Making the Grade: A Progress Report and Next Steps for Integrated Student Supports

By Kristin Anderson Moore, Hannah Lantos, Rebecca Jones, Ann Schindler, Jonathan Belford, and Vanessa Sacks

Child Trends

December 2017

Acknowledgments

Reports of this depth and length do not happen without the help of innumerable individuals. Though we cannot call out the principals and schools that participated by name, their insights and time were invaluable. The work they do daily as they commit to serving children with kindness, intention, and purpose inspires us.

While we worked on this report, we also worked with a team in the Center for the Improvement of Student Learning (CISL) at the Washington State Office of the Superintendent of Public Instruction (OSPI) to support the development of an integrated student supports protocol for the state of Washington. Their exciting work is available on their website.[1] Our conversations with the team in Washington – Andrea Cobb, Kelcey Schmitz, and Amber Palmer – helped shape our thinking about the questions remaining around ISS implementation and their insights and questions were greatly appreciated.

In addition, the report was reviewed by several key stakeholders both in and outside of Child Trends. We thank our Child Trends colleagues – Carol Emig, Kristen Harper, Deb Temkin, Alicia Torres, and Brent Franklin – who reviewed the report. The report was also reviewed by several colleagues outside Child Trends, including Reuben Jacobson at the Coalition for Community Schools, Heather Clawson at Communities in Schools, Betina Jean-Louis at Harlem Children's Zone, and Erin Sibley and Mary Walsh at City Connects. All of their insights and feedback were thoughtful, helpful, and made the report stronger.

Finally, we thank AT&T for its generous support for this study.

All photos courtesy of Allison Shelley/The Verbatim Agency for American Education: Images of Teachers and Students in Action.

[1] See the OSPI website for the Washington Integrated Student Supports Protocol (WISSP) and the legislative report that supported that work: http://www.k12.wa.us/CISL/ISS/default.aspx.

Table of Contents

Executive Summary

By Kristin Anderson Moore, Hannah Lantos, Kristen Harper, and Rebecca Jones

Introduction

In recent years, the education field has come to recognize the role of schools in supporting student health, safety, and well-being by developing ***integrated student support*** initiatives. These offer specific services and supports to students and their families to build a foundation for academic success. These initiatives, referred to as community schools and wraparound supports as well as integrated student supports models, help schools connect struggling children with secure housing, medical care, food assistance, tutoring, and other critical supports. While they are understood to be vital components of community efforts on behalf of children and families, they also further our nation's collective efforts to close education opportunity gaps, raise graduation rates, and better compete on the international stage.

Child Trends evaluated these initiatives in a 2014 overview of the evidence regarding integrated student supports (ISS)—implementation models in which schools secure and deliver coordinated, school-based supports that target various barriers to student achievement.[1] In general, ISS relies on five essential elements to support service delivery: community partnerships, student support coordination, integration into the school setting, needs assessments, and data tracking. The 2014 overview clarified that ISS was an emerging field of practice. With limited rigorous evaluations, Child Trends' researchers posited that ISS was a promising way to improve academic outcomes and see a substantial return on investment.

Since then, interest in ISS models has grown. Educational achievement remains a major vehicle for individual and family success. Although the high school graduation rate has risen over the past decade, the United States still lags behind other countries, and large disparities persist in academic outcomes. ISS models aim to bolster academic performance by recognizing the importance of addressing students'

[1] Moore, K.A., Caal, S., Carney, R., Lippman, L., Li, W., et al. (2014). *Making the Grade: Assessing the Evidence for Integrated Student Supports. Child Trends*. Bethesda, MD. Available at: https://www.childtrends.org/publications/making-the-grade-assessing-the-evidence-for-integrated-student-supports/.

nonacademic needs. Indeed, the 2015 reauthorization of the federal Elementary and Secondary Education Act (the Every Student Succeeds Act, or ESSA) encourages implementation of ISS for the first time. As written, ESSA now expressly permits schools and school districts to incorporate ISS into Title I targeted assistance programs for eligible students at risk of failing state academic achievement standards, and into Title VI, Part A activities that support student health and safety. Further, ESSA now makes available new federal formula dollars to states (under Title VI, Part A) to implement models that address student health, which could be utilized to support broader ISS models.

With ISS now codified in federal law and expanding across the country, school districts and principals are in need of a more current review of the evidence to guide school implementation. To this end, Child Trends updated its review with a synthesis of findings from relevant resources—including evaluations, child development research and theory, implementation reports, interviews with principals, benefit/cost analyses, and analyses using the Social Genome Microsimulation model.

Key Findings

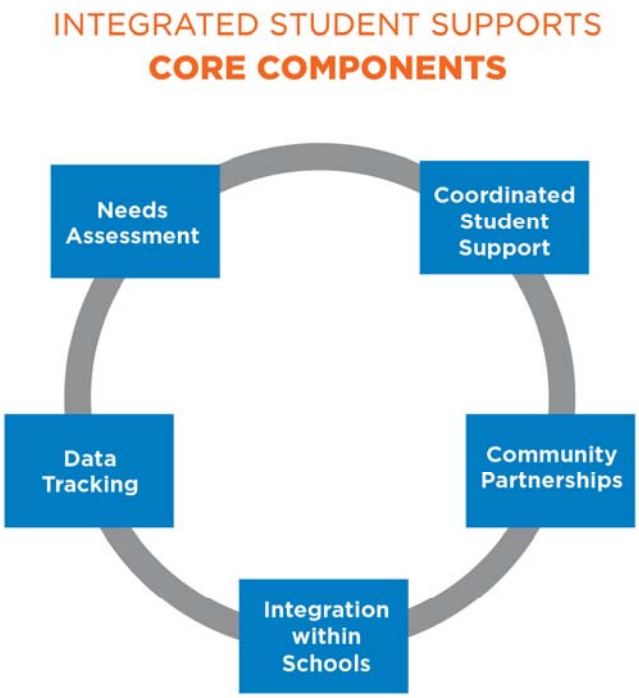
Based on this updated review, the authors are optimistic about the effectiveness of ISS. The report highlights a growing evidence base in support of ISS while serving as a reminder to the field that the evidence is not yet complete.

- Evaluation studies find a mix of positive and null (non-significant) findings, but there are virtually no negative effects across the evaluations.
- Several strong evaluations find support for particular ISS models, including City Connects, Communities in Schools in Chicago, the Harlem Children's Zone's Promise Academy, and Diplomas Now.
- New evidence from an application of a microsimulation model, which allows for a forecast of long-term outcomes—as well as evidence from four benefit/cost studies—finds that students' participation in effective ISS interventions will have long-term benefits.
- In addition to this evidence, the ISS model continues to rest on a solid base of research and best practices from child development research and theory.
- While the five essential components of ISS models (Figure 1) continue to support service delivery, identification of the specific, concrete elements that comprise successful implementation of each ISS component—and how they are implemented—is evolving slowly among researchers and educators. This work represents the critical frontier for research and practice.
- High-quality program implementation is important and will require adequate resources.
- Nonacademic outcomes are rarely measured as part of the evaluations, even though they are central to the conceptual model, which limits our understanding of the mechanisms driving ISS success.

What Are Integrated Student Supports?

ISS models recognize that students' unmet nonacademic needs can undermine their academic success. However, the types of nonacademic needs addressed vary across programs and across evaluations. In general, the supports provided under ISS models can include academic supports, housing assistance and food supplies, medical care, and mental and behavioral health services, and may go beyond student needs to provide critical services to parents and families. Moreover, the lack of consistency in the language used to describe ISS makes it challenging to discern which core services are necessary to make the ISS approach effective. Nevertheless, whatever the terminology, there is now widespread recognition that positive investments to address nonacademic needs are essential to student success.

Figure ES1. Core Components of Integrated Student



Expansion of ISS Models Across the United States

Every state in the country now has schools that use ISS models. Formal programs—such as Communities in Schools, City Connects, or community schools more broadly—have contributed to the rapid nationwide expansion of ISS models in the last decade. However, ISS models have also expanded informally, school by school, because experienced principals and staff who work directly in schools recognize the importance of supporting students' nonacademic needs in structured and systematic ways. While academic success remains the primary goal of educators, they recognize (based on their on-the-ground experience) that addressing both academic and nonacademic needs is necessary to reach this goal. Because ISS programs are most likely to operate in schools that serve large numbers of low-income students and students of color, they have the potential to reduce disparities by improving the academic outcomes of some of the most vulnerable students.

Key Findings Explained

Evaluation studies find a mix of positive and null (non-significant) findings, but there are virtually no negative effects across the evaluations.

The evidence base for ISS grew from approximately 11 rigorous evaluation studies (either randomized control trial or quasi-experimental design studies) in the 2014 review to a total of 19 in this 2017 update. The evaluation findings are promising and suggest that the ISS model is tipping results in the right direction. Specifically, this updated review of evaluation studies indicates that ISS interventions have mostly positive or null (statistically non-significant) results, and that negative findings are rare. There were only two negative outcomes among these 19 rigorous evaluations. Positive results can be seen across the studies for a variety of outcomes, including attendance, grades, test scores, graduation, and GPAs. Additionally, we continue to see positive results when different measures are used to examine similar outcomes, suggesting that these results can withstand varied types of measurement. However, these positive results are interspersed with numerous null results, suggesting that ISS is a promising but not yet proven approach.

Several strong evaluations find support for particular ISS models.

The evaluation studies with the strongest methodologies find more consistently positive impacts, including the evaluations from CIS in Chicago, City Connects, Diplomas Now, and the Harlem Children's Zone's Promise Academy. This likely reflects both the strength of these programs and the choice of an appropriate evaluation design. A lack of positive results in an evaluation, either negative or null, could mean that the program was not effective or was poorly implemented, or that the evaluation was inappropriately designed. Examples of poor design include studies that did not include enough participants to measure change, outcomes that were inappropriate for the inputs of the program, or a comparison group that was not truly similar.

New evidence from an application of a microsimulation model, which forecasts long-term outcomes—and evidence from four benefit/cost studies—finds that students' participation in effective ISS interventions will have long-term benefits.

Four benefit/cost studies have been conducted to date. Although all four studies used very different approaches and estimation methods, each shows strong returns on investment (ROI). Based on these studies, ROI estimates range from $3 to more than $14; that is, for every dollar invested, a return of at least $3 and up to $14 can be anticipated.

Child Trends augmented findings from these benefit/cost studies with analyses from microsimulations that use the Social Genome Model (SGM) (developed by the Brookings Institution with input from Child Trends, and now managed by Child Trends and the Urban Institute). Results from rigorous evaluations were incorporated into the SGM to assess whether and how ISS enhances income at age 29. These

analyses suggest modest but real improvements in the estimated incomes of individuals in their late twenties, due to better math scores, higher graduation rates, lower rates of incarceration, and a lower incidence of teen pregnancy.

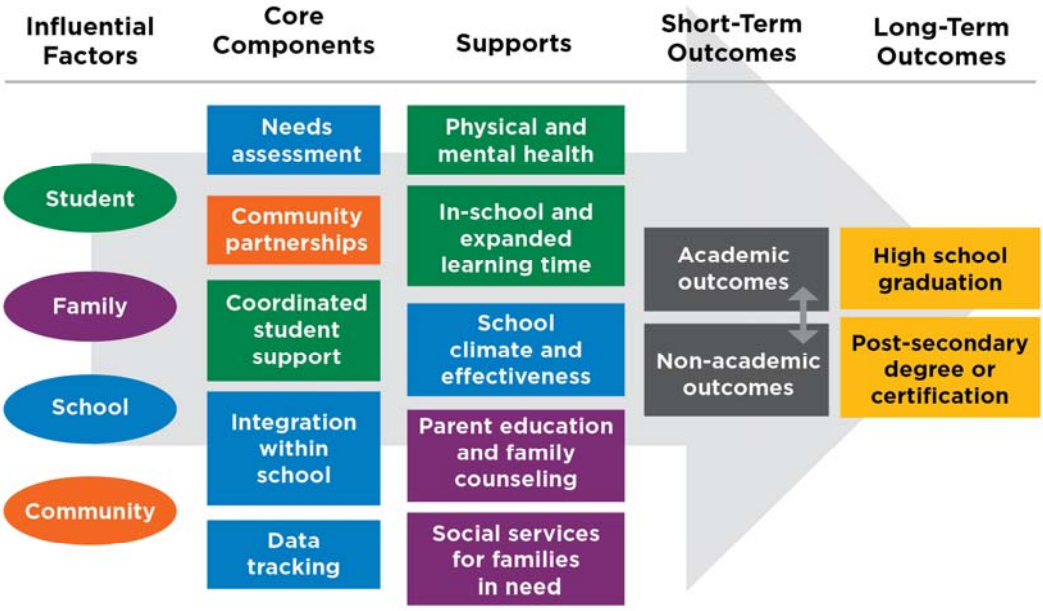
The ISS model reflects principles and best practices from child development research and theory.

One last key finding about the approach overall highlights what we already know about ISS models: they are aligned with widely accepted child development research and theory. For example, ISS models align well with the following bodies of research and theory:

- *Whole child model*: health, behavior, emotional, and academic factors are all recognized as important for children's development
- *Ecological approach*: ISS is consistent with models that acknowledge the unique ways in which child-, family-, school-, and community-level factors contribute to each student's academic success
- *Life course perspective*: ISS recognizes that earlier education experiences, including academic and nonacademic school experiences, affect later accomplishments
- *Child-centered*: ISS recognizes that programs should focus on students' needs (rather than those of the school or adults), and acknowledges the value of tailoring interventions and approaches to the needs of each individual child
- *Social determinants of health*: ISS acknowledges how contextual inequities can drive health inequities because the environment, services, and people surrounding a child can impact their health
- *Social and emotional competencies*: ISS recognizes that students' social-emotional skills affect their academic success
- *Soft skills*: ISS can support the delivery of services to build interpersonal and intrapersonal skills (like effective communication or conflict management), and recognizes their importance to success in work and life
- *Positive Youth Development (PYD)*: ISS is consistent with models that emphasize supportive approaches over punitive or didactic approaches, and acknowledges their added effectiveness in engaging students and helping them achieve their goals

The next three key findings focus on questions that remain to be answered by future research. Answers to these questions will allow ISS models to more strongly impact students' academic and nonacademic outcomes.

