multilevel systems and applying these principles to the development of theories and statistical models. In this brief overview, however, we summarize primarily those parts of MMST necessary to highlight BLoS and its relation to our work on the SEL Challenge.

Despite the increasingly popular use of the term *multilevel*, the concept of multidimensional multilevel systems is relatively absent from mainstream research. Contemporary scientific methods and theories are typically based implicitly or explicitly on either single-level theories or unidimensional-levels-system theories. Single-level theories assume that all units of analysis exist and function on a single "level" (Cacioppo & Berntson, 1992). Unidimensional-levels-system theories assume that all units of analysis exist and function within and across a single, unidimensional series of "levels" (e.g., organs, organisms, organizations) (Peck, 2007). However, the evolution, development, and functioning of the human brain highlight both the limits of unidimensional LoOrg schemes and why additional multilevel systems appear to be necessary for understanding the relations among so-called biological, psychological, and social phenomena.

LoOrg refer explicitly to the materially-nested, hierarchical arrangement of the basic constituents of the material world (e.g., molecules, macromolecules, cells, tissues, organs, organ systems, organisms, populations). This hierarchical arrangement reflects several key features of LoOrg, including (a) units at adjacent levels vary, by orders of magnitude, in size and timescale such that "there is a real break, or boundary, in the world at every jump *across* a level" (Salthe, 1985, p. 122); and (b) interactions *within* levels are generally stronger than interactions *across* levels (Simon, 1996); and (c) causal effects generally do not "skip over" levels but "travel through" adjacent levels (aka, the principle of

"nontransitivity of effects across levels;" Salthe, 1993, p. 45). For example, the effects of cells on organismic functions (e.g., behavior) are generally mediated fully by organs and their functions.

If all health-related biological, psychological, and social phenomena conformed to the principles of LoOrg, our theoretical and modeling tasks would be relatively simple. However, evolutionary processes have yielded a series of highly consequential forms of biological complexity (e.g., hypothalamus, amygdala, neocortex) that are known, collectively, as the human brain. Although the brain has been described as a hierarchically-organized multilevel system, or neuroaxis (Berntson & Cacioppo, 2008; Bowden & Martin, 1995; Bronson, 1965; Derryberry & Tucker, 1991; Jackson, 1884; Lewis & Todd, 2007; MacLean, 1990), none of these descriptions conform to the concept of LoOrg used here (Peck, 2007). For example, although the neocortex and the brainstem are at opposite ends of the neuroaxis, they do not correspond to distinct LoOrg (e.g., because they consist of material from exactly the same LoOrg, such as cells and tissues).

In addition, the brain has also been described as a heterarchically-organized (e.g., functionally nested) multilevel system (Berntson, Boysen, & Cacioppo, 1993; Berntson & Cacioppo, 2008; Lewis & Todd, 2007; MacLean, 1990; Mesulam, 2000) and these functionally-nested multilevel brain systems have been referred to as *basic levels of self* (BLoS; Roeser et al., 2006) and *levels of representation* (LoRep) (Peck, 2007). In contrast to materially-nested LoOrg, functionally-nested multilevel systems have been described as "control hierarchies" (Pattee, 1973a, p. 75), or "command hierarchies" (Salthe, 1993, p. 45), that typically function as parallel distributed processes (Holland, 1995; McClelland & Rogers, 2003). For example, although the heart is not nested materially within the brain, the brain and the heart influence each other by passing "signals" through intermediary physiological systems. In stark contrast to the nontransitive (i.e., through adjacent levels) relations across LoOrg, LoRep (e.g., BLoS) are characterized by causal effects that can run directly (i.e., transitively) from any level to any other level (Berntson & Cacioppo, 2008; Salthe, 1985). (Note, also, that LoRep have a contextual manifestation that is not discussed there; see Peck, 2007, p. 1141. on "contextual levels of representation").

In these terms, although the neocortex and brainstem cannot be at different "levels of organization" because they both involve the same orders of LoOrg complexity (e.g., cells, tissues, organs), the neocortex and brainstem *are* at different LoRep because, in BLoS terms, each level of self (e.g., iconic vs. symbolic) corresponds to a qualitatively distinct way of representing information, each of which was achieved over the course of millions of years of biological and social evolution. In other words, from the perspective of biological evolution, there are a relatively clear set of BLoS marked by increasingly flexible processing capacities as we move up the neuroaxis from the brain stem to the neocortex (Berntson & Cacioppo, 2003; Bronson, 1965; Derryberry & Tucker, 1991; Herrick, 1949; Lewis & Todd, 2007; MacLean, 1990; Schneirla, 1949). Further, units within each BLoS tend to

represent the same organismic and environmental conditions in qualitatively different ways (cf. Grene, 1988; McClelland, McNaughton, & O'Reilly, 1995; Pattee, 1973b; Salthe, 1985; Sherry, 2006). For example, *symbolic* representations (e.g., beliefs) centered in the neocortex differ qualitatively from *iconic* representations (e.g., schemas) centered in the limbic system, and these different kinds of representations can implicate conflicting or converging behavioral responses to the same environmental conditions (cf. Boysen & Berntson, 1995; Epstein, 1990; Flannelly, Koenig, Galek, & Ellison, 2007; McClelland, Koestner, & Weinberger, 1989; Roeser et al., 2006; Schultheiss, 2001).

We use the term iconic representation where referring to the species-typical, sub-cortical capacity for the development of "sensory-affective and affective-motor schemas that become increasingly differentiated and integrated into higher order sensory-affective-motor scripts as a function of direct experience with the immediate environment" (Roeser et al., 2006, p. 402). In these terms, "within the iconic level, schemas are the most fundamental unit of information, and simple sensory-affective and affective-motor schemas combine to form more complex sensory-affective-motor schemas, which combine to form more complex sensory-affective-motor scripts" (Peck, 2007, p. 1139). We use the term symbolic representation where referring to the species-typical, neo-cortical capacity for the development of "declarative (e.g., beliefs about things) and procedural (e.g., beliefs about how to do things) knowledge (primarily verbal in nature) that becomes increasingly differentiated...and integrated...over time" (Roeser et al., 2006, p. 403). In these terms, "within the symbolic level of representation, beliefs are the most fundamental unit of information, and basic beliefs differentiate and integrate across time to form more complex belief systems such as attitudes that combine to form goals that combine to form plans" (Peck, 2007, p. 1139).

Principles associated with BLoS (e.g., transitive interactions and parallel distributed processes) can be used to understand the behavioral implications of simultaneously operating factors such as (a) beliefs about the relevance of program-offering content (hypothesized to be represented primarily within the symbolic system) and (b) social rejection sensitivity (hypothesized to be represented primarily within the iconic system). For example, two youth with similarly favorable attitudes toward program-offering content (and similarly favorable context quality) may vary in their tendency to remain fully engaged in constructive interactions with program staff as a function of their differential sensitivity to social rejection. In more general terms, such cross-level BLoS dynamics have been described in terms of "dual-process" models of self-regulation and impulse control (Ayduk et al., 2008; Deutsch & Strack, 2006; Hoffman, 2009; Kahneman & Frederick, 2007; Kochanska & Knaack, 2003; Rodriguez, Mischel, & Shoda, 1989; Rothbart, 2007; Schultheiss, 2001; Sherman et al., 2008; Sloman, 1996; Smith & DeCoster, 2000; Strack & Deutsch, 2004), but few of these models are based on well-developed multilevel systems theories. Consequently, we believe that understanding the relations among biological, psychological, and

social factors associated with social and emotional learning will require (a) distinguishing hierarchical LoOrg from heterarchical BLoS and (b) attending explicitly to more than simply "dual" processes (Peck, 2007).

In order to illustrate some distinctions between hierarchical and heterarchical systems and describe in more detail several different kinds of intraindividual structures and processes and their relations to each other and behavior, we will now show (in Figure C-1, below) and describe a more elaborated version of the right-hand side of Figure 3 (from Chapter Three), which highlighted three key components of the multilevel intraindividual person system (i.e., symbolic, iconic, and phenomenological representation) and their relations to the multilevel context system. In Figure C-1, we show how symbolic, iconic, and phenomenological representations are related to other components of the multilevel intraindividual person system; particularly, sensory input, "controlled" reflective processing, and behavioral output.

The upward-pointing, curved, blue arrow at the bottom left of Figure C-1 indicates the path followed by information from the immediate physical and social context that is available to the intraindividual person system. This contextual information is processed via the five human somatic sensory systems (e.g., visual and auditory) as a set of parallel inputs into the brain stem, depicted here as the temperamental representation system. The temperamental system, as used here, refers to the species-typical brain structures and processes responsible for baseline levels of physiological arousal, sensitivity, and reactivity (Buss & Plomin, 1984; Roeser, Peck, Nasir, et al., 2006; Thayer, 1989). From the brainstem, sensory pathways fan out and on, as parallel processes, to a wide range of limbic (i.e., iconic) and cortical (i.e., symbolic) regions responsible for the initial decoding and integration of sensory input.

During this initial sensory-processing, percept-generation, and meaning-making phase, which has been referred to as *primary appraisal* and lasts for approximately 400 milliseconds, none of the information being processed is available to conscious awareness (shown at the top, center of Figure C-1). In other words, primary appraisal processes take hundreds of milliseconds and occur completely outside of conscious awareness. Despite this lack of awareness, however, iconic and symbolic representations relevant to the pattern of incoming sensory information are being activated, and spreading activation within and between the iconic and symbolic systems begins to elaborate the meaning of the incoming sensory information in relation to physiological needs (e.g., hunger), experiential memories (e.g., past trauma), and personal goals (e.g., to learn).

Once activated, these *affectively-charged* schemas and *valenced* beliefs generate the initially unconscious emotional and behavioral response dispositions that are input simultaneously (i.e., in parallel) to the phenomenological representation system. After symbolic or iconic information is activated (or, *re-represented*) phenomenologically—which generally takes at least 400 milliseconds from the time

of initial sensory input—thoughts, feelings, and behavioral-response dispositions are accessible to conscious awareness, even if they do not become the focus of conscious awareness. At this stage, behavioral-response dispositions can be expressed as overt behavior, regardless of whether or not those response dispositions have become the focus of awareness. (Note: The distinction between phenomenologically-activated representations that are in versus out of the focus of awareness is indicated by the dashed line that runs horizontally through the middle of the box labeled, "Phenomenological Representation"). Overt behavior generated automatically by activated schemas and beliefs is generally referred to as habitual or impulsive behavior.

However, once activated phenomenologically, thoughts, feelings, and behavioral-response dispositions that become the focus of conscious but *passive awareness* (as indicated by the upward-pointing, curved, blue arrow at the top and center of Figure C-1) provide the raw material for reflective processing, or what has been referred to as *secondary appraisal*. It is at this point, for example, that behavioral-response dispositions can potentially be suppressed or inhibited by active awareness (Roeser & Peck, 2009), as indicated by the downward-pointing, curved, blue arrow at the top and center of Figure C-1. Active awareness has been described variously as controlled processing (Shiffrin & Schneider, 1977), effortful control (Rothbart, 2007), and executive attention (Posner & Rothbart, 2000). For example, if seeing cookies in the lobby triggers the urge to consume cookies, actively focusing awareness on a newly formed goal to replace sugar-filled foods with vegetables can inhibit the dominant cookie-consumption response. In these terms, the highest form of youth participation in their own development can be described in terms of actively focusing awareness on specific thoughts and feelings so that they can further develop (e.g., revise or extend) their understanding of themselves and their world.

Similarly, consistent with James' (1890) distinction between *I* and *Me*, planning and reflection are secondary appraisal processes that can be characterized as *active awareness* (i.e., *I*) focusing selectively on specific parts of the stream of phenomenological representations (i.e., *Me*). Whereas primary appraisal involves the relatively rapid and automatic activation of previously learned beliefs and schemas, some of which may become the focus of passive awareness (i.e., thoughts or feelings generated by activated beliefs or schemas are noticed consciously), secondary appraisal involves focusing awareness selectively and repeatedly on subsets of phenomenological content in efforts to activate or construct new content (e.g., a new plan) and encode (i.e., store) that new content in long-term memory (e.g., symbolic representation). Actively focusing awareness both reactivates the selected content for further conscious processing and allows the activation levels of unselected content to diminish and eventually fade completely unless that content is reactivated by active awareness or by spreading activation due either to related phenomenological content that is focused on or to the ongoing stream of relevant environmental stimulus information that re-triggers that content. Engaging in such secondary

appraisal processes (i.e., *Type 2 agency*) allows individuals to activate, construct, and encode new goals and plans that are then available to become activated automatically by relevant environmental stimuli, thereby affording them an active, agentic form of self-regulation that may otherwise appear to be passive (i.e., *Type 1 agency*).

Finally, whereas we distinguish explicitly between awareness and the phenomenological objects of awareness, similar to how James (1890) distinguished between *I* and *Me*, most dual process theories in psychology focus instead on the distinction between what they term, for example, explicit versus implicit (Schacter, 1987), rational versus experiential (Epstein, 1990), analytic versus holistic (Jung, 1923), or reflective versus impulsive (Lieberman, Gaunt, Gilbert, & Trope, 2002) processes. From a BLoS perspective, the first term used in each of these cases refers primarily to secondary appraisal processes involving symbolic representations, whereas the second term refers primarily to secondary appraisal processes involving iconic representations. In other words, most dual process theories focus implicitly on the distinction between symbolic and iconic representations, not on the distinction between awareness and the phenomenologically-activated objects of awareness. Given that we view input to phenomenological representation as a parallel distributed process involving whatever temperamental, iconic, and symbolic content is activated simultaneously at any given time, we find it useful to distinguish among more than two different aspects of the overall intraindividual information processing system. Using this more differentiated model allows us, for example, to consider the extent to which (a) conscious thoughts and feelings can be informed simultaneously by automatically activated iconic (e.g., affective) and symbolic (e.g., valenced) memories (aka, "knowledge"), and (b) the behavioral implications of such activated memories or knowledge can be influenced consciously by sustaining or shifting the focus of awareness on or to select aspects of that knowledge. In this view, although youth's thoughts, feelings, and behavior occurring at the point of service tend to be governed heavily by past experiences (including recently activated emotional trauma), or primary appraisal processes, the potential for learning alternative forms of thought, feeling, and behavior is highly contingent on contextual scaffolding (e.g., staff behavior) designed to promote conscious engagement with pre-existing forms of thought, feeling, and behavior and their modification through both co-regulated and increasingly self-regulated secondary appraisal processes.

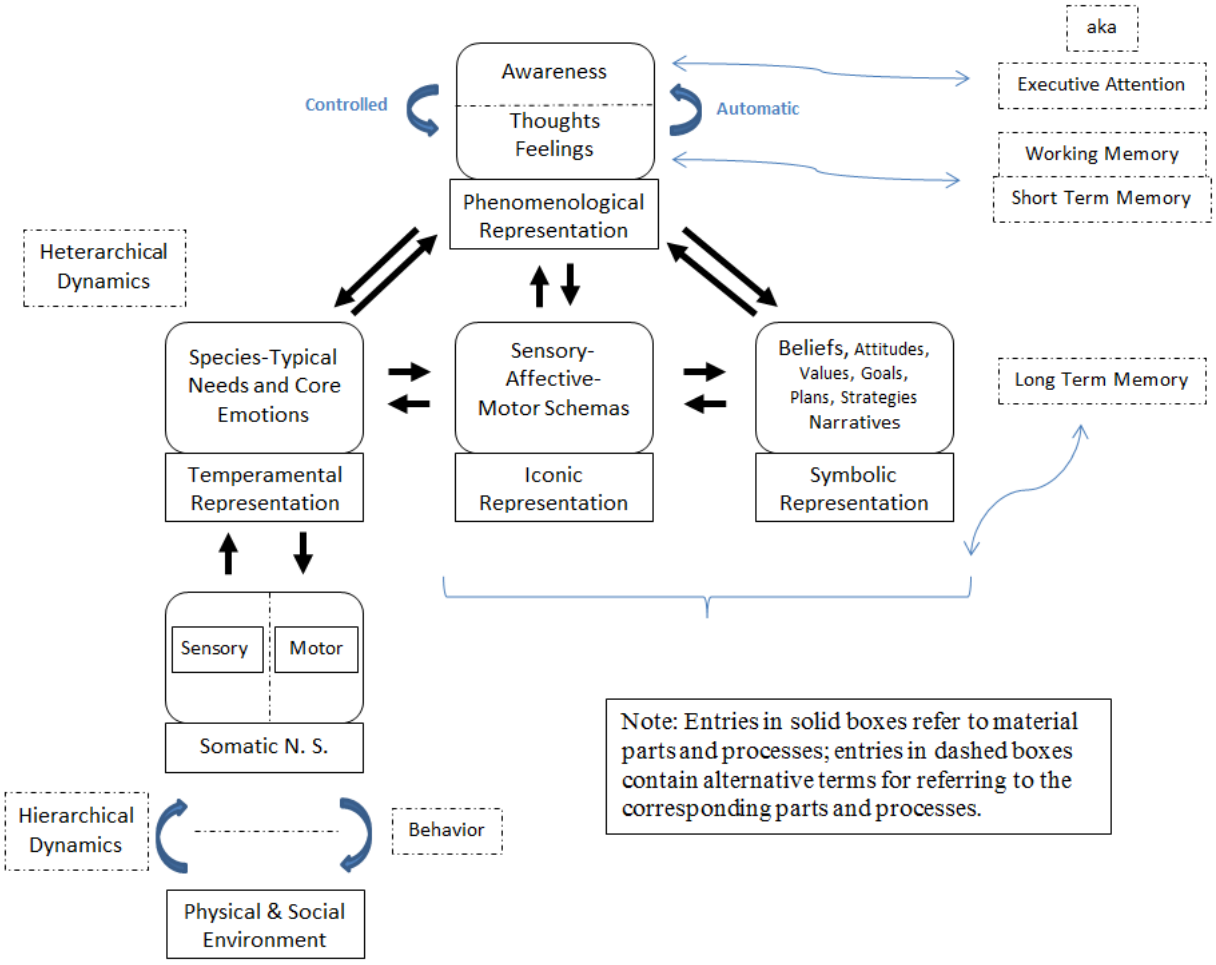

Figure C-1. **Multilevel Intraindividual Person System: Contents, Structures, and Processes**

Appendix D - Scale and Item Descriptives for All Measures, All Time Points

Organizational Measures

Table D-1. Accountability

PROMPT: How true are the following statements regarding accountability for quality services? (1=Almost never true, 3=True about half of the time, 5=Almost always true)	*Mean (SD)*
Accountability SEL (N=8)	**4.55 (.482)**
Our program is held accountable for the quality, including point-of-service quality (i.e., relationships, instruction)	4.53
Our program is routinely monitored by higher level administrators	4.44
In our program all staff are familiar with standards of quality	4.59
Collaboration across programs is strongly encourages by network administrators	4.63
Site supervisors in our network share a similar definition of high quality services	4.55
Accountability De-identified Reference Database (N=48)	**4.61 (.541)**

Data Source: SEL Staff Survey

Table D-2. Staffing Model

Prompt: Please rate the extent to which the following statements are true for staff in your program. (1 = Almost never true of staff, 3 = True for about half of staff, 5 = Almost always true of staff)	*Mean (SD)*
Capacity SEL (N=8)	**4.41 (.591)**
Staff come to the program with adequate training or experience	3.91
Staff stay at our program for a long time	4.48
We have enough staff and/or student-to-staff ratios are good	4.65
New staff get an adequate orientation	4.26
Staff have enough time to attend meetings or do planning	4.50
Staff are designing and delivering activities consistent with program goals and objectives for students	4.65
Capacity De-identified Reference Database (N=42)	**4.44 (.423)**

Data Source: SEL Staff Survey

Table D-3. School Day Content

PROMPT: Please indicate the proportion of program youth for which the following statements are true in your program. (1=Almost none, 3=About half, 5=Almost all).	*Mean (SD)*
School Day Content SEL (N=8)	**2.72 (.895)**
I help manage formal 3-way communication that uses the afterschool program to link students' parents with school-day staff and information	3.15
I participate in meetings for afterschool and school day staff where linkages between the school day and afterschool are discussed and/or where academic progress of individual students are discussed	3.02
I participate in parent-teacher conferences to provide information about how individual students are faring in the afterschool program	1.98
School Day Content De-identified Reference Database (N=102)	**3.50 (.672)**

Data Source: SEL Challenge Staff Survey

Table D-4. Horizontal Communication

PROMPT: Please respond to the following statements by selecting the number that most nearly represents how often the following practices occur in your program: (1=Never, 3=Every few months, 5=At least weekly).	*Mean (SD)*
Horizontal Communication SEL (N=8)	**4.40 (.420)**
I co-plan with another member of staff	4.77
I discuss teaching problems or practices with another staff member	4.89
A co-worker observes my session and offers feedback about my performance	3.86
I work on plans for program policies or activities with other staff	4.44
I observe a co-worker's session and provide feedback about their performance	4.00
Horizontal Communication De-identified Reference Database (N=170)	**3.75 (.756)**

Data Source: SEL Challenge Staff Survey

Table D-5. Vertical Communication

PROMPT: Please respond to the following statements by selecting the number that most nearly represents how often the following practices occur in your program: (1=Never, 3=Every few months, 5=At least weekly).	*Mean (SD)*
Vertical Communication SEL (N=8)	**4.28 (.669)**
My supervisor gives me helpful feedback about how I work with youth	4.16
My supervisor is visible during the offerings that I lead or co-lead	3.68
My supervisor knows what I am trying to accomplish with youth	4.73
My supervisor challenges me to innovate and try new ideas	4.32
My supervisor makes sure that program goals and priorities are clear to me	4.50
Vertical Communication De-identified Reference Database (N=171)	**3.78 (.759)**

Data Source: SEL Challenge Staff Survey

Table D-6. Job Satisfaction

PROMPT: Please rate the extent to which the following statements are true for you (1=Almost never true, 3=True about half of the time, 5=Almost always true).	*Mean (SD)*
Job Satisfaction SEL (N=8)	**4.22 (.556)**
In most ways, this job is close to my ideal	4.44
The condition of my current job is excellent	4.06
I am satisfied with this job	4.21
If I could change my career so far, I would not change anything	4.17
Job Satisfaction De-identified Reference Database (N=172)	**3.69 (.536)**

Data Source: SEL Challenge Staff Survey

Table D-7. Climate

PROMPT: Please respond to the following statements by selecting how often the following practices occur in your program. 1=Rarely, 2=Sometimes, 3=Often, 4=Most of the time, 5=Always	*Mean (SD)*
Climate SEL (N=8)	**1.52 (.770)**
The amount of work I have to do prevents me from doing a good job	
Climate De-identified Reference Database (N=68)	**1.67 (.603)**

Data Source: SEL Challenge Staff Survey; 2007 YPQI W1 Staff Survey

Point-of-Service Measures

Table D-8. Growth and Mastery

PROMPT: Please indicate the proportion of students in your program for which the following goal statements are true (1=Almost none, 3=About half, 5=Almost all).	*Mean (SD)*
Growth and Mastery Skills SEL (N=8)	**4.53 (.177)**
We will expose students to experiences which are new for them	4.56
Students will have responsibilities and privileges that increase over time	4.33
Students will work on group projects that take more than five sessions to complete	4.52
All participating children and youth will be acknowledged for achievements, contributions and responsibilities	4.79
Students will identify a skill/activity/pursuit that the feel they are uniquely good at	4.41
Growth and Mastery Skills De-identified Reference Database (N=103)	**3.74 (.553)**

Data Source: SEL Challenge Staff Survey

Table D-9. Youth Program Governance

PROMPT: Please indicate the proportion of program youth for which the following goal statements are true. "1" = Almost none, "3" = About half, "5" = Almost all	*Mean (SD)*
Youth Program Governance SEL (N=8)	**3.79 (.674)**
Youth have opportunities to begin their own projects, initiatives, and enterprises	4.19
Youth are involved in selecting the content or purposes of activities and sessions	4.14
Youth contribute to the design, appearance, and aesthetics of the physical space	3.04
Youth Program Governance De-identified Reference Database (N=45)	**2.85 (.908)**

Data Source: SEL Staff Survey

Table D-10. Youth Organizational Governance

PROMPT: Please indicate the proportion of program youth for which the following goal statements are true. "1" = Almost none, "3" = About half, "5" = Almost all	*Mean (SD)*
Youth Organizational Governance SEL (N=8)	**2.76 (.651)**
Youth are involved in hiring new staff	2.74
Youth are involved in deciding how the organization's budget is spent	1.81
Our students experience afterschool sessions led or supported by PAST AFTERSCHOOL STUDENTS who are paid staff or volunteers	3.14
Our students help to provide public recognition of community volunteers, organizations, and businesses that contribute to the afterschool program	3.36
Youth Organizational Governance De-identified Reference Database (N=58)	**2.50 (.945)**

Data Source: SEL Staff Survey

Table D-11. Curriculum Planning

PROMPT: Please indicate the proportion of program youth for which the following goal statements are true. "1" = Almost none, "3" = About half, "5" = Almost all	*Mean (SD)*
Curriculum Planning SEL (N=8)	**4.12 (.403)**
The session is planned in advance and written out in a lesson plan format	4.47
The session is targeted at specific learning goals for the individual student, or for a school curriculum target or for a specific state standard	4.11
The session builds upon steps taken in a prior activity or session	4.66
The session is based on recent feedback from students about where they need support	4.16
The session combines academic content with the expressed interests of students	3.22
Academic Planning De-identified Reference Database (N=97)	**4.05 (.611)**

Data Source: SEL Staff Survey

Belief Measures

Table D-12. Emotion Management Belief Scales

	Time 1	*Time 2*	*Time 3*
Emotion Management Domain	3.56 (.49)	3.53 (.43)	3.80 (.53)
Optimism	**3.27 (.64)**	**3.40 (.64)**	**3.54 (.69)**
In uncertain times, I usually expect the best	3.29	3.59	3.61
If something can go wrong for me, it will	3.93	3.01	3.22
I'm always optimistic about my future	3.81	3.82	4.01
I hardly ever expect things to go my way	2.29	2.95	3.20
I rarely count on good things happening to me	3.08	3.26	3.36
Overall, I expect more good things to happen to me than bad	3.78	3.93	4.09
Identification of Emotions	**3.77 (.74)**	**3.90 (.64)**	**3.99 (.67)**
I am aware of my emotions as I experience them	3.85	4.03	4.06
I easily recognize emotions as I experience them	3.89	3.93	4.10
I am aware of the non-verbal messages that other people send	3.86	3.90	4.09
I know what other people are feeling just by looking at them	3.48	3.76	3.66
Reappraisal Factor	**3.64 (.74)**	**3.78 (.74)**	**3.90 (.73)**
I control my emotions by changing the way I think about the situation I'm in	3.59	3.81	3.92
When I want to feel less negative emotion (such as sadness or anger), I change the way I'm thinking about the situation	3.47	3.72	3.90
When I want to feel more positive emotion (such as joy or amusement), I change the way I'm thinking about the situation	3.71	3.82	4.00
When I want to feel more positive emotion (such as joy or amusement), I change what I'm thinking about	3.76	3.88	4.00
When I want to feel less negative emotion (such as sadness or anger), I change what I'm thinking about	3.68	3.68	3.91
When I'm faced with a stressful situation, I make myself think about it in a way that helps me stay calm	3.55	3.82	4.00