Figure ES2. Logic Model of the Five Core Components of ISS Models



While the five essential components of ISS models (Figure 1) continue to support service delivery, identification of the specific, concrete elements that comprise each component—and how they are implemented—is evolving slowly among researchers and educators. This work represents the critical frontier for research and practice.

Interviews with principals across the country highlight that the core components identified in 2014 (Figure 1) continue to describe the ISS model's approach.[2] However, an understanding of the concrete elements and strategies that effectively translate ISS models from theory to practice is evolving slowly. This involves an understanding of the critical elements that must be present in every model (so that fidelity can be defined), and of how high-quality implementation of these elements affects student success.

Unfortunately, the ideal process for implementation of ISS programs is not yet clear. In a time of limited budgets, schools want to know which practices are essential and which are not: are certain key elements required for ISS models to be successful? For example, do children need to have a positive relationship with a teacher in the school building for any of the other elements to work? Is one relationship enough? Do schools need to have a full-time ISS coordinator on-site? Interviews with principals suggest that having a coordinator dedicated to integration and coordination can make the difference between high and low impact for an ISS model in a school. However, these questions remain unanswered

[2] This model was developed in 2014 based on reviews of existing programs and input from stakeholders.

quantitatively because most evaluations do not include variables in their analyses about the specific mechanisms at work.

One factor that undermines high-quality implementation in schools is insufficient understanding of the essential elements for each of the five components. The other factor is a lack of awareness of what "quality" means for the core components and their constituent elements. As shown in Figure 2, the conceptual model relies on the expectation that the intermediate factors (nonacademic outcomes) will improve, and that these improvements will lead to better academic outcomes. However, poor or inconsistent program implementation may explain why some schools see results and others do not.

Accordingly, the educational field must discuss how to build a stronger body of research. Importantly, which elements are critical for a high-quality ISS implementation that ensures more consistently positive effects? This work represents a vital frontier for research and practice.

High-quality program implementation is important and will require adequate resources.

Studies of early childhood and youth development programs consistently demonstrate that high-quality implementation is associated with more positive outcomes. One Communities in Schools study that examined this topic continues to stand out, finding that a poorly implemented ISS program was no better than no program at all. Interestingly, each of the six implementation studies reviewed here highlighted different aspects of implementation, ranging from higher teacher-to-student ratios, to fidelity to the defined model, to a focus on specific outcomes identified in the organization's theory of change. The programs reviewed here for their implementation of various ISS models augmented our understanding of which key program parts are important for positive outcomes. However, as noted above, the key elements of quality are only beginning to be defined and examined.

For example, adequate resources are clearly required to carry out implementation tasks: a needs assessment, coordination, data collection, programming to meet needs unaddressed elsewhere, etc. School staff and principals may move forward with this work out of necessity, but doing it well over time will require dedicated ISS staff. In large schools, more than one staff person may be needed. Without staff who can dedicate their time to this work, these models are difficult to build and sustain.

Other critical elements may include staff who are committed to the ISS student-centered approach, the use of data to identify needs and monitor progress, a supportive and violence-free school, and the provision of services to students (and even to families) when barriers undermine learning. However, these elements must reflect current hypotheses based on the broader research literature, and they must be empirically tested.

Nonacademic outcomes are part of the conceptual model but are rarely measured in evaluations, which limits understanding of the mechanisms that drive ISS success.

While evaluations increasingly suggest that ISS may have positive impacts on academic outcomes, most evaluations included in this study did not examine *nonacademic* outcomes with much depth or nuance. While it is critical to monitor academic outcomes, some evaluations focus on these almost exclusively, which limits our understanding of the impact of ISS programs on nonacademic well-being. This is problematic because it is necessary to specify, measure, assess, and analyze data on nonacademic competencies to understand the critical mechanisms that lead to academic success. Are social skills the critical mechanism improved by ISS models, leading in turn to improvements in academic outcomes? Alternately, is the critical mechanism a student's concept of self, or their persistence or grit? This relative neglect of nonacademic outcomes is beginning to change, but there is still little consistency across studies regarding the competencies that are assessed or how they are measured when included. These nonacademic outcomes are part of the theory of change for ISS models (Figure 2), but until evaluations assess them fully and with consistency, there is insufficient evidence that the theory is wholly or partially correct. Most importantly, policymakers, principals, and school staff lack evidence-based information about the concrete practices to be implemented.

Four Key Areas for Further Research

First, **evaluation methodology impacts researchers' ability to state conclusions.** Decisions about evaluation design, comparison or control groups, measurement, length of implementation or follow-up for the study, and statistical analyses affect the kinds of conclusions that can be drawn. Some null findings likely stem from the inadequate methodologies used for analyses.[3] (Evaluation methodologies must be appropriate for each program in terms of timing, types of data, outcomes, etc.) Using different approaches in future evaluations may allow researchers to tease apart small but significant effects in a way that current studies were unable to do.

Second, **many evaluations continue to use slightly (or very) different measures of outcomes**, and measures may be obtained from different sources (e.g., student reports versus school records). When results differ with different measures, it is difficult to disentangle whether there is truly an effect or whether the effect is specific to certain outcomes. Encouraging greater use of the same measure or measures across studies would allow findings to be comparable.

Third, **studies tend to examine each outcome in isolation**. Researchers may control for confounding factors but infrequently conduct analyses that examine the unfolding process by which ISS models may affect outcomes. Structural equation models, for example, would allow analysis of intermediate/mediating nonacademic variables and how they relate to longer-term academic outcomes.

[3] Tables with the full findings can be found in chapter 4 of the report as well as in the appendices.

It is essential to include and study these nonacademic outcomes, as they will improve our understanding of whether various ISS models work as theorized.

Finally, if these models do work as theorized, what explains the difference between successful and unsuccessful programs? To answer some of the remaining questions about ISS, **more focus will be needed on program implementation**. Specifically, what explains success in some schools but not others that use the same approach? Are some implementation strategies more likely to result in better outcomes? We need to better understand implementation approaches and quality to identify critical factors and support achievement of higher-quality implementation by principals and teachers.

Key Takeaways for Stakeholders

Policymakers

Federal, state, tribal, and local policymakers can implement policies that are supportive of ISS. At the local, state, and tribal levels, policymakers can **provide resources for school-based coordinators**, help **develop lists of services** available in different communities, or require that schools **plan for integrated and coordinated supports** to students. Their state mandates can also explicitly emphasize the importance of integrated nonacademic supports in schools.

Federal agencies can support implementation of ISS provisions by providing **technical assistance products and services** that explain best ISS practices, aligning implementation with other popular student support frameworks and programs (e.g., Multi-Tiered Systems of Support and Social and Emotional Learning), and ensuring fiscal support for ISS implementation under federal formula programs. Further, such entities can **support research that might answer remaining questions,** and provide discretionary grant dollars to **states and districts to develop and sustain integrated models.**

Additionally, federal and state policymakers can make it easier to **link or braid funding streams** in schools—such as Medicaid, housing support, or Temporary Assistance for Needy Families—to meet the needs of students and their families. Some states are considering innovative ways to braid funding so that people in different fields (housing, healthcare, schooling, juvenile justice, etc.) can more easily work together.

Practitioners

Practitioners include teachers, principals, school staff, and staff in departments of education. Based on this updated review, principals and teachers now have further evidence that ISS models can be effective. In addition, **ISS aligns with research and theories on child development.** Using these theories to develop an integrated and coordinated support system for students in schools will likely result in better outcomes for children and their families.

Principals and teachers should explore ways to **align student support initiatives** meant to improve student development, health, and safety. Efforts to implement ISS need not compete with other models or programs that they employ or have heard about, such as Multi-Tiered Systems of Support or Positive

Behavioral Interventions and Supports. Rather, these approaches can build on one another, and educators can plan ways to pursue implementation as a single cohesive system.

Principals and teachers need to **collect data** to monitor the effect of ISS models on their schools and students. This would allow them to know that students are being reached and supported and can help the field identify the essential elements for a successful ISS school.

Experience on the ground suggests the **importance of having an ISS coordinator** in the school. Principals and teachers already work long hours, and few can assume the demands of building an integrated model that performs a needs assessment, develops community partnerships, coordinates student supports, integrates services within the school, and monitors progress for individual students and the school. While a coordinator would require funding, our interviews with principals suggest that it is crucial to successful implementation.

Researchers / evaluators

Several findings are relevant to researchers/evaluators. Researchers should prioritize understanding the **key mechanisms** that drive ISS models' success in the design of future evaluations. Using the **same outcomes across studies** would advance the field because researchers could more easily make comparisons. Many outcomes (both academic and nonacademic) in the various studies differ, making cross-evaluation comparisons difficult.

To advance the field, it is essential that researchers use the **most rigorous appropriate design** (given the timing of the study, data available, and program design constraints). A rigorous study design with data that do not match the program can result in null findings, which does a disservice to the program and the field.

Researchers and evaluators are learning the importance of **building school-level capacity** by helping schools conduct needs assessments, develop data systems, and identify ways to use performance management data to monitor student performance and identify ways to improve outcomes. Once these practices are in place, impact or outcome evaluations may be more productive.

Finally, there is a need to conduct quantitative studies (quite limited to date) that explore mechanisms of success with depth and nuance. Rigorous qualitative work also has much to add to the research literature on key ways in which ISS works.

Funders

This review suggests that funders should support evaluations that are **appropriately designed to accurately measure results**. Conducting a randomized control trial prematurely, or with methods or measures that do not align with critical questions, is not useful to programs or to the field. It would be more useful to conduct implementation or outcome evaluations that identify the critical mechanisms that make the ISS approach effective. Funders should not rush to randomized studies if the timing, available data, or study design is not conducive to a rigorous, quantitative study design. Answers to any

remaining questions about process, implementation, and qualitative findings are also needed and could be supported prior to a randomized control trial.

Achieving results **can take time**. Funders must invest in developing good programs, recognizing that both effective implementation and thoughtful evaluation take time. Changes in educational outcomes will not happen within a year of changing systems and practices within schools. Ideally, funders will support schools in conducting needs assessments, coordinating student supports, developing community partnerships, integrating student services, and using data to monitor progress.

Funders may seek to support a **consortium of researchers and practitioners to work together to identify critical constructs** for future evaluations, and provide a common set of measures for the field.

Conclusion

As a result of Child Trends' review of integrated student support models, the authors are cautiously optimistic about the potential for this approach to improve student outcomes, especially in schools with concentrations of at-risk students. Our caution is based on the large number of null findings, as well as the lack of evidence regarding the concrete elements that make different models successful or how they must be implemented.

With these cautions in mind, we nevertheless find that ISS models represent a promising approach to supporting students that aligns existing knowledge about child development with additional insight from dedicated, experienced practitioners. Moreover, as the knowledge base accumulates, positive or null findings are common, with rare negative findings. ISS interventions combine research-based learning with practitioner wisdom: they are student-centered, address the whole child in a positive way, develop students' soft and hard skills alike, and acknowledge both the struggles and the resilience seen in families, schools, and communities. Implementation of ISS models should remain flexible to changing needs, identify services and supports within the community, use data to identify needs and monitor progress, and conduct rigorous evaluations when appropriate.

Chapter 1: Introduction and Background

By Kristin Anderson Moore, Hannah Lantos, Rebecca Jones, and Ann Schindler

Introduction

Teachers and school administrators who interact with children daily know that nonacademic issues can undermine academic success. Research increasingly supports these practitioners' insight by confirming that nonacademic factors in a young person's life influence their ability to concentrate, learn,[i] process information,[ii] and behave well in class; in turn, these influence academic and life success and overall well-being.[iii] Students who suffer from poor physical[iv] or mental[v] health, who are homeless,[vi] who experience instability at home[vii], or who come to school hungry[viii] do less well in school.[ix]

Recently, policymakers, researchers, and education officials have begun to recognize that educators and schools can play critical roles in addressing children's barriers to learning.[x, xi] As policymakers and administrators have identified the linkages between children's negative life situations and academic outcomes, policy and programmatic approaches have begun to address nonacademic barriers, and both government and private funders are investing more resources to remove these barriers. Prominent education models (including Harlem Children's Zone's Promise Academy or Turn-Around for Children) and federal legislation (including the Every Student Succeeds Act [ESSA]) place a strong emphasis on the social, emotional, and health needs of students. They also emphasize the use of Multi-Tiered Systems of Support (MTSS)—such as Positive Behavioral Interventions and Supports (PBIS), integrated student support models (ISS), or Response to Intervention (RTI)[xii] (an overview of these is included in Appendix 5)—where a tiered model is used such that all students have a base level of services and needier students are targeted with more services, more intensive services, or different services that better meet their needs. Integrated student supports use this approach to bring in outside-of-school supports for students who need them. In 2014 Child Trends defined ISS models as "a school-based approach to promoting students' academic achievement and educational attainment by coordinating a seamless system of wraparound supports for the child, the family, and schools, to target students' academic and nonacademic barriers to learning."[xiii]

The underlying assumption that has guided work on ISS is that educational outcomes will improve when a variety of barriers to successful academic achievement are removed. These barriers might exist at the level of the student, the school, the family, or the community. These barriers might also be addressed by providing services that go beyond academic inputs such as tutoring. For example, interventions might be implemented for individual students to build their social skills and character, address trauma, or provide healthcare and mental health services to the student. Or, interventions might work at the school level to prevent bullying, reduce suspension and expulsion, and improve school climate. When needed, services might assist families with finding stable housing, obtaining healthy food, and/or finding a job or getting

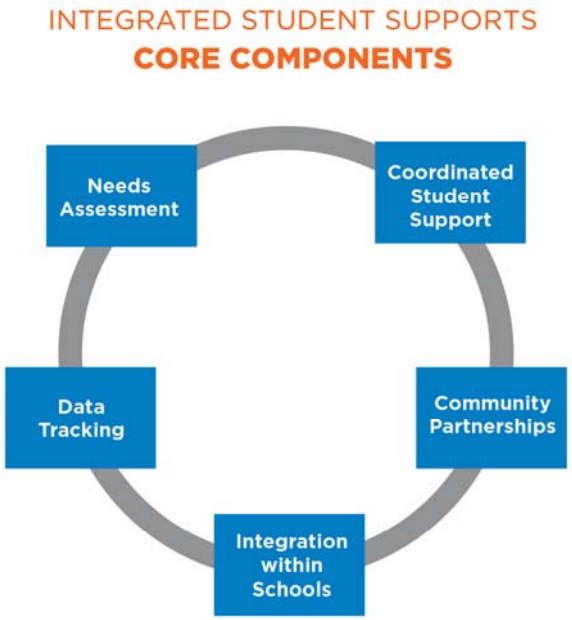
job training. Many existing resources highlight how integrating supports into schools aligns with ESSA regulations,[xiv] and we hope that this report highlights how to integrate, what questions about successful integration remain unanswered, and what is important to remember as integration moves forward.

Overview of Findings from the 2014 *Making the Grade* Report

To assess the quality and depth of the evidence available at the time, Child Trends published *Making the Grade: Assessing the Evidence for Integrated Student Supports in 2014*. This report aimed to assess whether integrating nonacademic services into academic settings has a positive effect on children's schooling outcomes. In addition, Making the Grade raised questions that would help inform future research and evaluation. All of the ISS models that Child Trends reviewed for the 2014 report (listed in Appendix 1) aimed to connect children and families in need to resources in the community and the school.

The 2014 report conducted three important types of analyses. (Chapter 2 provides a more complete summary.) First, after a thorough review of research and existing programs, Child Trends researchers defined ISS by developing a conceptual model that depicts the processes that underlie all the models and programs (see Figure 1, below, for the conceptual model developed in 2014). This figure identifies the five core components through which ISS can enhance student outcomes in both the short- and long-term. These five core components are: conducting a needs assessment, coordinating supports across the school and outside organizations, developing community partnerships to meet needs outside of the school, integrating supports and processes within the school, and collecting data to report on reach and outcomes.

Figure 1: Core Components of the Integrated Student Supports



Second, the report reviewed existing evidence on what variables are most important to student success in school and conducted new analyses using the National Educational Longitudinal Study (NELS) of 1988. Child Trends researchers concluded that there is no "silver bullet" that improves academic outcomes; rather, it is the power of combining many positive, developmentally appropriate assets that results in improved outcomes. Each individual factor has a relatively small effect, suggesting a need for the comprehensive approach ISS provides.

Finally, the report reviewed existing rigorous evaluations of ISS programs to assess whether the scientific evidence found that the ISS approach improves academic outcomes. Outcomes, implementation, and cost-effectiveness evaluations were examined if they used a statistically rigorous methodology.

Ultimately, five important conclusions were drawn:

- There is emerging evidence that ISS models can contribute to student academic progress.
- Available studies find a positive return on investment.
- ISS is a student-centered approach firmly grounded in the child and youth development research and literature.
- ISS is aligned with empirical research on the varied factors that promote educational success.
- High-quality implementation is essential to producing positive outcomes.

Child Trends' researchers concluded that ISS models are a "promising approach for helping more disadvantaged children and youth improve in school and have a brighter path in life." They also noted that the ISS approach is "solidly based in the literature on child and youth development, practitioner experiences, and studies of education."

However, that report concluded that "the evidence base is emerging" and that the approach is "promising," rather than unquestionably effective, based on several limitations observed in the available literature. Many of the studies produced nonsignificant findings where differences between intervention and comparison groups were statistically similar in terms of attendance, behavior, or course grades. Also, few studies were able to track outcomes beyond 1–2 years, making it hard to assess whether improvements persisted; few employed similar statistical methods or compared the same outcomes; and several outcomes were only studied once. This made outcomes with only one finding or conflicting findings challenging to interpret.

In addition, implementation of the models is quite varied. All models incorporate the five general elements depicted in Figure 1, and their goals are always to enhance children's opportunities for school success. Beyond this, the models evaluated differed in terms of what they provide to whom, how they are funded, and the types of support provided by outside organizations.

Additionally, few nonacademic outcomes were assessed, and most evaluations did not focus on whether the expected pathways actually matter. That is, we did not yet have evidence that improving any

particular nonacademic outcome translates into stronger academic outcomes. Therefore, although ISS is based on solid research and theory—indicating that family and community and nonacademic factors are the pathways that mediate the relationship between inputs and school performance—few evaluations explored these nonacademic mediators. Without these analyses, we cannot identify what worked and what did not. Moreover, without assessment of those intermediate outcomes, the findings are difficult to understand and contextualize. This is particularly true for null findings: was the theory incorrect, were the programs not good enough, or were they not implemented for long enough to detect differences?

Further, only a few evaluations incorporated implementation evaluations that explored issues of quality and fidelity. Consequently, the precise "ingredients" that comprise each of the common elements have not yet been identified. Therefore, educators do not know which key ingredients foster success. As new sites seek to implement the ISS approach, they understandably want to know what to do and how to do it.

Finally, the three cost-effectiveness studies available at the time used different quantitative approaches and different measures of cost, and different community supports were included. In addition, programs differ in how many services they provide in-house versus in the community. These differences made it hard to compare and draw precise conclusions about costs and benefits. That said, all three studies did report very positive returns on investment.

Education Matters

Although the four-year high school graduation rate has inched upward, from 73 percent in 2001[xv] to 79 percent in 2010[xvi] to 83 percent in 2015,[xvii] students' educational progress has remained sluggish. Researchers consistently find that reading and math proficiency have a positive association with high school completion and college attainment,[xviii] but performance on assessments in both of these areas has recently declined. In 2015, National Assessment of Educational Progress (NAEP) reading scores for eighth- and twelfth-graders, which are a measure of reading proficiency, declined for the first time in ten years after consistent but moderate increases.[xix] Similarly, NAEP mathematics scores for eighth-graders decreased after almost 20 years of increases, which began initially when accommodations were permitted for students with disabilities and those with limited English proficiency. The linkages between high school and further higher education are also fairly weak. By 2015, just over one-third of young adults ages 25–29 had obtained a bachelor's degree or higher despite overall upward trends in high school graduation and college enrollment.[xx]

Completing high school and continuing to further higher education are strongly correlated with income. According to the Bureau of Labor Statistics, annual earnings for someone with less than a high school degree in 2016 were approximately $26,000, while those with a high school degree earned nearly $36,000 and those with a bachelor's degree earned $60,000.[xxi] Given the strong relationship between income or opportunity and educational attainment, it is increasingly important to ensure that all children are provided with educational opportunities that position them for success. Shifting from a "one size fits all" school model to an ISS model that works toward providing for the individual needs of each child is one way to promote positive outcomes for all youth.

In addition to concerning recent trends, racial and ethnic disparities remain of significant concern. Despite the percentage of high school graduates increasing over time, significant disparities persist. American Indian/Alaska Native students have the lowest high school completion at 72 percent, while black and Hispanic students are similar at 75 percent and 78 percent, respectively. White students have a high school completion rate of 88 percent, while 90 percent of Asian/Pacific Islander students graduate from high school.[xxii] Numbers are particularly low for young adults who come from lower-income families,[xxiii] young men of color,[xxiv, xxv] or those from communities that are predominantly racially or ethnically segregated (and often very low-resourced).[xxvi] Unfortunately, these disparities are linked to other life outcomes. Researchers have found that higher levels of educational attainment are associated with higher wages, better health, higher levels of socio-emotional well-being, lower unemployment, and lower risk of living in poverty.[xxvii] These gaps have widened over time, suggesting that ISS models may also be an important strategy to reduce disparities experienced by some of the most vulnerable children in the United States.

The Last Three Years

In the three years since the initial report, both up-take of ISS and legislation have changed in important ways that make revisiting the evidence timely as the needs remain. First, more students are served by schools that are integrating nonacademic supports into the school. For example, today, Communities in Schools (the largest ISS provider in the United States) serves 200,000 more children than just two years ago (for a total of 1.5 million students in 2,300 sites).[xxviii] Second, in December 2015, Congress passed the Every Student Succeeds Act (ESSA). This measure passed with bipartisan congressional support to reauthorize the 50-year-old Elementary and Secondary Education Act (ESEA). In ESSA, for the first time, legislators encourage the implementation of integrated student supports. As written, ESEA now expressly permits schools and school districts to incorporate ISS into Title I targeted assistance programs for eligible students at risk of failing state academic achievement standards; and Title VI, Part A activities to support student health and safety.

With ISS codified in federal statute, federal officials have also made new funds available to states to implement models that address student health. After eliminating ESEA formula dollars for safety and health in 2009, legislators designated $400 million for the new Student Support and Academic Enrichment program in 2017—providing states with a new source of funding to promote student nutrition, physical activity and fitness, and social emotional learning. Together with the language in support of ISS implementation, these funds provide states with a new foundation to address the academic and behavioral challenges facing schools today—from bullying, to school violence, to school discipline, to chronic absenteeism. As federal, state, tribal, and district officials are still in the early stages of implementing the reauthorized federal law, it is critically important that policymakers, practitioners, and communities have ready access to the latest research on ISS to inform broad-based planning on how schools will support children and their families.

Thus, a better understanding of whether ISS models have moved from “promising” to clearly effective is essential to help states and local school districts prioritize what programming to implement under ESSA.

Child Trends undertook an update to the review to assess whether new research findings solidify, enhance, or revise the evidence for ISS.

Updated Review of Research on Child Development

In 2014, Child Trends' initial review of research identified major theories about child and youth development based in research, and compared their precepts with the elements of the ISS model: the whole child perspective, a child-centered focus, a life-course perspective, and the ecological model. These research-based theories supported our conclusion that ISS is a promising model because it fits so well into all of them.[xxix] In fact, everything we know about child development from the theoretical literature supports an integrated approach to supporting all aspects of a child's life. This continues to be the case as we update our review of theoretical perspectives in the field of child development. These are reviewed briefly below.

The **whole child model** recognizes that children's development is multi-faceted. One cannot study children's education without understanding how their physical health and safety, psychological/emotional development, and social and behavioral development affect cognitive development and educational achievement. That is, the whole child perspective recognizes that development in one domain affects development in another.[xxx,xxxi,xxxii] Eye glasses provide an example: a child who cannot see the board very likely cannot see the math examples that the teacher demonstrates in front of the class, creating at least one barrier to academic success. Recognizing that their physical ability to see impacts their ability to learn sees them as a whole child rather than just a learner or a patient in the optometrist's office.

A **child-centered focus** treats each child as an individual with unique strengths and needs, with the understanding that one uniform approach for every child cannot meet every child's individual needs.[xxxiii] This is particularly important in underserved communities and underserved schools. While most research on ISS does not look at racial or ethnic minorities, the research that is available indicates that the factors that are important to the general student population are also important for students of color. With the passage of the Every Child Succeeds Act (ESSA), Title 1, Part A funding requires that low-performing schools, which tend to also serve minority populations, must implement interventions that fall in Tiers 1, 2, or 3 evidence-based categories. Another aspect of this child-centered focus is that the student – not the adult faculty or the school itself – is the center of the school's mission.

Additionally, researchers have found that experiences—good and bad—in the early stages of development affect development and well-being in later stages of life.[xxxiv,xxxv] This is the basis of **life course models**, which argue that life experiences build on one another over time. For instance, if a student had a kindergarten teacher who criticized their early efforts to read, they may be fearful of reading—not because they are unable but because they were told they were unable. Understanding a student's background is important to effectively serve children in schools.

The **ecological model** recognizes that children do not live in a vacuum, but are influenced by many factors. Importantly, it places children at the center of concentric circles to illustrate that a child's

development is influenced by their biology, family, friends, community, and school environment, as well as the larger society. In other words, how a child does in school is impacted by various factors, including their innate ability, their family's values toward education, whether their friends are supportive of school success, the financial and other resources that exist in their community, and much broader state, tribal, and federal education policies. This perspective is already included in some prominent education approaches like the "Whole School, Whole Community, Whole Child (WSCC)" model,[xxxvi] which aims to situate children inside of schools and then communities when thinking about necessary interventions. Another model where this thinking dominates is the community school model. Community schools do this work by becoming a central part of the community, open to everyone, and creating partnerships between the school and community to meet academic, health and social service, and community development and engagement needs.[xxxvii] The Coalition for Community Schools framework builds visually on the ecological model with concentric circles.[xxxviii] Recent research finds that schools implementing the community school model seek to provide and address those opportunity gaps in schools where poverty and racism impact communities with fewer resources.[xxxix]

In addition to these theories, we highlight several additional approaches that are relevant to ISS. These include the social determinants of health, social and emotional learning (SEL), soft skills, equity, prevention, and implementation science. Again, all theoretical literature points toward ISS as a useful approach to support children in schools.