Data Source: SEL Challenge Youth Belief Surveys

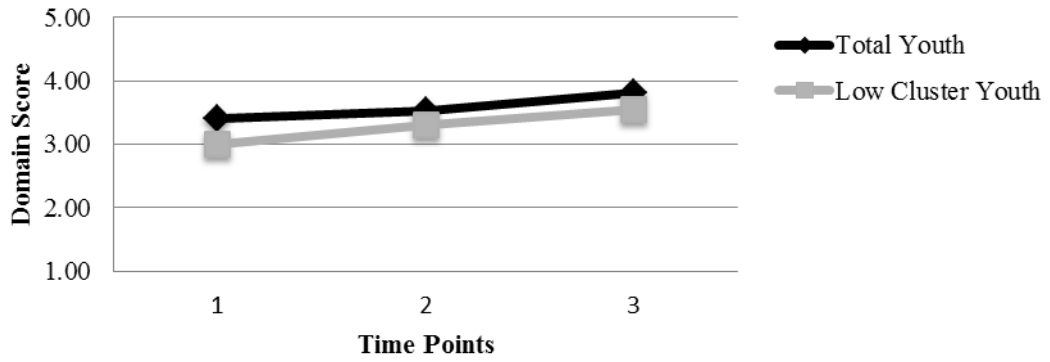

Figure D-1. **Emotion Management Belief across Three Time Points with Low Cluster Comparison**

Table D-13. Empathy Belief Scales

	Time 1	*Time 2*	*Time 3*
Empathy Domain	3.87 (.67)	3.89 (.65)	4.01 (.77)
Adolescent Empathy	**3.87 (.67)**	**3.89 (.65)**	**4.01 (.77)**
I feel bad when someone gets their feelings hurt	3.63	3.59	3.83
I understand how those close to me feel	3.82	3.93	4.06
It is important to me to understand how other people feel	3.99	3.84	4.03
I am happy when others succeed	3.94	4.04	4.13

Data Source: SEL Challenge Youth Belief Surveys

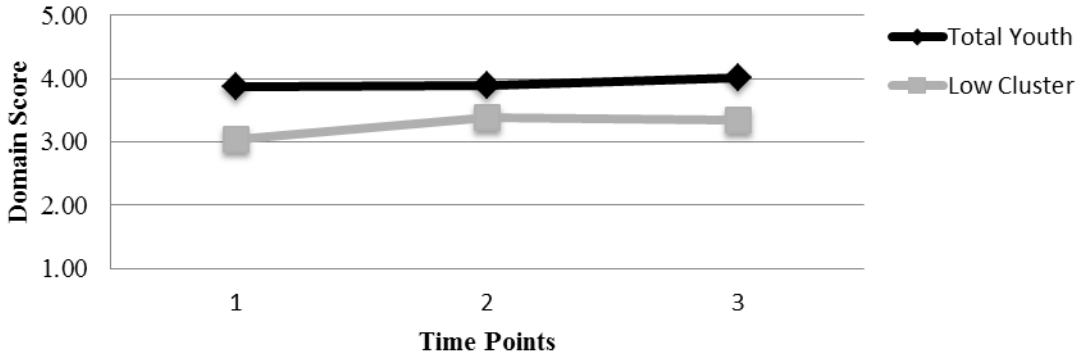

Figure D-2. **Empathy Belief across Three Time Points with Low Cluster Comparison**

Table D-14. Teamwork Belief Scale

	Time 1	*Time 2*	*Time 3*
Teamwork Domain	3.99 (.57)	4.06 (.52)	4.15 (.65)
Adolescent Social Competency	**3.99 (.57)**	**4.06 (.52)**	**4.15 (.65)**
I avoid making other kids look bad	3.72	3.77	3.80
If two of my friends are fighting, I find a way to work things out	3.91	3.93	4.14
When I work in school groups, I do my fair share	4.25	4.11	4.42
Do you get along well with people of different races, cultures, and religions?	4.45	4.42	4.63
Do you listen to the ideas of others?	4.29	4.14	4.44
Do you control your anger when you have a disagreement with a friend?	3.63	3.80	4.06
Can you discuss a problem with a friend without making things worse?	3.73	3.89	4.10
Do you follow the rules when you are at a park, theater, or sports event?	4.1	4.16	4.37
Do you respect other points of view, even if you disagree?	4.00	3.96	4.30

Data Source: SEL Challenge Youth Belief Surveys

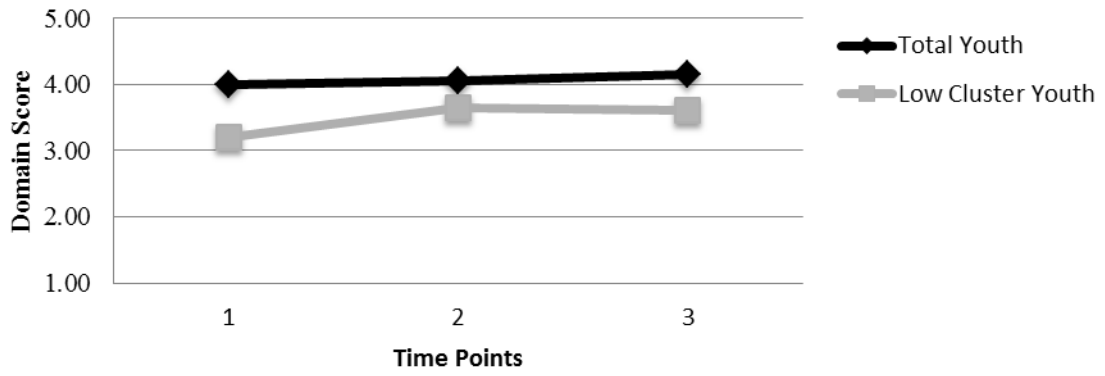

Figure D-3. **Teamwork Belief across Three Time Points with Low Cluster Comparison**

Table D-15. Responsibility Belief Scales

	Time 1	*Time 2*	*Time 3*
Responsibility Domain	3.74 (.57)	3.80 (.60)	3.95 (.58)
Adolescent Diligence and Reliability	**3.74 (.57)**	**3.80 (.60)**	**3.95 (.58)**
Do you work harder than other people?	3.75	3.83	4.01
Do you do as little work as you can get away with?	3.44	3.30	3.68
Do you finish the tasks you start?	4.02	3.97	4.30
Is it hard for you to finish the tasks you start?	3.32	3.21	3.66
Do you give up when things get difficult?	3.59	3.67	4.04
Can people count on you to get tasks done?	3.95	4.08	4.33
Do you do things that you say you are going to do?	3.89	4.04	4.24

Data Source: SEL Challenge Youth Belief Surveys

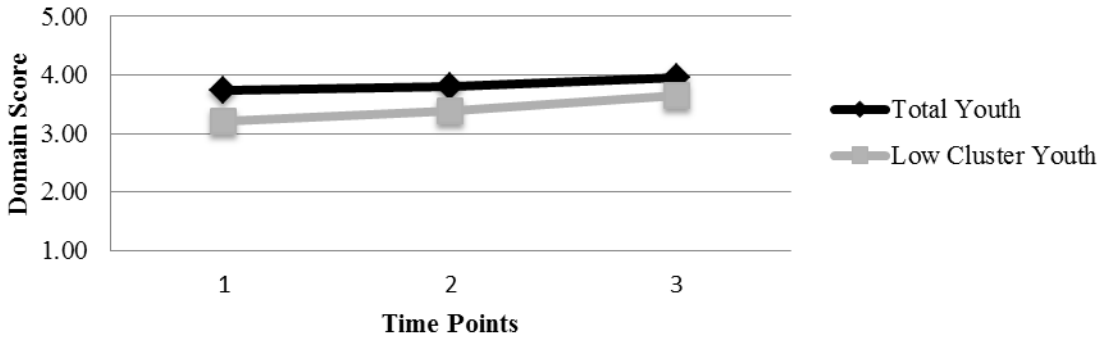

Figure D-4. **Responsibility Belief across Three Time Points with Low Cluster Comparison**

Table D-16. Initiative Belief Scales

	Time 1	*Time 2*	*Time 3*
Initiative Domain	**3.90 (.64)**	**4.13 (.58)**	**4.28 (.58)**
Adolescent Initiative Taking	**3.99 (.69)**	**4.06 (.65)**	**4.17 (.63)**
I am willing to risk failure to reach my goals	3.59	3.73	3.99
When I work at something, I care about doing my best	4.23	4.33	4.49
I like coming up with new ways to solve problems	3.97	3.96	4.23
I am a leader, not a follower	4.22	4.20	4.52
Adolescent Purpose	3.76 (.80)	4.23 (.70)	4.38 (.68)
My life has no meaning	4.16	4.26	4.47
My life will make a difference in the world	3.28	4.15	4.35
I am doing things now that will help me achieve my purpose in life	4.04	4.23	4.53

Data Source: SEL Challenge Youth Belief Surveys

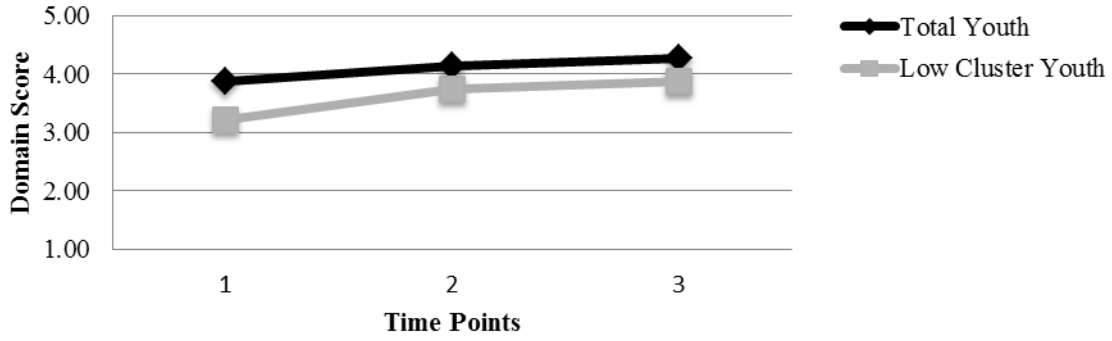

Figure D-5. Initiative Belief across Three Time Points with Low Cluster Comparison

Table D-17. Problem Solving Belief Scales

	Time 1	*Time 2*	*Time 3*
Problem Solving Domain	***3.82 (.60)***	***3.89 (.55)***	***4.03 (.55)***
Adolescent Goal Orientation	**3.87 (.69)**	**3.94 (.57)**	**4.07 (.59)**
I develop step-by-step plans to reach my goals	3.46	3.54	3.87
I have goals in my life	4.34	4.38	4.63
If I set goals, I take action to reach them	4.01	4.10	4.36
It is important to me that I reach my goals	4.33	4.32	4.61
I know how to make my plans happen	3.87	3.99	4.09
Do you make plans to achieve your goals?	3.82	3.99	4.18
Do you have trouble figuring out how to make your goals happen?	3.17	2.98	3.18
Problem Solving Strategies	**3.76 (.65)**	**3.83 (.67)**	**3.98 (.67)**
I try to understand the purpose for a task/project before I get started	3.94	4.01	4.21
I try to imagine all of the parts of a task/project that I have to complete	3.92	3.90	4.18
Pictures, diagrams or graphs help me understand the parts of a task/project	3.96	4.03	4.30
I select and organize information before getting started on a task/project	3.78	3.88	3.93
I start a task/project by brainstorming lots of ideas about how to do it	3.82	3.72	3.94
I start a task/project by thinking about the end result and then work backward through the steps to get there	3.20	3.31	3.51

Data Source: SEL Challenge Youth Belief Surveys

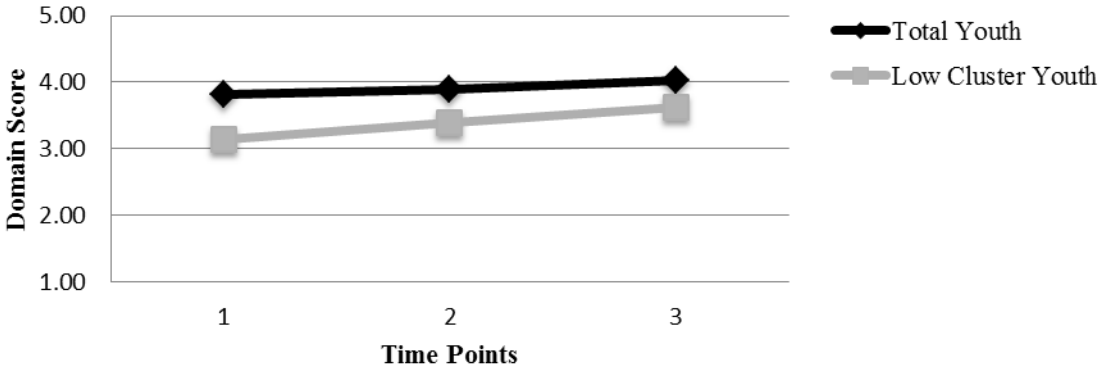

Figure D-6. **Problem Solving Belief across Three Time Points with Low Cluster Comparison**