One theoretical frame prominent in the public health field and very relevant for ISS research is the **social determinants of health (SDH)**. The World Health Organization defines SDH as "the conditions in which people are born, grow, work, live, and age, and the wider set of forces and systems shaping the conditions of daily life. These forces and systems include economic policies and systems, development agendas, social norms, social policies and political systems."[xl] The social determinants of health are relevant for ISS work because understanding why some children (or their families) may have more health struggles can shed light on what resources are necessary to support them and their families. For instance, children living in polluted neighborhoods may be more likely to be absent from school due to asthma.[xli] As shown by this definition and example, there is overlap between the social determinants of health, the ecological model, and the whole child model. All try to capture the ways in which context and different dimensions of people's lives affect their health and opportunities to be healthy, successful, and financially stable. We include this here because it is important to recognize that different fields of study that are relevant to ISS use different language. SDH is commonly used in the public health literature, and ISS has been described as a public health approach to education.[xlii, xliii] Thus, it is useful to understand the public health language more explicitly.

It is also important to highlight the social, emotional, and soft skills that children can acquire in ISS schools because these skills are increasingly sought in the community and the labor market. The Collaborative for Academic, Social, and Emotional Learning (CASEL) defines **social and emotional learning** (SEL) as "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."[xliv] Many researchers have found a positive association between social and

emotional learning and academic outcomes.[xlv,xlvi,xlvii] The theory hypothesizes that, by improving how youth cope with anger, stress, and disappointment and supporting them in developing a growth mindset,[xlviii] academic outcomes should improve. An issue of *The Future of Children* focuses extensively on how SEL can improve schooling outcomes, reduce disparities, be taught in or outside of school time, and be reinforced with developmentally appropriate strategies across the life course.[xlix] Additionally, cost estimates have found substantial benefits to investments in SEL, with net present values exceeding, and often greatly exceeding, current cost levels.[l]

Building on skills stressed by SEL, **soft skills** tend to focus on job market competitiveness—the importance of having not only technical and academic skills when applying to jobs, but also the "softer" skills of communication, negotiation, emotion management, flexibility, etc. Children growing up today face an economy in which low-skilled jobs are becoming rarer and higher-order skills (specifically soft skills)[li] are increasingly necessary for employment with a livable wage.[lii,liii,liv,lv] Many jobs in today's market do not explicitly require mastery of academic content and instead require an ability to communicate, learn new skills and content, and work collaboratively.[lvi] Many young people lack opportunities to gain these skills, making opportunities to develop and practice soft skills in school increasingly important. Child Trends' report on soft skills from 2014 found that, increasingly, the evidence suggests that soft skills can be as (or more) important than either academic or technical skills in terms of predicting employment and earnings in the long run.[lvii, lviii]

Work on soft skills overlaps with the SEL research in focusing on the types of skills like communication, emotion management, empathy, etc. that are needed to be successful and move up in the workplace in today's economy. ISS, with its integrated approach, may be able to weave these skills throughout a school and its programs. Without including these soft skills as well as social and emotional skills in measures of outcomes, studies may be missing a critical component of success.[lix]

In addition to theories that focus more on understanding individual children, their contexts, and their strengths and challenges, a major focus of ISS interventions is **equity**. This is important because the federal legislation is also focused on creating more equitable opportunities for all children: ESSA requires Title I, Part A funding for low-performing schools to implement interventions that fall in Tiers 1–3.[lx] Schools that seek to implement an ISS model with quality and integrity tend to be those that are also trying to address opportunity gaps in schools due to poverty and racism and their placement in communities with fewer resources.[lxi] By using the ISS model to address individual needs, these schools are ensuring that the needs of every child are addressed, including students of racial, ethnic, or economically disadvantaged backgrounds.

Finally, we want to emphasize an increased focus in research and evaluation on **prevention** and **implementation**. Prevention is prominently situated in the ISS work because the goal is to identify students who might need supports *before* they have serious problems. One example is the early warning indicators that schools increasingly use. High rates of absences, for example, are associated with lower rates of high school graduation.[lxii] As early as kindergarten, absenteeism is associated with lower achievement in subsequent grades.[lxiii] ISS models frequently identify these children and provide them with the necessary individual and familial supports to *prevent* drop-out. In addition to prevention being

a key component of the ISS approach, high-quality implementation should be as well. We address this in more detail in Chapter 5, but note here that studies have concluded that high **implementation** has a positive relationship with positive outcomes.[lxiv,lxv] While programs and reform should be implemented with fidelity to ensure positive outcomes, considerations of variations in school context also matter.[lxvi]

All of these approaches reflect research centered on improving the skills, well-being, and resiliency of children—especially those who are most disadvantaged. Many overlap in their understanding of key drivers of children's success and of inequity; however, all of the approaches to learning currently getting attention in the literature are aligned with the ISS model as a tool to improve the academic outcomes of our nation's children.

Outline of This Report

This updated report is organized as follows: Chapter 2 summarizes findings from the previous report. Chapter 3 describes the methodology of the analyses conducted for this report. Chapter 4 reviews the new literature in the field, describes the micro-simulation, and discusses the evidence for outcomes evaluations, incorporating findings from previous and new evaluations. Chapter 5 reviews evidence for implementation evaluations, also incorporating findings from previous and new evaluations. Chapter 6 reviews evidence around cost-benefit analysis. Chapter 7 will finish with a brief summary, discussion, and recommendations.

Chapter 2: Summary of 2014's *Making the Grade* Report

By Kristin Anderson Moore, Hannah Lantos, Rebecca Jones, and Ann Schindler

A major goal in the 2014 report was to triangulate different types of evidence to identify factors that determine academic success and examine the alignment of these factors with the ISS model. The 2014 report used seven different strategies to assess whether integrated student supports were effective. These seven strategies represent complementary approaches that allowed the authors of the initial report to cross-check findings from multiple approaches. The seven strategies were:

- Examination of ISS models in practice
- Synthesis of current educational research
- New empirical analyses of high school graduation and postsecondary attendance
- Assessment of alignment of ISS with child development theories and frameworks
- Review of outcome evaluations
- Assessment of implementation evaluations
- Examination of cost-benefit analyses

We will briefly review findings from each of these seven approaches.

First, the research team examined models for which integrated student supports were provided in practice and determined the five essential components common to ISS models.[4] At the beginning, all ISS models conduct some type of needs assessment. This assessment is often done for individual students and sometimes with or about families, to assess which needs are unmet or uncoordinated. It may also be done at the school and community levels to understand what resources already exist or are not available.

After the needs assessment, schools that use an ISS model develop partnerships within their communities to better utilize existing resources. This happens with other youth-serving organizations and with other people and service providers that may not be youth-focused. The needs that ISS models

[4] All models reviewed in the 2014 report are described in the appendix.

address range from behavioral health support, to housing and food support, to family violence prevention (and more). A variety of local providers can help meet these needs.

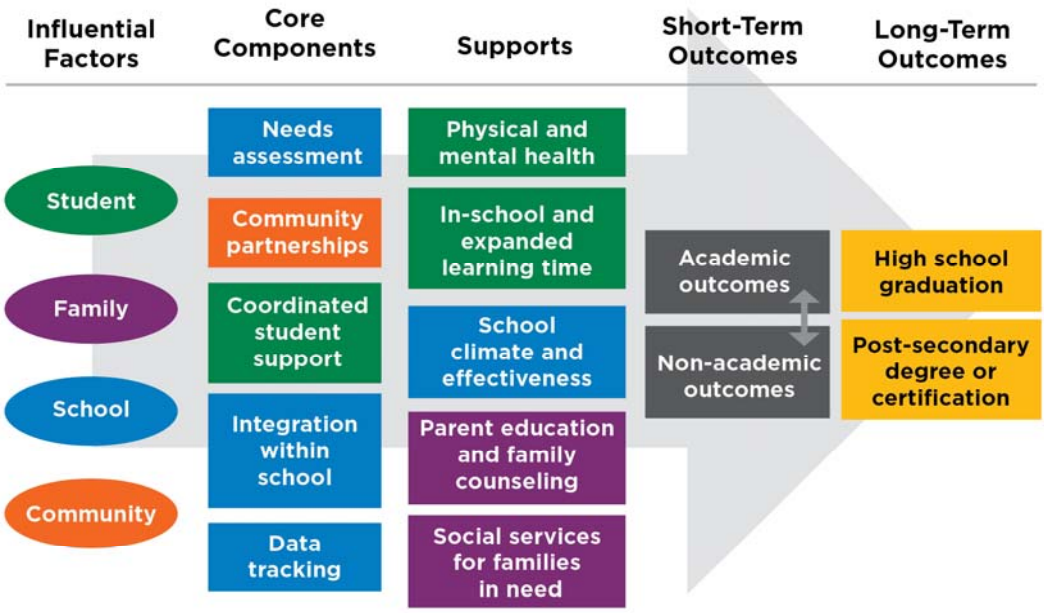
Once existing resources and potentially unmet needs are identified, ISS models focus on coordinating the supports they provide so that children receive necessary supports when needed.

Next, all models focus on creating integration within the school so that teachers, students, counselors, and others who are involved in supporting specific students are aware of what is going on with each child and what needs might remain.

Finally, all ISS models emphasize the collection and use of data to track and monitor students over time so that improvements can be understood and remaining problems can be quickly addressed. These core elements may not unfold in lockstep order, but stakeholders who reviewed model components confirmed that these elements generally represent the core of an ISS model.

These five core components, and how they fit into a broader ISS logic model, are shown below in Figure 2. Within these boxes are a number of different constructs. For instance, both grades and test scores should be included within the academic outcomes box, while the nonacademic outcomes box may include outcomes such as mental health or behavioral outcomes. The tables in the outcomes chapter provide examples of how many different indicators could be included in either of these two boxes. Additionally, there are likely to be implied arrows between boxes here, although we know little about the strength of these relationships. In this report, we have an added an arrow to connect the academic and nonacademic outcomes in the figure below. Additionally, there are likely arrows on the left side of the diagram from student to family to school and to community, and vice versa—as described in the ecological model. Importantly, the arrow in the background focuses on the long-term outcomes that ISS models seek to achieve: increased high school graduation and/or postsecondary degree or certification.

Figure 2. Logic Model of the Five Core Components of ISS Models



Second, the research team synthesized current educational research exploring the factors that affect educational outcomes at four levels: the individual student, the family, the school, and the neighborhood. This very thorough review of that literature can be found in the first report. Child Trends researchers examined studies of educational attainment and achievement to determine which factors have the greatest influence. The factors fell into five categories: individual, family, peer, school, and neighborhood factors.

Promising individual factors included student attendance and engagement, as well as student health and well-being. For example, children who experience high levels of discrimination in school are more likely to believe that they do not belong there. Key family factors contributing to educational achievement and attainment included parental expectations and parenting behaviors. Students whose parents have high expectations for academic achievement are more likely to enroll in postsecondary education. Peer factors can be either negative or positive influences. While researchers have found associations between peer influences and education, more research is needed on this specific relationship to better understand the mechanisms through which peer pressures can be helpful or harmful. School factors that influence educational achievement include the socioeconomic status of the students who predominately attend the school, the quality of student/teacher relationships, school size, and a safe school climate. For example, in schools where students have positive relationships with teachers and administrators, there is evidence of better student behavioral[lxvii] and academic outcomes,[lxviii] while bullying in schools is associated with lowered academic performance[lxix] and increased dropout.[lxx] In the 2014 report, Child Trends concluded that these factors align well with the ISS conceptual model shown in Figure 2. The research is consistent in identifying both non-school and school-based factors as influences on academic achievement.

Third, Child Trends researchers conducted new empirical analyses. One limitation in much of the existing educational literature is that many analyses focus on one main factor and explore only that factor's association with educational attainment. However, the field often fails to examine cases where a variety of factors may have a small (and sometimes negligible) impact; together, these small impacts may add up and have a large effect on educational outcomes. To address this gap in the research literature, Child Trends researchers analyzed data from the National Educational Longitudinal Study (NELS) to identify factors predictive of high school graduation rates and postsecondary enrollment. NELS is a dataset that follows student outcomes of eighth graders from the class of 1988 for twelve years after their eighth-grade graduation. There were two outcome variables of interest: high-school completion and postsecondary enrollment.

From the 7,500 available independent variables, researchers selected 154 as high-potential variables that were also malleable—that is, these variables could be changed by effective programming. A series of logistic regressions were run to determine which variables best correlate with the two outcome variables. Results of the logistic regressions found that many important factors are related, each in a small way, to high school graduation, but that few have large effect sizes. Furthermore, in the analyses run, logistic regressions for smaller sub-groups of black and Latino students found that—despite relatively few differences between black students, Hispanic students, and students overall—there were several important differences. For example, completing homework has a larger impact on the

graduation rates of black students than for students as a whole, suggesting that (for myriad reasons) black students may need more support to get homework completed and turned in on time.

Overall, the research team concluded that these findings align with ISS models because ISS does not focus on just one issue; rather, ISS schools address the unique and multiple needs and concerns for each individual child.

Fourth, Child Trends researchers reviewed the literature on child development. Again, this was one way to triangulate the evidence and explore whether child development theory supported ISS models. As noted earlier, the research team determined that ISS models align well with five well-supported child development theories. Child Trends researchers found resonance between each of these perspectives and the ISS approach. The five theories that researchers reviewed included:

- The whole child perspective acknowledges the importance of multiple domains to children's well-being and development, including physical health, emotional well-being, social development, and academic or cognitive achievement.
- A child-centered focus recognizes that each child is an individual with unique assets and needs and that the child is the focus, not the adult.
- A life course perspective finds that earlier life cycle experiences impact later events in the life cycle.
- The ecological model recognizes that a child's development is influenced by a wide range of factors, including biology, family, peers, neighborhoods, and the larger social and economic context.
- Positive youth development research finds that programs that use supportive intervention strategies that are developmentally appropriate are more effective than didactic programs or negative interventions, or those that try to scare young people away from certain behaviors.

Finally, researchers analyzed outcomes, implementation, and cost-benefit evaluations. Evaluations were only included for programs that existed in more than one state and that had either a randomized control trial (RCT) or rigorous quasi-experimental design (QED) study with a comparison group. Three types of studies were examined:

- Outcomes and impact evaluations
- Implementation findings
- Cost-effectiveness

Four criteria were defined to select studies for inclusion in the review of outcomes evaluations. Specifically, studies were selected if they operated in more than one state, served students from pre-K – 12th grade, and utilized community partners to support students and families, and if there was a rigorously designed (RCT or QED) evaluation. The 2014 review of ISS models that had been rigorously evaluated found promising (albeit inconsistent) results, suggesting that the model has the potential to impact long-term student outcomes and well-being.

Student progress (measured by credit completion, grade retention, high school dropout, and promoting power—a ratio of seniors in a high school relative to the number of freshman four years prior), school attendance (measured by chronic absenteeism, absenteeism, and attendance rate), and academic achievement (measured by reading/ELA achievement, math achievement, and overall GPA) were all studied as outcomes in one or more of the nine identified studies. Several positive impacts were found, including on credit completion, promoting power, grade retention, and high school dropout rate. Evaluations of attendance were promising as well, although somewhat less conclusive: some found impacts on attendance while others did not. Some also found impacts for students in specific age groups while other ages were unaffected.

In general, the QED studies often found more consistently positive results for academic achievement than the RCTs. In these studies, ISS programs were found to affect math achievement and literacy, although the math results were more consistent. Similarly, none of the RCTs found impacts for GPAs, but QEDs that studied GPAs as an outcome did find effects. Reviewers also noted that standardized tests vary across different states, so these are more useful when comparing models within the same state, or among states that use tests that are either the same or more closely aligned to one another.

Based on this extensive and comprehensive review, the 2014 report concluded that the ISS approach was promising but that many questions remained unanswered. Only one study was able to randomize the receipt of ISS at the school level, while all others randomized students *within* a school. This means that some students were randomized into a group that received a higher level of services(tier 2), compared to the rest of the student who received a school-wide base level of services (tier 1). This means that any results found in these RCTs are the impact of additional "tier 2" services, as opposed to being the results of the ISS model per se. However, negative findings were few and far between.

Additionally, few studies explored the effect of ISS on the variables hypothesized to be the intermediate or mediating variables. These are often nonacademic mediators like school engagement, or improvement to the child's situation or well-being following receipt of additional supports (e.g., improved mental or behavioral health or more stable housing for the family once resources are provided). If the ISS models do not actually affect the outcomes they are hypothesized to affect, then the models are not working. However, if the models affect these intermediate variables but do not yet show impacts on the academic outcomes of interest, researchers should explore whether that is because the academic outcomes take longer to be affected or because the ISS theory of change is incorrect. Without measuring and studying these intermediate outcomes, we do not know which is true. Additionally, it is essential to remember what the previous report found—that no single nonacademic outcome will likely be the silver bullet, but that the multiplicity of needs and supports is what will improve overall academic outcomes.

In sum, although there was promising evidence that ISS could improve educational outcomes at the time of the initial report, the review found little definitive information across the studies about what specific outcomes are affected, as many studies used different measures and outcomes.

In addition, there was a lack of definitive information regarding best practices in implementation. The goal of the current update is to identify new studies and explore whether the last three years have produced evidence that might help fill in the blanks.

In this report, we update our findings by analyzing ten more (mostly newer) studies of seven different programs, in addition to the earlier identified studies. We also expand analysis of the findings about factors that might explain some of these differences. This should provide valuable insight to policymakers, principals, school district administrators, and state departments of education as they begin to implement the integrated approaches supported by the new ESSA legislation.

Chapter 3: Methodology Used to Review Outcomes Evaluations

By Hannah Lantos, Kristin Anderson Moore, and Rebecca Jones

Chapter Overview

This chapter briefly reviews the methodologies used in the analyses in this report. Three types of studies are included: outcome evaluations, implementation evaluations, and benefit-costs studies. First, the criteria for inclusion of outcome evaluations is reviewed. Second, the chapter covers the criteria used for implementation studies. While benefit-costs analyses were the third type of study analyzed, only one new benefit-cost analysis was published; a description of its methodology is included in Chapter 6. Finally, we conclude with a description of the qualitative data collection and coding conducted with principals. A detailed review of the methods of the Social Genome Model is included in Appendix 4.

Outcome Evaluations

This update keeps three of the criteria used in the 2014 report to select studies for the review: schools needed to serve students from kindergarten through 12th grade, models needed to utilize community partners, and evaluations needed to employ a random assignment or quasi-experimental approach. However, we dropped the criteria that a program be implemented in multiple states and required only that models be nationally recognizable. This revision allowed us to include evaluations of schools that operate in only one place. We have also included working papers that have undergone peer review.

Three criteria were used to define a rigorous design for both RCTs and QEDs:

1) An experimental design (random assignment design or, for QEDs, a matched comparison group)
2) An intent-to-treat analysis (the evaluations from City Connects are the exception to this criterion, as they use propensity score matching—a "treatment on the treated" approach—but were included because this is a rigorous, quasi-experimental design)
3) No serious problems in terms of confounding (for example, the presence of another education program in the school that cannot be controlled for)

We required QED studies to meet three additional criteria. First, they needed to have low attrition rates, as defined by the Office of Adolescent Health's guidance[lxxi] and explained in more detail in the What Works Clearinghouse.[lxxii] Second, groups needed to establish baseline equivalency indicating that the intervention and comparison groups were similar on key *a priori* identified variables. Third, analyses

needed to statistically control for age (or grade), gender, race/ethnicity, and baseline outcome measures (if those were measured). Nonexperimental studies were excluded.

The outcomes of interest in these studies include:

- GPA
- Grade progression
- Math and ELA grades
- Math and ELA test scores
- Attendance

To identify and select outcome evaluations that met our criteria, we conducted a four-stage process. First, we looked for updated evaluations of the programs included in the 2014 report and reached out to their program staff to ask whether there were any new, updated evaluations.

Second, we conducted a review of the research literature to identify new evaluations (of both previously identified and new programs), in both the education and public health research literatures. This review included both peer-reviewed journals and other studies, such as evaluations funded by the federal government or a foundation that met our methodological criteria. To find this literature, we searched library databases at two major universities (Columbia University's Teachers College and Johns Hopkins University, including ERIC, Ebscohost, JSTOR, Project Muse, Proquest, and Sage), with search phrases such as "integrated student support/s," "community school/s," "wraparound services," "outcome/s evaluation," "implementation evaluation," and "collective impact." We also looked at other organizations' published evaluations in the gray literature (such as studies conducted by AIR, MDRC, ICF, and NORC) to see if they had conducted program evaluations of which we had not previously been aware. As more and more schools begin to incorporate support for nonacademic needs into their mission and programming, it is challenging to identify new programs that are truly ISS models without evaluations that describe them in detail.

Third, we reached out to education experts (specifically those in the ISS field) and other stakeholders to ask about other new evaluations. Some of these stakeholders had been identified in 2014 and participated in our stakeholder roundtable, but we also included new connections, people at newly identified programs, and participants from a roundtable discussion on ISS hosted by the Economic Policy Institute (EPI) in September 2016. Finally, we reviewed the database of the Coalition for Community Schools to see if they had identified any additional studies.

We identified 11 new evaluations of seven programs which, combined with the 10 evaluations of three different ISS models from the first report, add substantially to the existing knowledge base and information on program variety. The programs are listed and their evaluation methods are described in Table 1, below.

Table 1. All Studies Included in the Review of Outcome Evaluations

Study	Notes
2017 Report	
Randomized Control Trials	
Communities in Schools in Chicago, IL[lxxiii]	Randomized at the school level
Communities in Schools, National Study[lxxiv]	Randomized at the student level
Diplomas Now, National Study[lxxv]	Randomized at the school level
Harlem Children's Zone Promise Academy (HCZ PA) - Middle School Outcomes, New York, NY[lxxvi]	Randomized at the student level
Harlem Children's Zone Promise Academy (HCZ PA)- high school and postsecondary outcomes, New York, NY (a follow-up study)[lxxvii]	Randomized at the student level
Quasi-experimental Designs	
City Connects, Boston, MA[lxxviii]	Interrupted time series
City Connects, Boston, MA[5, lxxix]	Propensity score matching
City Year, National Study[lxxx]	Propensity score matching
Communities in Schools, Texas and North Carolina[lxxxi]	Comparative interrupted time series (CITS)
Say Yes to Education, National Study[lxxxii]	Propensity score matching
Talent Development, National Study[lxxxiii]	Comparative interrupted time series (CITS)
2014 Report	
Randomized Control Trials	
Comer School Development Program, Prince George's County, MD[lxxxiv]	Randomized at the school level
Communities in Schools (CIS) in Austin, TX[lxxxv]	Randomized at the student level
Communities in Schools (CIS) in Jacksonville, FL[lxxxvi]	Randomized at the student level

[5] This paper is the published version of the previous working paper that was included in the 2014 report.

Study	Notes
Communities in Schools (CIS) in Wichita, KS[lxxxvii]	Randomized at the student level
Quasi-experimental Designs	
3 bi-yearly evaluations of City Connects (CCNX) in Boston, MA[lxxxviii]	Comparison schools
Comer School Development Program, Chicago, IL[lxxxix]	This was an RCT, but due to high attrition the authors found control schools to include and the study became a QED

Coding of studies

Studies were coded (in the case of the newly identified studies) and recoded (in the case of those previously identified) using NVivo 10. Two Child Trends researchers independently coded each study, identifying results for each of eight key outcomes:

- English language arts (ELA) grades
- ELA test scores
- Math grades
- Math test scores
- Attendance
- Grades/GPA
- Graduation
- Promoting power

Each outcome was coded as having improved, declined, or stayed the same, as measured by statistical tests in each article or report; the time horizon of the outcomes was also coded (one, two, three years after baseline). Codes were compared and disagreements about codes were discussed by the coders, who then concluded jointly on the final assigned codes. Researchers did not create new codes, but used the same ones from the 2014 report. This cross-checking identified few disagreements about which codes to use. The most common disagreement by coders was that one coder had identified subgroup analyses and another had missed them. The straightforward approach left little room for disagreement, as the authors of each report usually reported an improvement, decline, or neither; and there was little room for subjective misunderstanding. Statistical significance at the 5 percent level was used to define an impact or effect, although a few studies that used a cutoff at the 10 percent level are included in the following chapters with appropriate footnotes.

Implementation Evaluations

For the implementation studies, we included results from two different types of analyses. First, results from quantitative implementation studies are included for studies among the outcomes evaluations.

Second, we conducted a series of interviews with principals and program developers or leaders from across the country identified by City Connects, Communities in Schools, the DePaul University Center for Urban Education in Chicago, the New York City Office of Community Schools, and the Washington State Office of the Superintendent for Public Instruction.

Coding of interviews with principals

Two researchers from Child Trends participated in all interviews conducted with principals. We took notes and recorded each conversation so that we could go back to review. Each conversation was structured around the five core components identified in the 2014 report and shown previously in Figures 1 and 2. These conversations were focused on implementation of the ISS model, and researchers asked principals what each of the five components looked like in their school, whether the five core components captured everything they did at their school (or if something was missing), and which issues presented the biggest challenges to this work. The work was determined to be exempt by the Child Trends Institutional Review Board (IRB), as it was about school processes and there were no risks (such as job loss) in speaking with the interviewers. [xc]

At the end of each interview, the two researchers identified themes and whether any new themes were raised. After 11 interviews were completed, we collated themes repeated throughout interviews and modified the interview protocol slightly to explore specific themes that were appearing, and to ask other principals whether those themes resonated with them. For instance, we began to specifically ask about two levels of needs assessments: at the individual student level and at the school level. After 22 interviews, we were no longer hearing new themes and concluded that we had reached saturation. This was a modified grounded theory approach in which we coded themes immediately at the end of each interview in real time and added new topics to the list as we proceeded.

Five core components

In their interviews, we explicitly asked principals whether they implemented the five core components that Child Trends identified as essential for ISS models. However, for the rest of the interventions included in this study, we had to rely on information in the outcome or implementation evaluations about each program to determine if the five core components defined their different models. We are unable to say with confidence the degree to which the evaluated programs have considered and/or implemented these five core components. We do encourage programs, schools, and principals to think about these steps, although we do not expect all reports to be structured explicitly around them.

Conclusion

This report builds on the methodology of the 2014 report by including more studies and delving more deeply into implementation issues—what these models look like and what makes them successful from a principal's perspective. These models have rapidly expanded over the last three years and this report intends to give a snapshot of the state of ISS in the United States today and what remains unknown.

Chapter 4: Outcomes Evaluations

By Hannah Lantos, Rebecca Jones, Kristin Anderson Moore, Jon Belford, Vanessa Sacks, and Ann Schindler

Chapter Overview

In the 2014 report, Child Trends wrote: "There is emerging evidence, especially from quasi-experimental studies, that [integrated student supports] can contribute to student academic progress as measured by decreases in grade retention and dropout, and increases in attendance, math achievement and overall GPA. Findings for reading and ELA achievement are mixed."