Behavior Measures

Table D-18. Empathy Behavior Scales

	Time 1	*Time 2*	*Time 3*
Empathy Domain	3.25 (.54)	3.65 (.47)	3.86 (.34)
Value own/others perspectives	**3.25 (.54)**	**3.65 (.47)**	**3.86 (.34)**
Use inclusive language and socially preferred terms	3.26	3.76	3.97
Correct others' bias or use of exclusive language	2.77	3.3	3.93
Help ensure others are heard	3.23	3.67	3.55
Share personal information at an appropriate level and at an appropriate time	3.37	3.58	3.88
Reflect appropriate tone, gesture, feeling, and pacing during sensitive discussion	3.44	3.84	4
Can explain biases (e.g., adultism, sexism, ableism, classism, racism).	3.3	3.79	3.95

Data Source: SEL Challenge Staff Rating of Youth Behavior

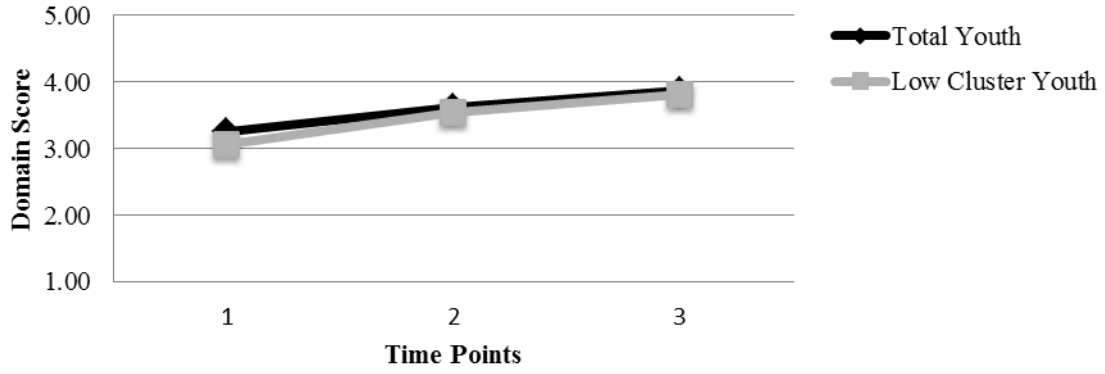

Figure D-7. **Empathy Behavior across Three Time Points with Low Cluster Comparison**

Table D-19. Emotion Management Behavior Scales

	Time 1	*Time 2*	*Time 3*
Emotion Management Domain	3.01 (.49)	3.58 (.44)	3.73 (.40)
Identify positive and negative emotions (e.g., excitement, anger, worry, joy)	**2.94 (.56)**	**3.65 (.54)**	**3.73 (.48)**
Identifies and names emotions	3.14	3.75	3.77
Has a wide descriptive vocabulary for emotions	2.79	3.65	3.74
Describes own emotional needs	2.88	3.55	3.67
Reason about causes and uses of emotion	**2.77 (.58)**	**3.45 (.34)**	**3.59 (.44)**
Identifies causes/triggers of emotion	2.81	3.52	3.70
Can guide self and others to address stress or extreme emotions	2.73	3.38	3.49
Manages emotions for functional purpose	**3.15 (.49)**	**3.58 (.49)**	**3.78 (.45)**
Checks for misunderstanding when negative (e.g., anger, frustration) emotions occur	2.63	3.36	3.54
Manages positive emotion (e.g., elation, pride) with confidence that doesn't belittle or exclude others	3.58	3.57	3.98
Maintains composure to constructively work through conflict or disagreement to maintain progress on task	3.11	3.37	3.84
Actively reaches out to others when they have emotional difficulties	2.96	3.64	3.52
Accepts feedback non-defensively, sees kernel of truth, restates with understanding to giver of feedback	3.12	3.64	3.77
Allows others to express and take responsibility for their own emotions (e.g., doesn't speak for them or immediately "rescue")	3.36	3.72	4.03

Data Source: SEL Challenge Staff Rating of Youth Behavior

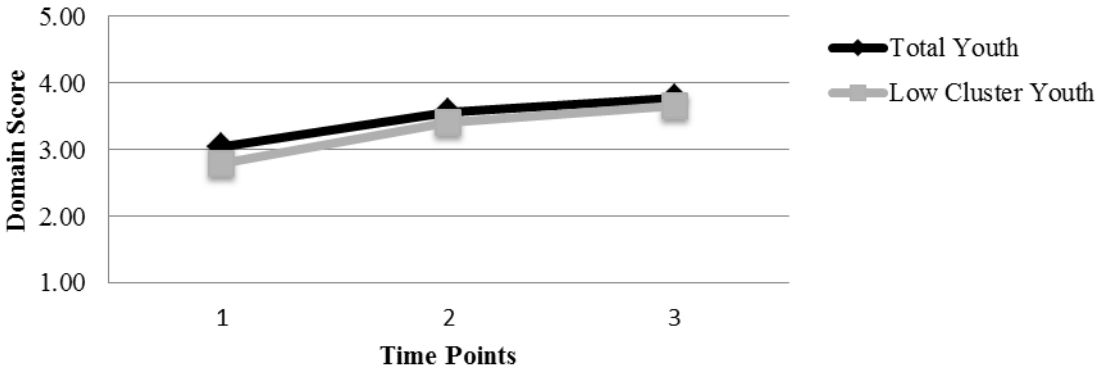

Figure D-8. **Emotion Management Behavior across Three Time Points with Low Cluster Comparison**

Table D-20. Teamwork Behavior Scales

	Time 1	*Time 2*	*Time 3*
Teamwork Domain	3.40 (.46)	3.75 (.41)	4.95 (.53)
Team Communication	**3.56 (.47)**	**3.93 (.44)**	**4.04 (.39)**
Clearly articulate point of view and/or ask follow-up questions for clarification	3.33	3.85	3.97
Communicate without dominating, interrupting, or showing disrespect for others' ideas	3.6	3.9	4.06
Demonstrate listening/interest through body language (e.g., posture, eye contact)	3.61	3.9	4.06
Upon disagreement, acknowledge and express interest in other's point of view	3.06	3.62	3.72
Refrain from gossip when having a problem and talk with the person directly	3.4	3.85	3.95
Show respect to group leaders	4.34	4.41	4.48
Supports action towards team goals	**3.27 (.41)**	**3.58 (.42)**	**3.82 (.39)**
Contribute ideas, skills, and commitment to group tasks	3.63	3.91	4.07
Offer to help those having trouble completing tasks	3.2	3.4	3.85
Monitor team progress on a task	3.05	3.53	3.66
Encourage youth and staff to be accountable for their roles and commitments	3.14	3.5	3.71

Data Source: SEL Challenge Staff Rating of Youth Behavior

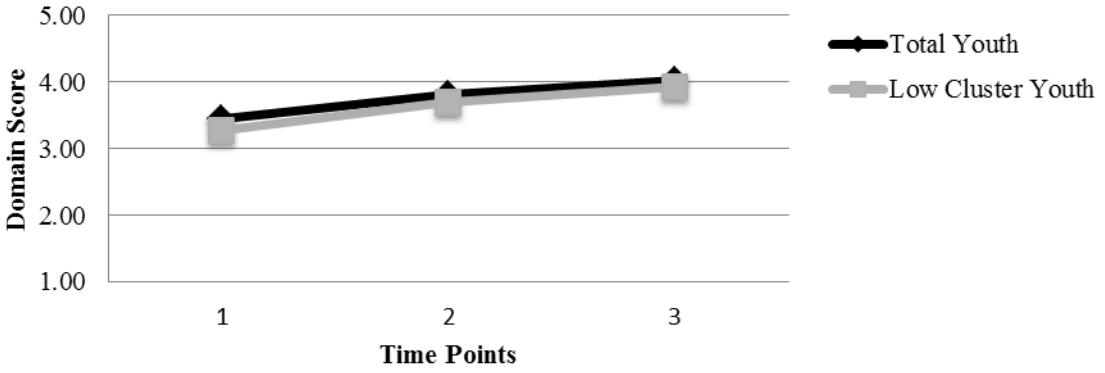

Figure D-9. **Teamwork Behavior across Three Time Points with Low Cluster Comparison**

Table D-21. Responsibility Behavior Scales

	Time 1	*Time 2*	*Time 3*
Responsibility Domain	2.70 (.35)	3.52 (.42)	3.71 (.33)
Fulfills Roles and Commitments	**3.15 (.42)**	**3.56 (.36)**	**3.72 (.32)**
Fulfill task role(s) with minimal supervision	3.19	3.57	3.61
Work toward mastery or excellence in fulfilling the role (e.g., better than compliance)	3.14	3.56	3.75
Acknowledge mistakes and be willing to address mistakes through action	3.22	3.73	3.92
Volunteer for additional tasks beyond the assigned role to further group goals	3.04	3.37	3.6
Adjusts and Negotiates Roles	**2.25 (.33)**	**3.48 (.51)**	**3.69 (.33)**
Make suggestions that would improve the role or overall task	3.27	3.67	3.88
Seek timely help from other youth or staff when roles become too challenging	2.97	3.46	3.7
Negotiate with staff/supervisor to adjust expectations for redefined commitments and deadlines	2.79	3.29	3.49

Data Source: SEL Challenge Staff Rating of Youth Behavior

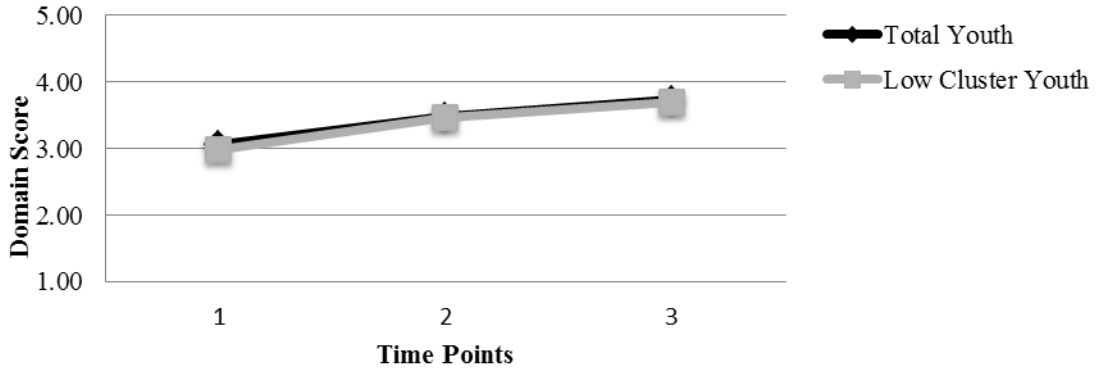

Figure D-10. **Responsibility Behavior across Three Time Points with Low Cluster Comparison**

Table D-22. Initiative Behavior Scales

	Time 1	*Time 2*	*Time 3*
Initiative Domain	3.18 (.53)	4.33 (.68)	4.33 (.68)
Develops and hones motivation for the OST task	**3.16 (.56)**	**3.56 (.45)**	**3.77 (.53)**
Articulates why the task and/or their specific role has personal value (e.g., feeling competent, social purpose, future career)	3.18	3.6	3.82
Finds positive opportunities (e.g., learning) in mistakes and failures	3.15	3.57	3.76
Can see positive opportunities in unexpected events or when plans go awry (e.g., "convert loss to win")	3.16	3.5	3.72
Perseveres through internal and external circumstances that challenge the OST work	**3.19 (.51)**	**3.55 (.46)**	**3.97 (.78)**
Stays focused on immediate (short-term) tasks despite difficulties or mistakes	3.15	3.56	3.75
Perseveres repeatedly through internal and external challenges for completion of the overall task	3.23	3.54	3.83

Data Source: SEL Challenge Staff Rating of Youth Behavior

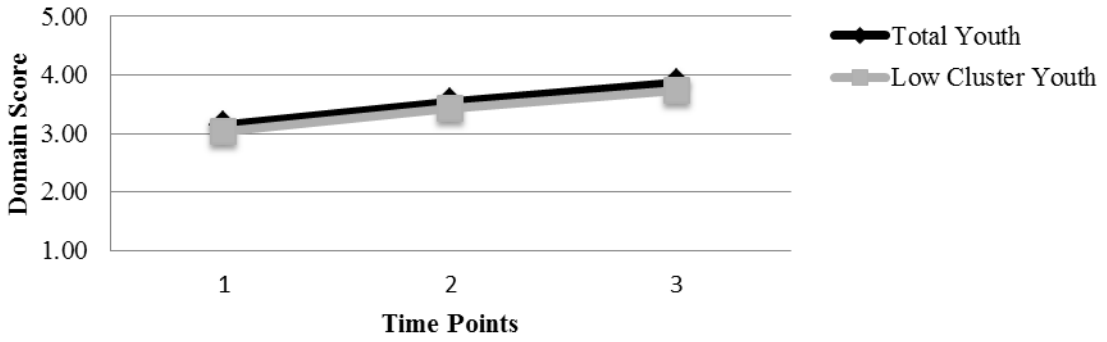

Figure D-11. **Initiative Behavior across Three Time Points with Low Cluster Comparison**