The conclusions in this updated 2017 report echo those in the earlier review: that the evidence in support of integrated student supports is promising but not conclusive. In this chapter, our caution stems primarily from the large number of null findings in the evaluation studies and a need to better understand what drives the positive findings across some models so that they can be replicated. Interestingly, for the present review, results from the RCTs and QEDs were similar. Previously, the QED studies found more consistently positive results for academic achievement than the RCTs. This time, outcomes from both RCTs and QEDs were more consistent for literacy, math, attendance, GPA (or grades), and graduation, with both null and positive results across the board.

Overall, most results are either positive or null (nonsignificant) across most of the outcomes in these evaluations. In one instance, a Communities in Schools (CIS) treatment group did worse than their peers in the control group on an academic outcome, but both groups improved over time, with the comparison group improving more rapidly. Additionally, the focus of the CIS program for this age group was more behavioral than academic, suggesting that improvements in academic outcomes—even if slower relative to the comparison school—were impressive. There was also one instance of standardized test scores being lower for City Connects participants, but negative outcomes were clearly quite rare. Additionally, as before, the outcomes that are measured vary. This lack of consistency makes it a challenge to compare results across studies. For example, some researchers examine grades, others look at test scores, while others focus on attendance or behavioral outcomes; this variation makes it harder to say that the evidence is very strong for any single outcome.

Methods

As noted in the last chapter, we used largely the same methodology for accepting studies as the 2014 report. That is, we required the outcome evaluations to be experimental—either a random assignment

design or a quasi-experimental design with control variables and a comparison group. One difference warrants mention: in the 2014 report, evaluations were only included if they were from programs that existed in multiple geographic locations. For this review, we removed that criteria because highlighting some findings from rigorous, evaluations of smaller programs seemed useful.

Rigorous Evaluations Summarized for Review

In total, 21 studies are included in this report; these evaluations cover eight different programs. The studies from both 2014 and 2017 are included. This allows for comparison over time, and a full understanding of how the evidence has built over time and what research questions remain to be answered given all the evidence so far. The 2014 report included four RCTs from two different programs and five QEDs from two different programs (for a total of nine studies from three programs). In 2014, although CIS had both an RCT and a QED, only the RCT was included in the outcomes analysis, while the latter was included in the analyses of implementation. For 2017, we added five RCTs from three different programs and six QEDs from five different programs. This time around, CIS has both RCTs and QEDs included in the list of evaluations. Also, the evaluation of City Connects included in the 2017 analyses was available in 2014 as a working paper, and has since been published under peer review. The peer reviewed version is included here while the working paper is not. In total, there are seven different programmatic models included here for their rigorous evaluation. Table 2 shows which studies were RCTs or QEDs and which academic outcomes were included in each evaluation.

Table 2. List of Studies Noting Academic Outcomes that Were Examined in Each Study

Studies that report school-level outcomes are marked with a "1" while those reporting student-level outcomes are marked with a "2"

	Program	Year, Site, or Outcomes	Math Grades	Math Test Scores	ELA Grades	ELA Test Scores	Attendance	Grad/ Promoting Power	GPA/ Grades	Grade/ Credits Completed
RCTs 2017	CIS	Chicago		2		2	2			
	CIS	Year 2 impact findings					2	2	2	2
	Diplomas Now						1		1	
	HCZ PA	High school students		2		2				
	HCZ PA	Middle school students		2		2	2			
RCTs 2014	CIS	Austin		2		2	2	2	2	2
	CIS	Jacksonville		2		2	2		2	2
	CIS	Wichita		2		2	2		2	2
	Comer	Comer, Prince		1			1		1	

	Program	Year, Site, or Outcomes	Math Grades	Math Test Scores	ELA Grades	ELA Test Scores	Attendance	Grad/ Promoting Power	GPA/ Grades	Grade/ Credits Completed
		George's County, MD								
QEDs 2017	CIS	National, 2017 (TX and NC)		1		1	1	1		
	City Connects	Dearing et al., 2016		2		2				
	City Connects	Walsh 2014	2	2	2	2			2	
	City Year			1		1				
	Talent Develop-ment			1		1	1			1
	Say Yes			2		2	2			
QEDs 2014	City Connects	Summary Report '08	2	2	2	2				

	Program	Year, Site, or Outcomes	Math Grades	Math Test Scores	ELA Grades	ELA Test Scores	Attendance	Grad/ Promoting Power	GPA/ Grades	Grade/ Credits Completed
	City Connects	Annual Report '10	2	2	2	2				
	City Connects	Progress Report '12	2		2			2	2	
	Comer	Comer Chicago		1 & 2		1 & 2				

There are two important patterns to note in this table. First, as mentioned above, studies examine a wide variety of academic outcomes and tend to use different measures even when they assess similar constructs. This makes it challenging to conclude systematically whether ISS is effective and how to interpret contrasting findings.

Second, course grades are not studied as an outcome by many of the programs; test score outcomes are more common across the studies. This presents both a strength and a potential weakness. The tendency is promising in that it facilitates comparisons across models or school sites; state test scores may be more comparable across schools, programs, and even classrooms than grades. However, if (as suggested above) we expect to see improvements in grades before we see test score improvements *because* grades reflect better behavior, effort, and engagement—which may be rewarded by teachers—then measuring only test scores may not capture some of the first changes in student behavior and performance. If so, we may miss one of the first pieces of evidence of impact in the studies that do not include grades or promoting power. It will be important to clarify this more explicitly in conceptual models moving forward. However, if we see grade improvement and no testing improvement, even over time, we might conclude that these models work through relatively subjective measures of learning.

Finally, this table only reviews the academic outcomes in each study. At the end of this chapter, we will also discuss the nonacademic outcomes included in each evaluation. Few studies examine intermediate outcomes with nuance. Having measures of mediators would enable researchers to assess whether programs are working as theorized. Again, the conceptual model assumes that providing resources will improve several nonacademic factors, which will, in turn, improve academic outcomes. However, understanding whether the models work depends on being able to test those mediating factors. Unfortunately, a common set of well-measured mediators is typically not assessed, making it hard to build a clear theoretical story in the literature. This represents a critical task for future studies.

Results

Given the varied and complex academic outcomes used across evaluations, summary tables are shown below that focus on outcomes—noting whether they are positive, negative, or null (not statistically significant)—by school level (elementary, middle, or high school). Detailed tables with notes about the specifics—especially helpful when, for example, one study found both positive and null effects—are included in Appendix 2. The tables in the appendices also highlight what, if any, sub-analyses were conducted, as well as their results.

The following tables are organized similarly. Each of the four tables presents findings for two of the eight academic outcomes. The top rows for each show results from RCTs, and the bottom rows show results from QEDs. Additionally, because two sets of outcomes are shown per table, the first outcome is in pink and the second is in blue. For each outcome, results are shown for elementary, middle, and high school. If a positive effect was found, this is a noted with a "+" and if a negative effect was found, this is noted with a "-" (if there was no effect, the word "null" is written). Some studies found both positive and null effects by age or for different tests; if this happened, both are noted. If a cell is grayed out, that

outcome was not included in that evaluation. A table with more detailed notes on each outcome is included in Appendix 2.

Findings for English and Language Arts

Table 3. English Language Arts Grades and Test Scores

	Program	Year, Site, or Outcomes	RCT or QED?	Literacy/Grades +, - , and null effects Full Sample Analysis			Literacy Test Scores +, - , and null effects Full Sample Analysis		
				Elem.	Middle	High	Elem.	Middle	High
RCTs 2017	Communities in Schools	Chicago	RCT				+	+	
	Communities in Schools	Year 2 impact findings	RCT						
	Diplomas Now		RCT		+	null			
	Harlem Children's Zone's PA	Dobbie and Fryer (high school students)	RCT						null
	Harlem Children's Zone's PA	Dobbie and Fryer (middle school students)	RCT					+ and null	
RCTs 2014	Communities in Schools	Austin	RCT						null
	Communities in Schools	Jacksonville	RCT					null	
	Communities in Schools	Wichita	RCT						null
	Comer	Comer Prince George's County, MD	RCT						

QEDs 2017	Communities in Schools	National, 2017 (Texas and NC)	QED				null	- and null	+
	City Connects	Dearing et al., 2016	QED				+ and null		
	City Connects	Walsh, 2014	QED	null	null		null	+ and null	
	City Year		QED				+ and null	+ and null	+ and null
	Talent Development		QED						+
	Say Yes		QED				null		
QEDs 2014	City Connects	Summary Report 2008–2009	QED	+ and null			+ and - and null	+ and null	
	City Connects	Annual Report 2010	QED	+			null	+	
	City Connects	Progress Report 2012	QED		+ and null				
	Comer	Comer Chicago	QED						+

Note: Designation of a direction or null finding is based on significance tests reported in each of the studies. Detailed, disaggregated results from each study for all outcomes are included in Appendix 2. Statistical significance is .05 or less.

English and language arts (ELA) grades. Findings depicted in pink in the first three columns of Table 3 depict results for studies that included ELA grades. As noted, few studies included grades, and all were from the City Connects Program. Only elementary and middle school students were included in these studies, as City Connects currently runs programs only at these levels. Results for grades were mixed. Specifically, some studies found a positive effect on ELA grades while the more recent studies found a null effect. No negative effects were found.

ELA test scores. As shown in the three blue columns on the right side of Table 3, more evaluations included ELA test scores as an outcome variable of interest. One study, the Communities in Schools QED evaluation, found negative and null impacts at the middle school level. However, all other evaluations (including other CIS evaluations) found positive or null effects at the elementary, middle, and high school levels. The CIS QED is one of the studies that is methodologically complex to interpret.

Specifically, this study compares receipt of tier 2 services relative to receipt of just tier 1 services. In so doing, it answers the question of whether additional tier 2 services improve outcomes more than simply receiving tier 1 services. Therefore, it is only able to assess the benefit of the additional services—rather than the benefit of being in a school with an ISS model relative to a school without one. Although the negative finding is concerning for middle school students, it does not lead us to make strong conclusions about ISS overall.

One other pattern is worth noting. As studies have accumulated, there are fewer differences between the results seen in RCTs and those seen in QEDs. In 2014, the results from QEDs were more positive than those from RCTs; however, this pattern is less starkly true now. In 2017, all QED evaluations included the ELA testing measure in their studies, while three of five RCTs did so. However, both QEDs and RCTs show a mixed pattern across all three age ranges, with results that are mostly null or positive (with one negative impact).

Findings for Math

Table 4. Math Grades and Math Scores

	Program	Year, Site, or Outcomes	RCT or QED?	Mathematics Grades +, - , and null effects Full Sample Analysis			Mathematics Test Scores +, - , and null effects Full Sample Analysis		
				Elem	Middle	High	Elem	Middle	High
RCTs 2017	Communities in Schools	Chicago	RCT				+	+	
	Communities in Schools	Year 2 impact findings	RCT						
	Diplomas Now		RCT		null	null			
	Harlem Children's Zone's PA	Dobbie and Fryer (middle school students)	RCT					+	
	Harlem Children's Zone's PA	Dobbie and Fryer (high school students)	RCT						+
RCTs 2014	Communities in Schools	Austin	RCT						null

	Communities in Schools	Jacksonville	RCT					null	
	Communities in Schools	Wichita	RCT						+
	Comer	Comer Prince George's County, MD	RCT					null	
QEDs 2017	Communities in Schools	National, 2017 (Texas and NC)	QED				null	null	null
	City Connects	Dearing et al., 2016	QED				+		
	City Connects	Walsh, 2014	QED	+ and null	null			+	
	City Year		QED				+ and null	+ and null	+ and null
	Talent Development		QED						+
	Say Yes		QED						
QEDs 2014	City Connects	Summary Report 2008–2009	QED	+			+ and null	+	
	City Connects	Annual Report 2010	QED	+			null	+	
	City Connects	Progress Report 2012	QED		+				
	Comer	Comer Chicago	QED						+

Note: Designation of a direction or null finding is based on significance tests reported in each of the studies. Detailed, disaggregated results from each study for all outcomes are included in Appendix 2. Statistical significance is .05 or less.

Math grades. Table 4 shows that few studies included grades for math. Again, all of these studies were from the City Connects Program, shown in pink in the first three columns of the table. The results are

more consistently positive than for ELA grades, as all studies found a positive impact on math grades, although some (depending on the age range) also found null results. For elementary school students, there were positive results in every study, while there was one positive and one null result for middle school students. As noted, City Connects does not have results from high school studies, as all their schools are elementary or middle schools.

Math test scores. In terms of math test scores, there are once again similarities to ELA results, in that more studies included math test scores in their outcomes than math grades. For this outcome, every evaluation found either positive or null effects at all three schooling levels. Overall, QEDs were, again, no more likely than RCTs to find positive impacts on math test scores. This is primarily driven by the fact that the three new RCT studies that include math test scores (from CIS Chicago and Harlem Children's Zone's Promise Academy) all found positive impacts across the elementary, middle, and high school levels. Additionally, four of five QED studies that include math test scores found a combination of null and positive effects.