Table D-23. Problem Solving Behavior Scales

	Time 1	*Time 2*	*Time 3*
Problem Solving Domain	3.09 (.49)	4.38 (.49)	3.97 (.32)
Intentional learns task related methods and tools	**3.14 (.50)**	**5.50 (.85)**	**3.78 (.31)**
Learn task related methods and tools	3.27	3.75	3.84
Assess own learning needs to successfully use task related methods and tools	3.02	3.51	3.71
Uses Problem-solving skills to develop, evaluate, and adapt a course of action	**2.71 (.38)**	**3.18 (.46)**	**3.55 (.32)**
Create plans including guidelines and steps	2.66	3.07	3.41
Problem-solve by evaluating options and potential solutions (e.g., backwards-planning, weighing pros and cons, anticipating results, considering if-then sequences, developing plan-Bs, considering worst-case scenarios)	2.75	3.16	3.67
Monitor progress toward shorter and longer term goals	2.82	3.32	3.59
Time Management	**2.87 (.43)**	**3.38 (.40)**	**3.71 (.27)**
Prioritize tasks and allot time accordingly			
External Stakeholders	**3.26 (.51)**	**3.69 (.49)**	**3.97 (.44)**
Adjust communication to a target audience	3.1	3.51	.44
Acts as an ambassador for the program and/or OST task	3.14	3.9	.58
Seeks opportunities to use or share learning beyond their own project for the benefit of others	3.29	3.66	.48
Reflects on learning and significance of results	**3.1 (.48)**	**4.17 (.54)**	**4.66 (.43)**
Reflect on how own/others actions influence task outcomes (e.g., cause & effect)	3.1	3.37	3.71
Assess and acknowledge their own level of developing skill	3.11	3.6	3.91

Data Source: SEL Challenge Staff Rating of Youth Behavior

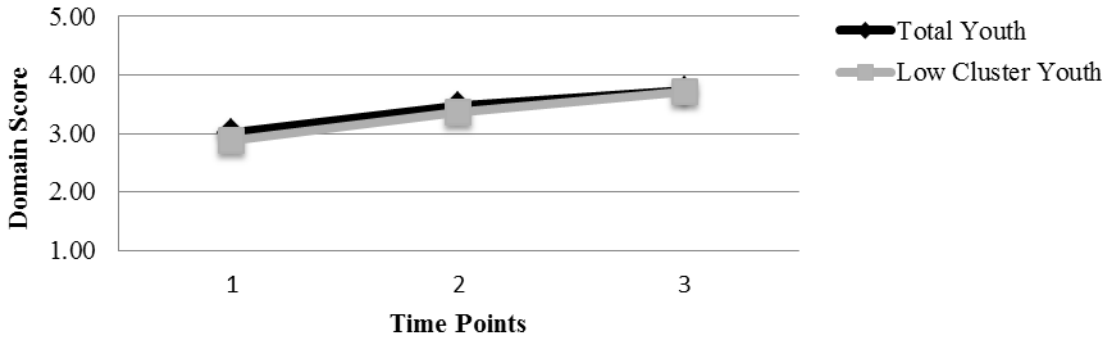

Figure D-12. **Problem Solving Behavior across Three Time Points with Low Cluster Comparison**

Appendix E - Technical Supplement for SEL Skill and Practice Measures

Appendix E provides detail on analyses and results for reliability and validity of the data produced using measures of SEL beliefs, SEL behaviors, and SEL practices. Basic scale and item descriptives (i.e., means and standard deviations by time point) for all measures are provided in Appendix D. Although Chapter 4 presents 12 composite scores for SEL skills, which combine information from 25 scales, all reliability and validity analyses in this appendix were carried out at the scale level; that is, 10 scales for SEL beliefs and 15 scales for SEL behaviors.

Summary of Results for Reliability and Validity of SEL Belief and Behavior Measures: With some important caveats, the data produced by the performance measures in the SEL Challenge 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. We characterize this evidence of construct validity as indicating that these measures are moderately precise. This also means that substantial improvements in the precision of these measures will be made in the next few iterations.

SEL Beliefs

Scale Reliability

Table E-1 displays Cronbach's alphas for the 10 belief scales. The item variables were group-mean centered in order to remove between-offering variance. In the absence of group-mean centering, the covariances used to calculate Cronbach's alpha would conflate the between-group variance with the within-group variance, and inferences about reliability may be wrong.

Table E-1. SEL Beliefs Scale Descriptives and Reliability

Scale	*Average Interitem Covariance*	*Cronbach's Alpha*
Optimism	.20	.58
Reappraisal Factor	.42	.82
Identification of Emotions	.33	.67
Adolescent Empathy	.28	.69
Adolescent Social Competency	.20	.75
Diligence and Reliability	.23	.71
Initiative Taking	.25	.62
Purpose	.31	.63
Goal Orientation	.28	.81
Problem Solving Strategies	.26	.71

Construct Validity

There are two options for dealing with multi-level data structure in traditional regression methods. The first is to adjust the standard errors to account for the differences in means and variances that likely exist among groups occurring at the same level. This involves the application of robust standard errors as is done for complex surveys that utilize clusters as the Primary Sampling Unit. The second is to explicitly model the multi-level structure using random effects. Both approaches were pursued. We conducted five confirmatory factor analysis (CFA) models for the SEL Beliefs data:

- Model 1 was the Emotion Management domain consisting of three constructs: Optimism, Emotion Reappraisal, and Identification of Emotions
- Model 2 combined the Empathy and Teamwork domains consisting of two constructs: Empathy and Social Competency
- Model 3 was the Responsibility domain with one construct: Diligence/Reliability
- Model 4 was the Initiative domain with two constructs: Initiative Taking and Purpose
- Model 5 was the Problem Solving domain with three constructs: Goal Orientation, Problem Solving Strategies, and Learning Effort

Each of the five models was estimated using baseline data. The model fit statistics are the primary focus to evaluate construct validity. Model fit statistics reflect the extent to which our theorized model – the items grouped together in the ways we think they are – actually approximates the information in the data. The five models fit statistics are:

- Chi-square – a chi-square statistic is computed to determine whether the correlation matrix produced by the theoretically-specified model is significantly different from the correlation matrix produced by the data. In this case, a low and non-significant value is desired, suggesting that the two matrices are not different from each other, Chi-Square is sensitive to sample size. If the sample size is too big, or the model has a large number of variables, the Chi-square is likely to be significant.
- Root Mean Square Error of Approximation (RMSEA) – RMSEA is commonly reported as an alternative fit index to the Chi-Square. A lower value (usually less than .06) suggests a good fit, whereas a value of 0 indicates a perfect fit.
- Comparative Fit Index (CFI) – CFI ranges from 0 to 1, and larger value indicates a better fit. In general, a CFI greater than .90 implies adequate model fit, whereas a value of 1 indicates a perfect fit.

- Tucker-Lewis Index (TLI) – TLI is similar to CFI in that both indices are affected by the average correlation in the data. A value of 1 indicates a perfect fit, and a TLI greater than .90 implies adequate model fit.
- Standardized Root Mean Square Residual (SRMR) – SRMR is a standardized difference between the correlation matrix produced by the data and the correlation matrix produced by the model. A value of 0 indicates a perfect fit, and a value less than .08 is generally considered a good fit.

In Table E-2, fit statistics are provided for each of the five models. Inferences vary from model to model depending on which statistic is used, but for every model, there are at least two of the five statistics that indicate acceptable fit.

Table E-2. Fit Statistics for Five Models and Eleven SEL Belief Constructs

	Chi-Square	RMSEA	CFI	TLI	WRMR
Model 1	182.96, df = 101, p < .001	.076	.927	.913	1.33
Model 2	123.87, df = 64, p < .001	.081	.922	.905	1.205
Model 3	46.39, df = 35, p = .942	.048	.903	.883	.49
Model 4	42.12, df = 13, p < .001	.126	.957	.93	.887
Model 5	127.29, df = 141, p = .789	0	1	1	.411

Growth Trajectories

The intent of the growth trajectory analyses was to determine if a positive growth trajectory on the 10 scale scores for SEL beliefs could be observed over time. The primary hypothesis tested was that scores on each scale will improve over time. A secondary hypothesis was: Youth who begin with low scores will show the greatest improvement over time.

The nature of the data – repeated measures nested within youth nested within offering – violates the assumptions of independence among youth skill reports and requires specialized methods that take the likely non-independence into account. The methodology employed below utilizes a three-level growth model that incorporate random effects for students (N = 153) and offerings (N = 8). In each model, youth is entered as a random effect (i.e., each student has his or her own starting point), and the effect of time is also allowed to vary by individual (i.e., each student has his or her own linear growth trajectory).

The youth and time random effects are summarized by variance components, which are statistics that detail the expected amount of student-to-student variability in starting points and growth. The "youth skill growth" coefficient is presented in the Slope column of Table E-3. The covariance between the youth and time random effects shows whether those who start out low on a scale tend to increase the

fastest (a negative covariance). This "equity" coefficient is presented as in "Intercept/Slope Correlation" column of Table E-3.

These analyses of the SEL belief data leads to the inference that scores do improve over baseline as time progresses. All of the slope coefficients are positive, and 7 of 10 are sufficiently large to achieve statistical significance in this small sample. An additional inference is that youth who have the lowest SEL skills to begin with show the greatest improvements over time. All of the intercept/slope coefficients are negative and in three cases – Optimism, Social Competency, and Purpose – sufficient to achieve statistical significance.

Finally, each of the growth models was re-estimated using measurement models for each of the scale scores that allowed for estimates of measurement error to be incorporated into the models. These models yielded a very similar pattern of results and provide a baseline for improvement of measures in future iterations.

Table E-3. SEL Beliefs Growth Trajectories

Domain	*Scale*	*Slope*	*Intercept/Slope Correlation*
Emotion Management	Optimism	.14**	-.09*
		(.042)	(.038)
	Emotion Reappraisal	.09*	-.09
		(.041)	(.060)
	Identification of Emotions	.08*	-.03
		(.041)	(.052)
Empathy	Empathy	.04	-.07
		(.041)	(.039)
Teamwork	Social Competency	.05	-.08*
		(.031)	(.038)
Responsibility	Diligence/Reliability	.10**	-.07
		(.032)	(.035)
Initiative	Initiative Taking	.05	-.04
		(.024)	(.022)
	Purpose	.13***	-.09*
		(.031)	(.037)
Problem Solving	Goal Orientation	.08*	-.10
		(.037)	(0.53)
	Problem Solving Strategies	.07*	-.08
		(.032)	(0.04)
	Learning Effort	.10**	.06
		(.030)	(.038)

*Note. Standard errors are in parentheses. * p < .05, ** p < .01, *** p < .001*

SEL Behaviors

Substantive and content validity

The SEL behavior measures were developed in collaboration with the expert practitioners who ultimately completed the assessments, with their own youth in each offering, at each of the three time points. Per Weikart Center's methodology for establishing reliability and validity of observation-based measures (Smith and Hallman et al., 2012), we initiated an item development process that focused on item clarity and evaluation of content and substantive validity. By content and substantive validity, we mean that the practice was rated by the expert practitioners as being both important to their work (content validity) and frequently occurring (substantive validity) in their curriculum design. Table E-4 presents results for expert ratings for item clarity, importance of the skill in the offering curriculum, and prevalence of skill demonstration in a typical offering session. Importantly, staff said that skills were likely to become more prevalent toward the end of the offering curriculum.

Table E-4. Substantive and Content Validity

Feature	*Average Rating*
Clarity of item working	1.92 (out of 2; clear or not clear)
Importance of skill in offering	2.55 (out of 3; not important to very important)
Skill prevalence at time one	1.3 (out of 3; 1= almost never; 3=in at least 80% of offering sessions)
Ending Skill Detection	2.4 (out of 3; 1=almost never; 3= in at least 80% of offering sessions)

Reliability

Table E-5 displays Cronbach's alphas for the 15 behavior scales. The item variables were group-mean centered in order to remove between-offering variance and improve the precision of reliability estimates.

Table E-5. SEL Behaviors Scale Reliability

Scale	*Average Interitem Covariance*	*Cronbach Alpha*
Emotion Management Behaviors	.35	.82
Reason about Causes	.31	.71
Manage Emotions for Functional Purpose	.19	.77
Values Own/Others Perspectives	.26	.84
Practices Respectful/Effective Communication	.19	.78
Coordinates and Supports Action Toward Team Goals	.30	.82
Fulfills Roles and Commitments	.27	.79
Defines, Adjusts, and Negotiates Roles	.26	.77
Develops Motivation for the OST Task	.37	.82
Perseveres through Internal/External Circumstances	.32	.79
Intentionally Learns OST-Task Related Methods and Tools	.21	.76
Uses Problem-Solving to Develop, Evaluate, and Adapt Course of Action	.30	.83
Connects with External Stakeholders	.37	.80
Reflects on Learning and Significance of Results	.34	.80

Inter-Rater Reliability

Inter-rater reliability in the SEL Challenge describes how consistently a pair of the same staff rated youth behavior at the three time points on the SEL behavior items. Table 6 presents results for rater consistency using three measures of inter-rater agreement: the intra-class correlation coefficient (ICC), Cohen's weighted kappa, and correlations (Kendall's tau for individual items, Pearson's *r* for scales).

The ICC is based on a single random effects model. It was not possible to include a second random effect for rater given that the raters differed across youth. The interpretation of the ICC is the amount of variability in between-youth scores relative to within-youth scores. The latter variance will be low, and the former high, when raters score the same youth in a similar fashion. One heuristic for interpreting ICCs is to view values above .75 as excellent, values from .60 to .74 as good, values from .40 to .59 as fair, and values below .40 as poor (D. V. Cicchetti, 1994).

The tables also report Cohen's weighted kappa for the individual items. No kappa is reported for the scales as they were based on averages and thus included non-integer values. The weighted form of kappa was calculated due to the fact that the scores were ordinal rather than nominal. Weighted kappa counts a disagreement between the pair of scores (1, 2) to be less severe than disagreement between the pair of scores (1, 5). Following the heuristics from Landis and Koch (1977), negative kappas mean no agreement, scores from zero to .20 indicate slight agreement, scores from .21 to .40 indicate fair agreement, scores from .41 to .60 indicate moderate agreement, scores from .61 to .80 indicate substantial agreement, and scores .81 and above indicate near perfect agreement.

The final measure of association between the two scores is the correlation. This is the nonparametric Kendall's tau-b for the ordinal items and Pearson's *r* for the full scales. There are few heuristics for using correlations to measure inter-rater reliability, but their metric is familiar to most researchers and easily understood

Table E-6. Inter-rater Reliability for SEL Behavior Items and Scales

	Time 1			***Time 2***			***Time 3***		
	ICC	*Kappa*	*Correlation*	*ICC*	*Kappa*	*Correlation*	*ICC*	*Kappa*	*Correlation*
SCALE: Identify positive and negative emotions.									
Identifies and names emotions.	.46	.25	.37	.33	.22	.35	.52	.36	.47
Has a wide descriptive vocabulary for emotions.	.53	.36	.46	.42	.32	.43	.62	.41	.57
Describes own emotional needs.	.32	.23	.29	.12	.18	.28	.27	.21	.34
SCALE:	**.52**		**.52**	**.37**		**.49**	**.51**		**.62**

Table E-6. Inter-rater Reliability for SEL Behavior Items and Scales (continued)

	Time 1			*Time 2*			*Time 3*		
	ICC	*Kappa*	*Correlation*	*ICC*	*Kappa*	*Correlation*	*ICC*	*Kappa*	*Correlation*
SCALE: Reason about causes of emotion.									
Identifies causes/triggers of emotion.	.45	.28	.37	.36	.33	.42	.43	.24	.40
Can guide self and others to address stress or extreme emotions.	.51	.33	.44	.17	.11	.29	.37	.27	.35
SCALE	**.52**		**.52**	**.32**		**.47**	**.38**		**.43**
SCALE: Manage emotions for functional purpose									
Checks for misunderstanding when negative emotions occur.	.4	.23	.36	.30	.17	.22	.27	.09	.24
Manages positive emotion with confidence that doesn't belittle or exclude others.	.37	.29	.35	.44	.26	.39	.59	.38	.58
Maintains composure to constructively work through conflict or disagreement to maintain progress on task.	.31	.19	.30	.15	.06	.14	.36	.18	.35
Actively reaches out to others when they have emotional difficulties.	.28	.20	.24	.21	.19	.26	.57	.37	.50
Accepts feedback nondefensively, sees kernel of truth, restates with understanding to giver of feedback.	.28	.21	.28	.40	.27	.36	.46	.27	.43
Allows others to express and take responsibility for their own emotions.	.20	.12	.21	.44	.28	.39	.44	.24	.32
SCALE:	**.38**		**.42**	**.46**		**.47**	**.5**		**.57**
SCALE: Values own/others perspectives and stories with sensitivity to context.									
Use inclusive language and socially preferred terms.	.55	.37	.47	.67	.51	.63	.36	.36	.47
Correct others' bias or use of exclusive language.	.42	.27	.39	.60	.50	.59	.46	.35	.48
Help ensure others are heard.	.19	.16	.20	.38	.21	.33	.19	.18	.29
Share personal information at an appropriate level and at an appropriate time.	.24	.18	.23	.41	.31	.43	.19	.20	.24
Reflect appropriate tone, gesture, feeling, and pacing during sensitive discussion.	.44	.30	.35	.40	.29	.37	.42	.25	.34
Can explain biases.	.47	.28	.42	.43	.34	.39	.52	.35	.51
Composes and shares their personal story.	.19	.1	.17	.39	.29	.41	.32	.27	.38
SCALE:	**.53**		**.55**	**.6**		**.64**	**.42**		**.60**

Table E-6. Inter-rater Reliability for SEL Behavior Items and Scales (continued)

	Time 1			*Time 2*			*Time 3*		
	ICC	*Kappa*	*Correlation*	*ICC*	*Kappa*	*Correlation*	*ICC*	*Kappa*	*Correlation*
SCALE: Practices respectful and effective communication within a team.									
Clearly articulate point of view and/or ask follow up questions for clarification.	.50	.33	.46	.31	.17	.32	.43	.32	.46
Communicate without dominating, interrupting, or showing disrespect for others' ideas.	.32	.24	.34	.44	.30	.40	.41	.37	.45
Demonstrate listening/interest through body language.	.40	.26	.38	.40	.27	.37	.50	.42	.55
Upon disagreement, acknowledge and express interest in other's point of view.	.24	.12	.24	.27	.13	.22	.35	.21	.27
Refrain from gossip when having a problem and talk with the person directly.	.33	.26	.34	.39	.27	.39	.34	.24	.33
Show respect to group leaders.	.33	.23	.28	.58	.44	.50	.57	.48	.67
SCALE:	**.48**		**.54**	**.48**		**.50**	**.48**		**.63**
SCALE: Coordinates and supports action toward team goals.									
Contribute ideas, skills, and commitment to group tasks.	.31	.22	.30	.50	.37	.47	.56	.47	.55
Offer to help those having trouble completing tasks.	.35	.22	.33	.42	.33	.42	.14	.15	.22
Monitor team progress on the OST task.	.34	.29	.37	.33	.21	.30	.23	.16	.28
Encourage youth and staff to be accountable for their roles and commitments.	.3	.18	.25	.27	.15	.23	.31	.22	.28
SCALE:	**.46**		**.50**	**.44**		**.45**	**.33**		**.45**
SCALE: Fulfills roles and commitments.									
Fulfill OST task role(s) with minimal supervision.	.20	.19	.25	.56	.44	.55	.41	.23	.38
Work toward mastery or excellence in fulfilling the role.	.45	.28	.40	.46	.32	.48	.33	.20	.38
Acknowledge mistakes and is willing to address mistakes through action.	.32	.21	.27	.30	.22	.32	.27	.15	.27
Volunteer for additional tasks beyond the assigned role to further group goals.	.21	.15	.19	.13	.30	.14	.32	.14	.29
SCALE:	**.35**		**.38**	**.47**		**.47**	**.4**		**.50**
SCALE: Successfully defines, adjust, and negotiates roles and commitments when required.									
Seek timely help from other youth or staff when roles become too challenging.	.27	.22	.21	.14	.04	.13	.14	.12	.16
Negotiate with staff to adjust expectations for redefined commitments and deadlines.	.10	.06	.05	.50	.31	.41	.58	.37	0.54
SCALE:	**.28**		**.26**	**.4**		**.42**	**.41**		**.55**

Table E-6. Inter-rater Reliability for SEL Behavior Items and Scales (continued)

	Time 1			*Time 2*			*Time 3*		
	ICC	*Kappa*	*Correlation*	*ICC*	*Kappa*	*Correlation*	*ICC*	*Kappa*	*Correlation*
SCALE: Develops and hones motivation for the OST task.									
Articulates why the OST task and/or their specific role has personal value.	.44	.27	.36	.37	.27	.33	.26	.24	.33
Finds positive opportunities (e.g., learning) in mistakes and failures.	.31	.25	.31	.41	.30	.38	.43	.29	.43
Can see positive opportunities in unexpected events or when plans go awry.	.35	.28	.34	.29	.27	.29	.34	.24	.37
SCALE:	**.42**		**.43**	**.41**		**.49**	**.35**		**.48**
SCALE: Perseveres through internal and external circumstances that challenge OST work.									
Stays focused on immediate (short-term) tasks despite difficulties or mistakes.	.32	.19	.31	.47	.32	.42	.45	.27	.43
Perseveres repeatedly through internal and external challenges for completion of the overall OST task.	.37	.22	.36	.27	.25	.29	.44	.33	.47
SCALE:	**.41**		**.43**	**.41**		**.40**	**.47**		**.52**
SCALE: Intentionally learns OST task-related methods and tools.									
Learn OST task-related methods and tools.	.39	.23	.40	.66	.52	.62	.40	.24	.41
Assess own learning needs to successfully use OST task-related methods and tools.	.37	.22	.40	.63	.49	.56	.33	.25	.44
SCALE:	**.41**		**.52**	**.67**		**.68**	**.41**		**.53**
SCALE: Uses problem solving skills to develop, evaluate, and adapt a course of action.									
Create plans including guidelines and steps.	.25	.13	.22	.51	.36	.48	.49	.28	.54
Problem solve by evaluating options and potential solutions.	.45	.28	.39	.59	.48	.52	.31	.24	.37
Monitor progress toward shorter- and longer-term goals.	.45	.32	.49	.50	.31	.43	.29	.18	.31
SCALE:	**.47**		**.49**	**.66**		**.67**	**.38**		**.47**
SCALE: Successfully manages time.									
Prioritize tasks and allot time accordingly.	.29	.19	.28	.3	.20	.29	.44	.26	.45
SCALE:	**.29**		**.34**	**.3**		**.29**	**.44**		**.53**
SCALE: Connects with external stakeholders									
Adjust communication to a target audience.	.21	.13	.18	.35	.21	.31	.23	.10	.25
Acts as an ambassador for the program and/or OST task.	.35	.22	.31	.51	.37	.49	.39	.29	.44
Seeks opportunities to use or share learning beyond their own project for the benefit of others.	.48	.33	.43	.46	.32	.41	.39	.30	.43
SCALE:	**.54**		**.57**	**.53**		**.56**	**.39**		**.53**

Table E-6. Inter-rater Reliability for SEL Behavior Items and Scales (continued)

	Time 1			*Time 2*			*Time 3*		
	ICC	*Kappa*	*Correlation*	*ICC*	*Kappa*	*Correlation*	*ICC*	*Kappa*	*Correlation*
SCALE: Reflects on learning and significance of results.									
Reflect on how own/others' actions influence OST task outcomes (e.g., cause & effect).	.23	.12	.22	.25	.13	.19	.45	.29	.45
Assess and acknowledge their own level of developing skill.	.32	.23	.34	.42	.30	.38	.35	.26	.41
SCALE:	**.28**		**.36**	**.39**		**.4**	**.44**		**.57**

Construct Validity

We conducted five CFA models using the baseline SEL behavior data:

- Model 1 was the Emotion Management domain consisting of three constructs: Identify Emotions, Reason about Causes/Uses of Emotions, Manage Emotions for Functional Purposes
- Model 2 combined the Empathy and Teamwork domains consisting of three constructs: Values Perspectives, Team Communication, Supports Action Toward Team Goals
- Model 3 was the Responsibility domain with two constructs: Fulfills Commitments, Adjusts Roles
- Model 4 was the Initiative domain with two constructs: Develops Motivation, Perseverance
- Model 5 was the Problem Solving domain with five constructs: Learns Task-Related Tools, Problem Solving, Manages Time, Connects with Stakeholders, Reflections on Learning

Each of the five models was estimated using baseline data, and the model fit statistics are the primary focus given that the interest is in construct validity. Model fit statistics reflect the extent to which our theorized model – the items grouped together in the ways we think they are – actually approximates the data. The five model fit statistics are described in the SEL beliefs section above. In Table E-7, fit statistics are provided for each of the five models. Inferences vary from model to model depending on which statistic is used, but for every model, there is at least four of the five statistics that indicate acceptable fit.

Table E-7. Fit Statistics for Five Models and Fifteen SEL Behavior Constructs

	Chi-Square	RMSEA	CFI	TLI	WRMR
Model 1	79.95, *df* = 93, *p* = .830	0	1	1.087	.311
Model 2	256.76, *df* = 249, *p* = .354	.015	.982	.981	.648
Model 3	28.50, *df* = 33, *p* = .690	0	1	1.015	.327
Model 4	3.81, *df* = 4, *p* = .431	0	1	1	.276
Model 5	57.00, *df* = 68, *p* = .826	0	1	1.036	.306

Growth Trajectories

As with the SEL beliefs measures, the primary hypothesis tested with the SEL behavior data was that scores on the 15 scales will improve over time. Again, secondary hypotheses were (a) youth who begin with low scores will be the ones who show that greatest improvement with time, and (b) youth in some of the eight offerings will demonstrate different pattern of performance than others.

The methodology employed below utilizes a three-level growth model that incorporate random effects for students (N = 153) and offering (N = 8). In each model, youth is entered as a random effect (i.e., each student has his or her own starting point), and the effect of time is also allowed to vary by individual (i.e., each student has his or her own linear growth trajectory).

The youth and time random effects are summarized by variance components, which are statistics that detail the expected amount of student-to-student variability in starting points and growth. The "youth skill growth" coefficient is presented in the Slope column of Table E-8. The covariance between the youth and time random effects shows whether those who start out low on a scale tend to increase the fastest (a negative covariance). This "equity" coefficient is presented as in "Intercept/Slope Correlation" column of Table E-8.

These analyses of the SEL behavior data lead to the inference that scores improve over baseline as time progresses. All of the slope coefficients are positive and sufficiently large to achieve statistical significance in this small sample. An additional inference is that youth who have the lowest SEL behavior scores to begin with show the greatest improvements over time. All but one of the intercept/slope coefficients are negative and in two cases – Values Perspectives and Team Communication – sufficient to achieve statistical significance.

Finally, each of the growth models was re-estimated using measurement models for each of the scale scores that allowed for estimates of measurement error to be incorporated into the models. These models yielded a very similar pattern of results and provide a baseline for improvement of measures in future iterations.