Findings for Grades and GPA

Table 5. GPA, Grades, Grade Completion, Credit Accumulation

	Program	**Year, Site, or Outcomes**	**RCT or QED?**	**GPA/grades +, - , and null effects Full Sample Analysis**			**Grade Completion/ Credit Accumulation +, -, and null effects Full Sample Analysis**		
				Elem	**Middle**	**High**	**Elem**	**Middle**	**High**
RCTs 2017	Communities in Schools	Chicago	RCT						
	Communities in Schools	Year 2 impact findings	RCT			null			null
	Diplomas Now		RCT		null	null			
	Harlem Children's Zone PA	Dobbie and Fryer (middle school students)	RCT						
	Harlem Children's Zone PA	Dobbie and Fryer (high school students)	RCT						
RCTs 2014	Communities in Schools	Austin	RCT			null			null

	Communities in Schools	Jacksonville	RCT		null			null	
	Communities in Schools	Wichita	RCT			null			+
	Comer	Comer Prince George's County, MD	RCT		null				
QEDs 2017	Communities in Schools	National, 2017 (Texas and NC)	QED						
	City Connects	Dearing et al., 2016	QED						
	City Connects	Walsh, 2014	QED		null				
	City Year		QED						
	Talent Development		QED						+ and null
	Say Yes		QED						
QEDs 2014	City Connects	Summary Report 2008–2009	QED						
	City Connects	Annual Report 2010	QED						
	City Connects	Progress Report 2012	QED		+				
	Comer	Comer Chicago	QED						

Note: Designation of a direction or null finding is based on significance tests reported in each of the studies. Detailed, disaggregated results from each study for all outcomes are included in Appendix 2. Statistical significance is .05 or less.

GPA or grades. Table 5 shows, in pink, the results for overall grades or GPA. Only three new studies (two RCTs and one QED) included these outcomes. With one exception, City Connects, all of these studies found null effects, suggesting that improving grades overall or increasing students' GPAs is a

complicated, difficult process. The City Connects study found positive effects for middle school students in terms of GPA or overall grades.

Credit completion or accumulation. The three columns to the right in Table 5 show, in blue, the studies that included outcomes of grade completion or credit accumulation. Few studies examined this outcome. The new Communities in Schools RCT did not find an impact of receiving tier 2 services on these outcomes, while some of the older CIS studies and the Talent Development study found mixed (positive and null) effects.

Neither set of outcomes was studied at the elementary school level. Credit accumulation and GPAs are more appropriate outcomes to study for older children, as younger children do not have to collect credits to graduate and do not have GPAs reported. However, it might be useful to study grades at the elementary school level to assess whether students are performing adequately over time.

Findings for School Attendance and for Graduation, Dropout or Promoting Power

Table 6. Attendance, Graduation, Dropout, Promoting Power

	Program	Year, Site, or Outcomes	RCT or QED?	Attendance +, - , and null effects Full Sample Analysis			HS Graduation, Dropout or Promoting Power +, - , and null effects Full Sample Analysis		
				Elem	Middle	High	Elem	Middle	High
RCTs 2017	Communities in Schools	Chicago	RCT	null	null				
	Communities in Schools	Year 2 impact findings	RCT		null	null			
	Diplomas Now		RCT		+	null			
	Harlem Children's Zone PA	Dobbie and Fryer (high school students)	RCT						
	Harlem Children's Zone PA	Dobbie and Fryer (middle school students)	RCT		null				

	Program	Year, Site, or Outcomes	RCT or QED?	Attendance +, - , and null effects Full Sample Analysis			HS Graduation, Dropout or Promoting Power +, - , and null effects Full Sample Analysis		
				Elem	Middle	High	Elem	Middle	High
RCTs 2014	Communities in Schools	Austin	RCT						null
	Communities in Schools	Jacksonville	RCT		null				
	Communities in Schools	Wichita	RCT			+			
	Comer	Comer Prince George's County, MD	RCT		null				
QEDs 2017	Communities in Schools	National, 2017 (Texas and NC)	QED	+	null	null			+ and null
	City Connects	Dearing et al., 2016	QED						
	City Connects	Walsh, 2014	QED						
	City Year		QED						
	Talent Development		QED			+			
	Say Yes		QED	null					
QEDs 2014	City Connects	Summary Report 2008–2009	QED						
	City Connects	Annual Report 2010	QED						

	Program	Year, Site, or Outcomes	RCT or QED?	Attendance +, - , and null effects Full Sample Analysis			HS Graduation, Dropout or Promoting Power +, - , and null effects Full Sample Analysis		
				Elem	Middle	High	Elem	Middle	High
	City Connects	Progress Report 2012	QED						+
	Comer	Comer Chicago	QED						

Note: Designation of a direction or null finding is based on significance tests reported in each of the studies. Detailed, disaggregated results from each study for all outcomes are included in Appendix 2. Statistical significance is .05 or less.

Attendance. Table 6 displays results for attendance in pink. This was by far the most commonly included outcome across studies, which makes sense because failing attendance is an early indicator of problems in school, at home, or both.[xci, xcii] Additionally, among states that have filed new plans under ESSA, attendance is the most common nonacademic outcome included in the plans. At the elementary school level, poor attendance can signify issues at home, with the family, or for the parents, as the latter ensure that young students regularly attend school and can set future standards for attendance.[xciii] In the higher grades, low attendance predicts lower performance and lower graduation rates, setting students on a long-term trajectory of lower income.[xciv] Tracking attendance is important to be able to address any issues as soon as possible, which is why many schools assess attendance in their "early warning" systems.

The findings for attendance are mixed. All RCTs included attendance, except for the middle school evaluation of the Harlem Children's Zone's Promise Academy. However, a statistically significant relationship between ISS and attendance was found in only two CIS evaluations: Austin, Texas and Wichita, Kansas. No other RCTs found any impact (positive or negative) on attendance. Communities in Schools, Talent Development, and Say Yes were the QEDs that included attendance. Interestingly, all these are more recent studies, while the older QEDs did not have attendance as an outcome measure—suggesting an increased understanding of the importance of attendance or a new emphasis on its measurement. These QED studies also had mixed results. Elementary students in CIS generally experienced positive effects, while Talent Development had a positive effect for high school students. Say Yes had null results at the elementary school level.

Graduation/dropout. The final outcome seen in these tables is for graduation or dropout. Only two programs included this as an outcome in their evaluations. City Connects and the QED of CIS found mostly positive associations with graduation.

Nonacademic Outcomes

It is important to study nonacademic outcomes for two reasons. First, ISS is based on the premise that improving nonacademic factors will ultimately lead to better academic outcomes. These may be referred to as the intermediate or intervening factors, or the mediator variables. They are important to assess and analyze because they can explain how the intervention works or does not work. For example, when studies do not measure these factors, it is difficult to explain null or small impacts on academic outcomes. Perhaps the ISS model is incorrect—or perhaps it *is* correct, but there has not been enough time to see change. In addition, including these measures allows both researchers and educators to assess whether ISS services are producing changes in intermediate nonacademic outcomes. Therefore, it is essential that future evaluations include nonacademic, mediating variables.

Unfortunately, including nonacademic outcomes in outcome evaluations does not yet happen frequently or consistently. Few studies included in this review had comprehensive measures of nonacademic outcomes, although many studies are starting to include some measures. Potential variables to include as nonacademic outcomes might be behavioral health improvements, successful responses to food instability in the home, or improved after-school opportunities for children to learn and explore. There is yet to be a clear consensus of *which* nonacademic outcomes are most important, but this depends heavily on the model a school(s) chooses to implement. Essentially, any nonacademic outcomes that each model attempts to improve should be measured in any evaluation of that model. The next step is to assess the expected linkage between these and the academic outcomes in mediation models. Importantly, none of the programs seemed to be associated with any negative effects on nonacademic outcomes.

In Tables 7, 8, and 9, results are shown for different types of broad nonacademic outcomes within larger groups—such as behavior, health, school climate, etc. We grouped outcomes into broad categories to explore patterns, recognizing that many studies did not include the exact same measures for these outcomes. Specifically, all school attachment outcomes are grouped together, as are all behavioral outcomes, and so on. Detailed results by study are included in Appendix 3.

Studies in these tables are also grouped according to the level at which the outcomes were measured. Outcomes measured at the individual level are shown in Table 7, in yellow; those at the family level are shown in Table 8, in blue; and those at the school level are shown in Table 9, in green.

In addition, studies are grouped by method of evaluation. RCTs are shown in the first column and QEDs in the second. Among RCTs, seven (of nine) studies included a nonacademic outcome; among QEDs, five (of nine) did the same. However, each study did not include every category, so the number of studies in each category is usually smaller.

Findings for Nonacademic Outcomes

Table 7 depicts the outcomes included in the evaluations. The most nonacademic outcomes by far were studied at the individual level. The two outcomes with the most consistent positive associations were student health/well-being and student-teacher/staff relationships. For health outcomes, one of two

RCTs and two of two QEDs found that ISS was associated with improvements in health and well-being. For students' relationships with staff and teachers, two of three RCTs and the sole QED that included this outcome found positive associations as well. Other outcomes were found to be associated with participation in an ISS program for just one or two studies (of a larger number of total studies).

Although the student health and student-teacher/staff relationship outcomes had the strongest evidence of improvement, there is reason to be cautious with this interpretation. First, these tentative patterns are based on a small number of studies (one to three), limiting our ability to draw strong conclusions. Additionally, there were often positive impacts in only one or two studies, even for outcomes that were included in more studies. Therefore, it is possible that the patterns will change as the number of rigorous evaluations accumulates.

Finally, behavior and social and emotional development were highlighted by many principals in their interviews with Child Trends researchers (see Chapter 5), as areas in which ISS contributes value. Many felt that their schools struggled the most, and that they would need the most support, in these areas—suggesting that these constructs should be prioritized in future evaluations.

Table 7. Nonacademic Outcomes at the Individual Level

	RCTs (7 of 9 include nonacademic outcomes)	**QEDs** (5 of 9 include nonacademic outcomes)	**Which studies found significant results?**	**Which studies found no significant results?**
Outcomes for Students as Individuals				
School Attachment/ Engagement	1 study of 6 found positive impacts on school attachment.	1 study (of 1) found positive impacts on school attachment.	**RCT**s: Communities in Schools: Year 2 Impact Findings found positive impact on school attachment. **QED**s: Comer in Chicago found positive impacts on school attachment.	**RCT**s: Diplomas Now; Communities in Schools in Jacksonville, Austin, and Wichita; and Comer in Prince George's County found no impact on school attachment. Harlem Children's Zone's Promise Academy did not look at this outcome at all. **QEDs**: None.[1]
Behavior	2 studies of 7 found	2 of 4 studies found	**RCT**s: Harlem Children's Zone's Promise Academy	**RCT**s: Communities in Schools (Year 2); Diplomas Now; and

[1] School attachment was measured in different ways. In the CIS report that found an impact, it was measured by engagement in school.

	RCTs (7 of 9 include nonacademic outcomes)	**QEDs** (5 of 9 include nonacademic outcomes)	**Which studies found significant results?**	**Which studies found no significant results?**
	significant declines in behavioral issues.	significant declines in behavioral issues.	found reductions in risky behavior (pregnancy, drug use, etc.) and Comer in Prince George's County found reductions in behavioral problems. **QED**s: City Connects 2008–2009 and City Connects 2010 found reductions in behavioral problems.	CIS in Austin, Jacksonville, and Wichita found no impact on behavior problems. **QED**s: Say Yes had marginally significant decreases in suspensions in the second year, but otherwise found no significant differences. Comer students had reported lower behavior scores; however, there is no indication that this gap widened over time, just that scores started out and stayed lower.
Socio-emotional Development	2 studies of 7 found relationships between socio-emotional development and ISS models. One relationship is positive and one is negative.	2 of 2 studies found positive relationships between socio-emotional development and ISS models.	**RCT**s: CIS Year 2 study found improvements in educational attitudes, while Harlem Children's Zone's Promise Academy found lower levels of grit. **QED**s: City Connects 2008–2009 found increases in effort and work ethic. City Connects 2010 had better work habit scores in grades 3 and 5 and better work ethic scores in grades 3, 4, and 5.	**RCT**s: Diplomas Now has no significant impacts on self-perceptions; CIS Austin has no significant differences in terms of personal responsibility, self-worth, or future aspirations; in CIS Jacksonville, results for personal responsibility are marginally significant; there are no significant differences for CIS Wichita in terms of personal responsibility, self-worth, or future aspirations; in Comer Prince George's County there were no significant differences in self-efficacy, satisfaction with self, or anger control between the Comer and non-Comer

	RCTs (7 of 9 include nonacademic outcomes)	**QEDs** (5 of 9 include nonacademic outcomes)	**Which studies found significant results?**	**Which studies found no significant results?**
				schools **QED**s: None.
Student Health and Safety	1 study of 2 found impacts on student health and well-being.	2 of 2 studies found impacts on student health and well-being (both positive).	**RCT**s: Harlem Children's Zone's Promise Academy found reductions in teen pregnancy but no impact on self-reported health. **QEDs:** City Connects 2008–2009 found that students in grades 4 and 5 scored higher on tests about unhealthy nutrition and overall well-being. In City Connects 2010, 2nd and 3rd graders learned more about nutrition, and 4th and 5th graders were less likely to engage in unhealthy eating behaviors.	**RCT**s: Comer in Prince George's County found no impacts on student health and well-being. CIS Year 2; CIS Austin, Jacksonville, and Wichita; and Diplomas Now did not include health outcomes. **QED**s: None

Familial outcomes were the least likely to either be studied or have impacts, but it is unclear why this was the case. However, future studies should incorporate family outcomes because they represent a key component of ISS models.