Table E-8. SEL Behaviors Growth Trajectories

Domain	*Scale*	*Slope*	*Intercept/ Slope Correlation*
Emotion Management	Identify Emotions	.36*** (.037)	--
	Reason about Causes/Uses of Emotions	.38*** (.047)	--
	Manage Emotions for Functional Purpose	.29*** (.031)	--
Empathy	Values Perspectives	.33*** (.035)	-.15*** (.033)
Teamwork	Team Communication	.21*** (.029)	-.08** (.027)
	Supports Action Towards Team Goals	.30*** (.029)	-.08 (.043)
Responsibility	Fulfills Commitments	.29*** (.035)	-.02 (.038)
	Adjust Roles	.41*** (.041)	-.04 (.066)
Initiative	Develops Motivation	.34*** (.042)	-.12 (.064)
	Perseverance	.36*** (.040)	-.11 (.062)
Problem Solving	Learns Task-Related Tools	.30*** (.035)	-.04 (.038)
	Problem-Solving	.38*** (.036)	.04 (.045)
	Manages Time	.38*** (.044)	-.04 (.064)
	Connects with Stakeholders	.26*** (.029)	-
	Reflections on Learning	.32*** (.038)	-.06 (.039)

*Note. Standard errors are in parentheses. * p < .05, ** p < .01, *** p < .001*

Quality of SEL Practice Measures (SEL PQA)

Tables E-9 and E-10 present descriptive information for observation-based measures of SEL practices. Scores in Tables E-9 and E-10 were created by taking the mean for three ratings for each offering and then taking the mean across the eight offerings. This average-of-averages score is presented for each item, scale and domain.

Table E-9 presents domain, scale, and item-level information for the Youth Program Quality Assessment (Youth PQA; Smith and Hohmann, 2005). Table E-10 provides information for a subset of Youth PQA items aligned to the content of the six SEL domains reflected in the SEL standards (See Table 4). This crosswalk of Youth PQA items to the six SEL domains demonstrates that coverage of the SEL standards using the pre-existing content of the Youth PQA items is incomplete and additional items will be developed in future iterations of the measure. Finally, Table E-11 presents bi-variate correlations for the six SEL PQA domain scores.

Table E-9. Descriptive Statistics for SEL PQA Domains, Scales, and Items

	Range	*Mean*	*SD*
Safe Environment	**.67**	**4.70**	**.19**
Psychological and emotional safety is promoted.	.00	5.00	.00
Positive emotional climate	.00	5.00	.00
Lack of bias	.00	5.00	.00
Healthy Environment: The physical environment is safe and free of health hazards.	1.50	4.83	.39
Free of health and safety hazards	2.00	4.74	.69
Clean and sanitary	2.00	4.83	.58
Ventilation and lighting	2.00	4.91	.42
Temperature	2.00	4.83	.58
Emergency Procedures: Appropriate emergency procedures and supplies are present.	3.00	4.26	.79
Posted emergency procedures	4.00	4.30	1.43
Fire extinguisher	4.00	4.30	1.29
First-aid kit	4.00	3.43	1.47
Other safety equipment	.00	5.00	.00
Supervised entrances	2.00	4.83	.58
Supervised access to outdoor space	2.00	4.78	.67
Accommodating Environment: Program space and furniture accommodate the activities	1.00	4.87	.27
Sufficient space	2.00	4.57	.84
Suitable space	2.00	4.91	.42
Furniture	.00	5.00	.00
Physical environment can be modified	.00	5.00	.00

Health and Nutrition: Healthy food and physical activity are provided.	2.00	4.57	.71
Plentiful food and drink	2.00	4.74	.69
Nourishing food and drink	.00	5.00	.00
Healthy food and drinks	4.00	4.29	1.21
Supportive Environment	**1.01**	**4.64**	**0.26**
Warm Welcome: Staff provides a welcoming atmosphere.	1.33	4.86	.35
Staff greet youth	4.00	4.57	1.04
Staff warm and respectful	.00	5.00	.00
Positive staff body language	.00	5.00	.00
Session Flow: Session flow is planned, presented and paced for youth.	1.60	4.76	.38
Start and end on time	4.00	4.22	1.17
Materials and supplies ready	2.00	4.90	.44
Sufficient materials	2.00	4.90	.44
Staff explain activities clearly	2.00	4.83	.58
Appropriate time for activities	2.00	4.91	.42
Active Engagement: Activities support active engagement	1.00	4.67	.39
Youth engage with materials or ideas.	2.00	4.83	.58
Youth talk about activities	2.00	4.74	.69
Balance concrete and abstract	2.00	4.74	.69
Tangible products and performances	4.00	4.39	1.41
Learning focus linked to activity	4.00	3.87	1.79
Skill Building: Staff supports children in building skills	3.50	4.16	1.06
Staff encourages youth to try new skills	2.00	4.30	.97
Staff models skills	4.00	4.13	1.46
Staff breaks down tasks	4.00	4.18	1.33
Support for struggling youth	2.00	4.85	.55
Encouragement Staff supports youth with encouragement.	1.33	4.77	.38
Staff uses non-evaluative language	2.00	4.48	.90
Staff asks open-ended questions	2.00	4.83	.58
Staff actively involved	.00	5.00	.00
Reframing Conflict: Staff uses youth-centered approaches to reframe conflict.	.00	5.00	.00
Approach calmly	.00	5.00	.00
Seek input from youth	.00	5.00	.00
Relationship between actions and consequences	.00	5.00	.00
Staff follow-up	.00	5.00	.00
Interaction	**2.63**	**4.02**	**.64**
Belonging: Youth have opportunities to develop a sense of belonging.	1.50	4.57	.49
Get to know each other	2.00	4.74	.69
Publically acknowledge achievements	4.00	3.91	1.60
Collaboration: Youth have opportunities to collaborate and work cooperatively with others.	2.67	4.19	.98

Opportunities to work collaboratively	4.00	4.65	.98
Interdependent roles	4.00	3.87	1.46
Shared goals	4.00	4.04	1.69
Leadership: Youth have opportunities to act as group facilitators and mentors.	3.33	2.91	.88
Group process skills	2.00	4.65	.78
Opportunities to mentor	4.00	1.61	1.12
Opportunities to lead a group	4.00	2.48	1.50
Adult Partners: Youth have opportunities to partner with adults.	3.00	4.39	1.12
Staff share control with youth	4.00	4.57	1.04
Staff provide an explanation	4.00	3.94	1.75
Engagement	**2.50**	**3.88**	**.73**
Planning: Youth have opportunities to make plans.	4.00	3.52	1.34
Opportunities to make plans	4.00	3.61	1.53
Planning strategies	4.00	3.43	1.34
Choice: Youth have opportunities to make choices based on their interests.	3.00	4.30	.97
Content choices	4.00	4.27	1.32
Process choices	4.00	4.30	1.29
Reflection: Youth have opportunities to reflect.	2.50	3.83	.70
Youth reflect on what they are doing	2.00	4.65	.78
Youth reflect in multiple ways	4.00	3.52	1.08
Youth provide feedback	4.00	2.83	1.90
Youth present to group	4.00	4.30	1.43
Total Score	**1.26**	**4.31**	**.32**
Instructional Total Score	**1.77**	**4.18**	**.43**

Data Source: SEL Challenge PQA Data

Table E-10. Scale Means and Standard Deviation for Crosswalk of Youth PQA Items to Six SEL Domains

SEL Domain	*PQA Items*	*Mean (SD)*
Emotion Management (10 items)	(Rf.1) Intentional reflection (SB.3) Staff models skills (SB.5) Support for struggling youth (Rf.2) Multiple reflection strategies (AE.2) Youth talk about activities (Ec.1) Staff uses non-evaluative language (Ec.3) Staff actively involved (AP.2) Expectations explained (Be.1) Opportunities for youth to get to know each other (ES.1) Positive emotional climate	4.50 (.05)
Empathy (10 items)	(Rf.1) Intentional reflection (SB.3) Staff models skills (Rf.2) Multiple reflection strategies (Ec.2) Staff asks open-ended questions (Be.1) Opportunities for youth to get to know each other (ES.2) Lack of bias (WW.1) Youth greeted (WW.2) Staff warm and respectful (WW.3) Positive staff body language (Be.2) Inclusive relationships	4.64 (.16)
Teamwork (9 items)	(Ld.2) Mentoring Opportunities (AP.1) Staff shares control with youth (Be.4) Public acknowledgement of achievements (Co.3) Shared goals (Co.2) Interdependent roles (Ld.3) All youth lead group (Co.1) Opportunities to work cooperatively (Ld.1) Practice group process skills (Rf.4) Structured opportunities to present to a group	3.79 (.26)
Responsibility (11 items)	(Rf.1) Intentional reflection (Ld.2) Mentoring Opportunities (AP.1) Staff shares control with youth (SB.3) Staff models skills (SB.5) Support for struggling youth (SB.1) Learning focus linked to activity (SB.2) Staff encourages to try skills (Co.2) Interdependent roles (Ld.3) All youth lead group (AP.2) Expectations explained (Be.3) Youth identify with program	3.86 (.27)

Table E10. SEL PQA Scale Means and Standard Deviation *(continued)*

Initiative (11 items)	(Rf.1) Intentional reflection (Pn.1) Opportunities to make plans (Ch.1) Content alternatives (Ch.2) Process alternatives (Be.4) Public acknowledgement of achievements (Co.3) Shared goals (Ec.2) Staff asks open-ended questions (Ec.1) Staff uses non-evaluative language (Ec.3) Staff actively involved (AE.3) Balance concrete and abstract (AE.4) Tangible products or performances	4.39 (.19)
Problem Solving (17 items)	(Rf.1) Intentional reflection (Pn.1) Opportunities to make plans (Ch.1) Content alternatives (Ch.2) Process alternatives (Ld.2) Mentoring Opportunities (AP.1) Staff shares control with youth (SB.3) Staff models skills (SB.5) Support for struggling youth (SB.1) Learning focus linked to activity (SB.2) Staff encourages to try skills (AE.2) Youth talk about activities (SF.4) Explains activities clearly (AE.1) Youth engage with materials or ideas (SB.4) Staff breaks down tasks (Pn.2) Multiple planning strategies used (Rf.3) Structured opportunities to provide feedback	4.01 (.28)

Table E-11. SEL PQA Correlation Matrix

	Emotion Management	*Empathy*	*Teamwork*	*Responsibility*	*Initiative*	*Problem Solving*	*Instructional Total Score*
Emotion Management							
Empathy	.60**						
Teamwork	.61**	.34					
Responsibility	.82**	.57**	.76**				
Initiative	.46*	.23	.81**	.55**			
Problem Solving	.60**	.44*	.61**	.81**	.61**		
Instructional Total Score	.55**	.33	.88**	.73**	.92**	.77**	

Source: SEL Challenge PQA Data
*** Correlation is significant at the .01 level*
** Correlation is significant at the .05 level*

Endnotes

[i] "Promising" typically indicates that the practice is both theoretically defined and has some supporting evidence of effectiveness - but not from a sufficiently rigorous design to move the designation to "evidence-based practice." We could easily make an argument that many of these practices could be considered evidence-based given the depth and rigor of the Larson et al. literature and the Weikart Center's prior work.

[ii] Contexts associated with poverty, traumatic experience, and chronic stress can cause dysregulation of emotion, motivation, attention, and behavior, and ultimately, limit youth experience of agency and developmental potential. As adolescents accumulate experience with successful self-regulation, they develop increasing confidence that, with effort, they can engage positive contexts and overcome challenging ones. SEL skills developed through experiences of successful self-regulation likely moderate the effects of earlier negative experience on outcomes in early adulthood. In short, the recovery and healthy development of adolescents suffering the effects of difficult SEL histories is likely to be fostered through exposure to contexts like the exemplary SEL offerings identified in the Challenge. These ideas are discussed in several disciplinary languages (for example: Blair & Raver, 2012; Bryck & Fisher, 2012; Côté, 2000; Côté & Levine, 2002; Curtis & Cicchetti, 2007; Evans & Fuller-Rowell, 2013; Murray, Rosanbalm, Christopoulos, & Hamoudi, 2015).

We identify two forms of agency (Peck, 2007) that go along with the practice of SEL skills. First, *Type 1 Agency* describes the mental experience of having your skills successfully put to work, more or less automatically, and often outside of the immediate focus of awareness. Type 1 Agency corresponds to the concept of *primary appraisal* (See Appendix C; cf., Lazarus, 1991; Scherer, 2001; C. A. Smith & Kirby, 2000). Many OST environments, particularly those designed for elementary-aged children, are specifically designed to help kids feel safe, interested in the content, and successful at demonstrating skills. Successful self-regulation increases the likelihood that the context can successfully activate mental processes that create engagement (rather than avoidance).

Second, *Type 2 Agency* describes forms of executive control where youth focus awareness on challenges and making well-informed decisions about what to do next. Type 2 Agency corresponds to the concept of *secondary appraisal* (See Appendix C; cf., Lazarus, 1991; Scherer, 2001; C. A. Smith & Kirby, 2000). In addition to effectively using skills more or less automatically, now the context should support the experience of more extended learning sequences such as trial, error, and adjustments or selection, optimization, compensation (cf. Baltes, 1997). OST environments for adolescents often include project-based curricula and youth control within their intervention designs, requiring higher levels of self-regulation. When doing SEL work, researchers and practitioners should attend equally to both forms of agency.

[iii] The Durlak et al meta-analyses were directly focused on SEL offerings in OST and mix of universal and targeted interventions in schools. The SAFE practices – sequenced, active, focused, and explicit – that were found to differentiate effective from ineffective programs are aligned with curriculum features identified in the SEL Challenge (Durlak & Weissberg, 2010; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Feldman Farb & Matjasko, 2012; Gottfredson & Wilson, 2003; Hattie & Timperley, 2007; Lipsey, Howell, Kelly, Chapman, & Carver, 2010; Porath-Waller, Beasley, & Beirness, 2010; Wheeler, Keller, & DuBois, 2010).

In addition, meta-analyses from several other disciplines and policy domains also lead to the conclusion that program experiences like those available in the SEL Challenge lead to positive effects on a wide variety of youth outcomes. (Apsler, 2009; Catalano, Berglund, Ryan, Lonczak, & Hawkins, 2004; Chappell, Nunnery, Pribesh, & Hager, 2011; Durlak & Weissberg, 2010; Durlak & Wells, 1997; Lauer et al., 2006; Li & Julian, 2012; Sambrano, Springer, Sale, Kasim, & Hermann, 2005; Springer et al., 2004)

[iv] It is evident that SEL skills are not only necessary for youth to successfully learn, as is suggested by the many meta-analyses linking self-regulation to learning outcomes, but that many of the SEL practices described in the standards are themselves best practices for teaching academic and other content. Several examples: (1) The Emotion Management and Initiative domains are focused explicitly on self-regulation of emotion and motivation. In the standards, SEL skills are organized as a hierarchy, with practice indicators across domains

addressing first declarative (naming things) and then procedural (how to do things) beliefs (e.g., EM3, EM6, E1, T7, R5, PS1). Practice indicators in all domains also describe the practice of scaffolding (Vygotsky, 1962), that is providing just enough support to enable a learner to achieve more than what they could without that support. (e.g.,T9b, R9, I5, I6, PS12, PS13). The project curricula from the offerings assured complex of goal structures that provided lots of opportunities for higher order problem solving (Oosterhof, Rohani, Sanfilippo, Stillwell, & Hawkins, 2008) for youth with roles on interdependent teams (Slavin, 1996).

[v] Despite what may be an emerging consensus about the nature of SEL and context factors that effectively promote SEL development, the diverse array of approaches to studying and promoting SEL reflect lack of consensus about the core nature of SEL and the context factors suited best to promote SEL development. This lack of shared understanding extends to the most important implications and social goals to be achieved by investing SEL infrastructure (Hoffman, 2009). As an example relevant to the fields of OST and k-12 schooling, the University of Chicago Consortium on Chicago School Research (UCCSR; Farrington et al., 2012) report on "The Role of Noncognitive Factors in Shaping School Performance," described SEL in terms of "five general categories of noncognitive factors related to academic performance" (p. 6): academic behaviors, academic perseverance, academic mindsets, learning strategies, and social skills. In contrast, the UCCSR report (Nagaoka et al., 2015) report on "Foundations for Young Adult Success" described SEL in terms of three "factors for success" (i.e., agency, competencies, and integrated identity) and 4 "foundational components" (i.e., self-regulation, knowledge & skills, mindsets, and values). In further contrast, the Collaborative for Academic Social and Emotional Learning (2015) report on "Effective Social and Emotional Learning Programs," described SEL in terms of "five interrelated sets of cognitive, affective, and behavioral competencies" (p. 2): self-awareness, self-management, social awareness, relationship skills, and responsible decision-making. Close analysis of the content of these general categories, as well as those used in the present report, reveals a wide range of possible psychological and social processes and mechanisms (e.g., the ability to selectively focus attention; the ability identify emotions in oneself and others; the ability to inhibit dominant responses) that tend to overlap both within and between the categories used by any given approach. Additional research and theoretical integration will be required to achieve consensus about both the nature of socio-emotional learning and the context factors that effectively promote SEL development.

Because the issue of context-person interaction is so complex, and the various terms so unclear, one of our primary objectives in the Challenge was to separate mental processes from behavior, and these, in turn, from context. This was difficult, and we were not successful in several instances.

[vi] The fields of education and human services are currently locked in an ideology that over-values impact evaluation and its methodological emphases (reliability of individual measures, external validity of research design) and under values "improvement science" (Hiebert, Gallimore, & Stigler, 2002; Yeager, Bryk, Muhich, Hasuman, & Morales, 2013) and its methodological emphases (feasibility, measurement of settings and change, criterion validity). Another consequence of this ideology is that researchers, policy makers, funders, and regulators are tuned to hear results from outcome studies, resulting in an evidence base for SEL that is overflowing with rigorous tests of impact on youth outcomes for incompletely developed interventions that consequently fail to go to scale (Dodge, 2011; O'Connell, Boat, & Warner, 2009). As Larson (2000) has aptly noted, "outcomes research is often the necessary evil that is done before anyone knows what to look for, and that has been the case in this domain as well. Evaluative data have been needed to justify funding for youth activities, even though we are not yet sure what the independent and dependent variables for this evaluative research should be" (p. 180).

[vii] This concept of a cumulative base of expertise has been called mid-level or middle-range theory (Merton, 1949) or, more recently, practical theory. Yeager et al. (2013) described practical theories as follows: "Practical theory is not a *disciplinary* theory, in that it does not seek to document novel features of human psychology or social processes that shape the way humans think or behave. Instead, a practical theory draws on both the wisdom of practice as well as insight from academic theories to guide practice improvement. While disciplinary theories emphasize novelty, counter-intuitiveness or fine distinctions… a practical theory uses only those distinctions or novel ideas that can reliably motivate practitioner action in diverse contexts. A practical theory is also not a *general education* theory.… The virtue of a practical theory is that each element is immediately recognizable to both practical experts and theoretical experts, each of whom deeply understands

the problem of practice, through their own lenses. Such theories function as useful guides for practice improvement while remaining grounded in current scientific knowledge" (pp. 17-18).

[viii] It is worth noting the most commonly identified attributes of successful individual skill building were supported by the programs designs. These were strong skill building programs. Within the standards and curriculum features, the following generic aspects of skill development are supported: *Skills beget skills* suggests that youth who enter a setting with greater skill are more likely to leave with greater skill because skills build on skills; *non-linearity of skill growth* suggests that skill development occurs in a scalloped shape, with repeated cycles of progress and backsliding and requires lots of practice; *simultaneous integration of multiple skills* implies that skills are not built one at a time but are deeply integrated into networks of skill or skill sets; *skill transfer* requires practice at transfer and suggests that for youth to learn to apply skills outside of the context in which they were learned, they have to practice using those skills in different contexts; *skill building is context dependent* in the sense that skills grow primarily in relation to context and are often most easily demonstrated in the contexts in which they were learned (Baker, 2014; Fischer & Van Geert, 2014; Heckman, 2008, 2013).

[ix] The Challenge study design is a fully mixed, sequential, equal status research design (Leech & Onwuegbuzie, 2009), meaning (among other things) that the study produces primary findings from both qualitative and quantitative aspects of the design. The design draws upon formative evaluation principles (Brown, 1992) that include attention to granularity of theory and description, diverse contexts and subgroups, description at multiple levels of organization as well as a focus on different types of individual change over time (Lavelli, Pantoja, Hsu, Messinger, & Fogel, 2008).

[x] Wyman's Teen Outreach Program (TOP) participated in several empirical research studies to evaluate both the behavioral outcomes and the process mechanisms that lead to positive outcomes for TOP participants. Read about the studies and future plans for additional research at: www.wymancenter.org. VOBS is a member of Outward Bound USA, an organization that assures consistent program standards across many experiential education providers. Read more at http://www.outwardbound.org. YW Boston's Youth Leadership Initiative is an adaptation of the national model of Anytown, a summer social justice experience developed and spread through the National Conference for Community and Justice (NCCJ). Read more at: https://nccj.org. TPP has been replicated by partner organizations in Charlotte, NC (2000-present); Los Angeles and Santa Barbara, CA (2001-2012); Chicago, IL (2003-4); Baton Rouge, LA (2003-2012); Rochester, NY (2008-present); Israel (2000-2007); and Cape Town, South Africa (2004-2008). Read more at http://the-possibility-project.org.

[xi] The Curriculum Features include five features and eighteen practice indicators. These features could have been referred to as "standards." We decided to distinguish the curriculum features because they were an exploratory result of our more intensive efforts to define promising practices in the six domains. We spent more time and methodological rigor on the development of the standards in the six SEL domains.

[xii] This learning cycle is widely attributed to John Dewey (Dewey, 1938) and/or Jean Piaget (Piaget, 1977), and an experimental literature supports educational interventions based on this model (Miettinen, 2000). This model also parallels the skill-building theory whereby individuals practice skills in relation to challenges and, upon mastery, are both more likely to successfully apply those skills in other settings and to engage with challenges of greater complexity (Fischer, 1980).

[xiii] This design has a history in adolescent group-work, camps, and community service interventions for high risk, high potential youth (Oden, Kelly, Ma, & Weikart, 1992) and has been usefully referred to as a "broad developmentally focused intervention" (Allen & Philliber, 2001) to emphasize the point that the interventions work across the wide variation in the SEL histories of adolescents.

This intervention design could easily be gleaned from the "blue book" synthesis of positive youth development research and the curriculum features named in the SEL Guide look a lot like the features of positive youth development programs named in this synthesis (Eccles & Gootman, 2002).

[xiv] An emphasis on process and/or content choices as an element of positive youth development and program quality is old news to the OST field. The way in which problem solving complexity – and the necessity for making consequential decisions that require adult support (i.e., co-regulation for complex choice) – was built into the project curriculum is well described in the "capabilities-complexity" model in Oosterhof et al. (2008).

[xv] In this report, we have not dealt explicitly with motivation as an aspect of mental skill or behavior, largely because motivation is difficult to define in terms of either skill or behavior. The point here is that beliefs about the self (e.g., who I am, what I know how to do, where I belong) contribute to the function of motivation (e.g., to promote engagement) as individuals encounter context elements that are in some way aligned with or fit to their existing belief systems. "Researchers interested in self and motivation … [have explored] how cognitive processes (e.g., attributions) and representations (e.g., beliefs) serve basic motivational and self-regulatory functions such as directing attention to, and speeding the processing of, particular features and domains of self-relevant experience, people, and situations. Eventually, the notion of self-representations provided a way of understanding how, despite limits on information-processing capacity, humans efficiently and often automatically motivate and regulate their experience and behavior (Bargh & Chartrand, 1999). It also provided a view of self (with its motivational and regulatory functions) that was consistent with James' (1890) description of the content-rich 'empirical me-selves' and is the basis for what are today referred to as self-schemas, self-theories, self-concepts of ability, self-efficacy beliefs, possible selves, beliefs about intelligence, educational aspirations, school values, achievement goals, and so on" (Roeser, Peck, & Nasir, 2006, p. 395). The Basic Levels of Self model (Roeser et al., 2006) defines self in terms of several levels of representation (i.e., temperamental, iconic, symbolic, and phenomenological) and indicates that contents and processes within these levels provides both energy and direction to behavior, regardless of the extent to which they invoke *awareness* (aka, the I-self). Specifically, "*motivation* (i.e., the energization of behavior) is a key *function* of self at each of these levels… [and]*regulation* (i.e., the direction of behavior) is also a key *function* of self at all of these levels and ranges from reactive and relatively automatic forms at the me-self levels to volitional and effortful forms at the I-self level" (pp. 417-418).

[xvi] For OST programs, the afterschool academic enrichment model is a clear example, wherein trips to museums and other cultural institutions build academic background knowledge (Marzano, 2004) that is likely to be activated in school day classrooms at some point in the future. More background knowledge creates more opportunities for contexts to activate skills (and motivation, as discussed in Endnote xv).

[xvii] A similar point is made in the two under-referenced meta-analyses from the field of the learning sciences that focus on the effectiveness of specific types of practice. Instructional practice associated with explicit attention to how symbolic level beliefs (e.g., knowledge, attitudes, goals, plans) are organized – meta-cognitive operations – were associated with largest effect sizes on skills (Hattie, 2009; Marzano, 1998).

[xviii] In fact, there is a growing "mindfulness" literature, involving both staff and students, that describes the positive effects associated with what we here refer to as volitional shifting and focusing of conscious awareness (Davidson et al., 2012; Jennings, Frank, Snowberg, Coccia, & Greenberg, 2013; Lovelace, Eilenfeld, & Francese, 2014; Roeser, 2016; Roeser et al., 2013).

[xix] Mastery experience, social modeling, persuasion, and positive affect are thought to build efficacy beliefs (Bandura, 1977, 1992). In particular, the literature on skill building and scaffolding learning suggest that the optimal *zone of proximal development* occurs when tasks are designed to place efficacy beliefs slightly above the actual ability necessary for a given task (Fischer & Bidell, 2006). The standards for SEL practices suggest that the SEL Challenge programs excel at these methods.

[xx] Belief and behavior measures for the SEL Challenge study were correlated near zero in all domains. The "correspondence" issue has been a subject of extensive study, suggesting that only as the referent of belief measures get more behavior- and context- specific does the correlation with behavior measures become substantively meaningful (e.g., Ajzen, 2005; Ajzen & Fishbein, 1977).

[xxi] The Weikart Center provides its clients access to performance measure comparison samples from a de-identified reference database. The database is an aggregation of de-identified data from all Weikart Center clients. Because not all measures are used with all clients, the number of sites represented in each normative comparison varies between 42 and 172 sites over several programs years. Although these numbers could be increased with additional resources, we believe that the comparison group levels will not change substantially with additional sites.

The Social and Emotional Learning Challenge was designed to identify promising practices for building skills in six areas: emotion management, empathy, teamwork, initiative, responsibility, and problem solving. The Challenge was a partnership between expert practitioners delivering exemplary programs in eight unique communities, a team of researchers, and a national foundation. This technical report includes a discussion of theory and methodology used to produce the SEL Challenge findings. This technical report is written for agency leaders, evaluators, funders, consultants, and policy makers who want to assess the validity of the SEL Challenge findings or mount their own local Challenges.

Preparing Youth to Thrive: Methodology & Findings from the SEL Challenge is brought to you by:

SC/e + David P. Weikart Center for Youth Program Quality