Table 8. Nonacademic Outcomes at the Family Level

	RCTs (7 of 9 include nonacademic outcomes)	**QEDs** (5 of 9 include nonacademic outcomes)	**Which studies found significant results?**	**Which studies found no significant results?**
Student Outcomes Measured in Families				
Academic Support at Home	No studies in 4 found impacts on academic supports at home.	The sole study found no impact on academic supports at home.	**RCT**s: No studies found a significant impact on academic supports at home. **QED**s: No studies found a significant impact on academic supports at home.	**RCT**s: CIS Austin, Jacksonville, and Wichita; and Comer Prince George's County found null effects on this outcome. CIS Year 2, Harlem Children's Zone's Promise Academy, and Diplomas Now did not include this outcome in their studies. **QED**s: Comer Chicago found no impacts on academic support at home.[2]
Parenting Techniques	The sole study found no impact on parenting techniques.	No studies included the use of positive parenting techniques in their evaluations.	**RCT**s: No studies found an impact on this outcome. **QEDs:** No QEDs included this measure.	**RCT**s: Comer in Prince George's County found no impacts on positive parenting techniques. **QEDs:** No QEDs included this measure.

[2] Comer Chicago reported a lower parent valuation of education. This finding did not widen, and parent valuation stayed consistently low. This outcome is null.

	RCTs (7 of 9 include nonacademic outcomes)	**QEDs** (5 of 9 include nonacademic outcomes)	**Which studies found significant results?**	**Which studies found no significant results?**
Parent-child Relationships	1 of 5 studies found impacts on positive parent-child relationships.	No QED studies included the use of positive parent-child relationships in their evaluations.	**RCTs**: CIS Year 2 found that case-managed students reported more caring relationships at home. **QEDs:** No QEDs included this measure.	**RCT**s: Diplomas Now and Communities in Schools in Austin, Jacksonville, and Wichita reported null findings with regard to the presence of positive parent-child relationships. **QED**s: No QEDs included this measure.

Finally, some evaluations included nonacademic outcomes at the school level. These were slightly more common than family-level variables and fell into two buckets: school climate and relationships between students and teachers or staff.

Table 9. Nonacademic Outcomes at the School Level

	RCTs (7 of 9 include nonacademic outcomes)	**QEDs** (5 of 9 include nonacademic outcomes)	**Which studies found significant results?**	**Which studies found no significant results?**
Outcomes Within the School Environment				
School Climate	1 study of 3 found positive impacts on school climate.	The sole study found positive impacts on school climate.	**RCT**s: Communities in Schools: Year 2 Impact Findings found positive impact on school climate. **QED**s: Comer Chicago found both positive and null findings concerning positive school climate.	**RCT**s: Marginally significant differences in positive school climate for Diplomas Now. No clear effects on school climate for Comer Prince

	RCTs (7 of 9 include nonacademic outcomes)	**QEDs** (5 of 9 include nonacademic outcomes)	**Which studies found significant results?**	**Which studies found no significant results?**
			Positive school climate increased for both students and staff, but was consistently lower in Comer schools than in non-Comer schools.	George's County. **QED**s: Not applicable.[3]
Student-Teacher and/or Staff Relationships	2 of 3 studies found positive impacts on student-teacher and/or staff relationships.	The sole study found positive effects on student-teacher and/or staff relationships.	**RCT**s: CIS Year 2 found tier 2 case-managed students reporting more caring relationships at school. Students enrolled in Diplomas Now were more likely to report having positive relationships with non-teacher staff members. There were no differences in the Comer study in these relationships for Comer participants compared to non-Comer students (+ and null). **QED**s: Several indicators about relationships between students and teachers had significant positive associations with participation in the Comer program, at both	**RCT**s: None of the variables about student-staff relationships were significantly higher for students in Comer schools. **QED**s: Not applicable.

[3] City Connects did not provide statistical differences, but present qualitative findings.

	RCTs (7 of 9 include nonacademic outcomes)	QEDs (5 of 9 include nonacademic outcomes)	Which studies found significant results?	Which studies found no significant results?
			the individual and the school level.	

Simulating the Long-Term Impacts of ISS Programs

Long-term studies of childhood and adolescent interventions can be both costly and time-consuming, and researchers must sometimes wait decades to identify adult outcomes. There is currently a lack of evidence on the long-term impacts of ISS programs, which is needed to understand the full benefits of these programs and assist policymakers in deciding whether to implement ISS interventions. However, carefully built microsimulation models like the Social Genome Model (SGM) allow us to observe the long-term outcomes of interventions like ISS based on a program's effects on youth.

The SGM was jointly developed by the Brookings Institution and Child Trends, and is now managed by the Urban Institute and Child Trends to inform policy discussions by modeling the development of children into adulthood. The SGM uses data from the 1997 National Longitudinal Survey of Youth (NLSY97), and includes characteristics and behaviors of youth from childhood into their early 30s. The model contains factors that affect success, including a respondent's family background, educational achievement, problem behaviors, substance use, college completion, criminal conviction, and earnings. These factors are sorted across six different life stages, from birth to age 29, and represent important contributors to development at each stage. The model can predict how altering one or more of these factors at a specific life stage can influence factors at later life stages. A more detailed description of the model and how it can be used is available in Appendix 4.

We use the SGM to predict the potential future outcomes of ISS participants using the effects of these programs on student math scores, graduation rates, teen pregnancy, and male incarceration.[4] The model employs a series of regressions to estimate how the effects of ISS interventions influence later-life outcomes at various life stages. Since the model only accounts for the significant effects of ISS programs, the simulated outcomes represent the potential long-term influence of these programs, rather than what can be expected from the average ISS program or what might be scalable to larger programs. The selected impacts were chosen because they come from high-quality ISS program evaluations and align well with variables in the SGM. However, no variables directly match teen pregnancy and incarceration in the SGM, so teen births and criminal conviction are used as proxies. Because ISS programs are usually targeted toward low-income students, we limited the sample in the model to youths whose mothers had only a high school degree or less. Models were run using both single and multiple effects of ISS programs. The results of the SGM simulations are reported in Table 10.

[4] The effects of ISS programs on teen pregnancy and male incarceration were obtained from a study on HCZ (Dobbie & Fryer, 2015). This was the only study to include these outcomes. In the HCZ study, sample sizes were relatively small for both males (N=233) and females (N=205), so it is unlikely that the impressive 100 percent decrease in male incarceration or 59 percent decrease in female teen pregnancy would be replicable in a scaled up ISS program.

Table 10. SGM Simulations of the Adult Outcomes of ISS Participants

Impact of ISS Programs on:	Life Stage		Effect Size	Intervention	High School Graduation (Age 19)	College Completion (Age 25)	Annual Family Income (Age 29)	Annual Personal Income (Age 29)
Math Scores	Middle childhood (ages 12–13)		+0.33 SD	City Connects (7th grade)[3]	+ 2.1 pp	+ 1.5 pp	+ $1,877.66	+ $913.03
Math Scores	Early adolescence (ages 14–15)		+0.45 SD	City Connects (8th grade)[3]	+ 1.0 pp	+ 0.3 pp	+ $28.90	+ $725.71
Graduation Rates	Adolescence (ages 16–19)		+11%	CIS[4]	-	+ 0.3 pp	+ $797.09	+ $470.87
Male Incarceration[1]	Adolescence (ages 16–19)		-100%	HCZ[5]	-	+ 0.3 pp	+ $1,674.59	+ $638.76
Female Pregnancy[2]	Adolescence (ages 16–19)		-59%	HCZ[5]	-	+ 1.5 pp	+ $599.06	+ $614.08
Male Incarceration[1] & Graduation Rate	Adolescence (ages 16–19)		-100% (Incar.) +11% (Grad.)	HCZ[5] & CIS[4]	-	+ 0.6 pp	+ $2,565.38	+ $1,162.56

Impact of ISS Programs on:	Life Stage		Effect Size	Intervention	High School Graduation (Age 19)	College Completion (Age 25)	Annual Family Income (Age 29)	Annual Personal Income (Age 29)
Female Pregnancy[2] & Graduation Rate	Adolescence (ages 16–19)		-59% (Preg.) +11% (Grad.)	HCZ[5] & CIS[4]	-	+ 1.9 pp	+ $1,387.69	+ $1,082.61

SD = Standard deviations
pp = Percentage points
[1] The SGM simulation uses criminal conviction as a proxy for incarceration.
[2] The SGM simulation uses teen births as a proxy for teen pregnancy.
[3] City Connects, Walsh et al., 2014
[4] CIS, Somers & Haider, 2017
[5] Harlem Children's Zone's Promise Academy, Dobbie & Fryer, 2015

These simulations suggest that effective ISS interventions on youth translate into beneficial outcomes in adulthood. Most notably, the simulations indicate that family and personal income are higher for ISS participants than nonparticipants. ISS participants also appear to have slightly higher high school and college completion rates. While there were also differences in other adult outcomes, these were generally quite small.

Interestingly, the effect of ISS on middle childhood math scores results in higher simulated adulthood educational attainment and earnings than is found for math scores in early adolescence, even though the effect size for the later life stage is larger. Perhaps the academic impacts of ISS programs on adult outcomes are larger when younger students are targeted because there is time for the positive effects to accumulate.

Additionally, combining the incarceration and high school graduation effects for males, as well as the teen pregnancy and high school graduation effects for females, results in better adult outcomes than for each impact alone. This suggests that the simulated differences in adult outcomes between ISS participants and nonparticipants are likely conservative, since ISS programs have other positive impacts that cannot currently be examined using the SGM. Regardless, these simulations give us a good glimpse at how the effects of ISS programs on youth may benefit them as adults.

Discussion

This review of ISS outcome evaluations finds results that are promising but not definitive. There are several null findings, but there are also hopeful, positive findings scattered across outcomes. It is noteworthy that there are almost no negative effects or impacts. Overall, although the evidence continues to build and there are several indications of positive effects, the field continues to lack a set of conclusive, consistent findings across outcomes or outcome types.

Based on our review, we have identified four factors that should be considered by evaluators and implementers going forward, as the evidence for ISS will be built by addressing these gaps.

First, the methodology that researchers choose for their evaluations impacts their (and their readers') abilities to make conclusions. Decisions about the evaluation design, comparison or control groups, measurement, and statistical analyses affect the kinds of conclusions that can be drawn. Some of the null findings we see in these results tables are likely the result of the different methodologies and measures used for analyses. For example, a short follow-up period may result in null findings because it takes longer to see impacts, and a comparison group that is poorly matched may mean the results are not valid estimates of the impact of the program. Taking care to use different study designs that match well with the design of the initiatives may allow researchers to tease apart small but significant effects in a way that current studies were unable to do.

For instance, the Communities in Schools 2017 evaluation randomized students within each school to either receive tier 2 services or not. However, it is hard to detect impacts when all tier 1 participants receive some base level of services. This means that the statistical models are testing for the additional impact of tier 2 services relative to tier 1 services, rather than the impact of the entire model (tiers 1 and 2) relative to receiving nothing. This makes the null results more understandable. A matched-pair cluster randomized study at the school level, while costly, would better address the need to match school populations while also allowing future evaluations to test for the effect of the entire model (tiers 1 and 2) instead of simply tier 2. This might allow evaluations to tease apart some of the differences between schools that invest the time, energy, and financial resources into creating the school culture shift that comes with implementing a Communities in Schools program.

Second, many evaluations continue to examine different outcomes and/or use different measures, including those obtained from different sources (e.g., student report versus school records). For example, measures for attendance include both attendance and absenteeism. The latter is a school-level variable while the former is an individual-level variable, and they capture different potential issues. For example, one school could have over 90 percent attendance every day, but have a handful of students who are chronically absent, while another school could have lower average daily attendance that is evenly distributed across the student body such that no group of students misses a significant number of days.[xcv] Although these both capture a measure of student presence in school, they are slightly different. When results differ for different measures, it is difficult to disentangle whether there is truly an effect or

whether the effect is specific to certain outcomes. Encouraging greater usage of the same measure or measures across studies would allow findings to be more comparable.

Third, there is variation in the length of follow-up period. These results show effects on grades and less consistent effects on testing scores. This may be because grades are more subjective—and reflective of student behavior and effort—while testing may be considered more objective because all students within a state receive the same test. If we believe that an improvement in grades is the first step on a longer-term trajectory of improved learning,[xcvi] we would expect that better grades would be the initial change, and that improved test scores would develop over a longer period of receiving support services. If so, it is possible that the length of time during which students are followed is insufficient, and that with longer follow-up some studies that examine test scores may see a shift. This is particularly reinforced by the next chapter's findings, which suggest that results strengthen as the program has been implemented in a school for a longer time.

Fourth, studies tend to examine each outcome in isolation. Researchers may control for confounding factors, but infrequently conduct analyses that examine the unfolding process by which ISS models may affect outcomes. Structural equation models, for example, would allow analysis of intermediate/mediating nonacademic variables and how they relate to longer-term academic outcomes.

Finally, more focus is needed on program implementation, which would provide answers to some of the very thorny remaining questions about ISS. Specifically, what explains success in some schools but not in others that use the same program? Are some implementation strategies more likely to result in better outcomes? How does leadership matter? What difference does the vulnerability of the student population make? How well do systems work to identify student needs? The lack of implementation findings in most of these quantitative evaluations—which will be discussed in more detail in the next chapter—leaves many outstanding questions unanswered. We need to better understand implementation approaches and quality to identify critical factors, and identify how to best support principals and teachers to achieve higher-quality implementation.

Conclusions

The cumulative body of rigorous evaluations continues to find positive patterns. First, several evaluations (City Connects, Harlem Children's Zone's Promise Academy, City Year, and CIS in Chicago) find promising results in terms of math and literacy grades and/or test scores. For these evaluations, the direction of results is clearly positive even though some specific years or grades yielded null results. The evaluations of these programs are highly rigorous and are described very clearly, such that measures and processes are transparent. They also show more consistently positive results than other studied programs.

Second, we see very few negative effects of participation in an ISS program for students. Even for programs without a consistently positive story to tell, we see null effects for academic outcomes more often than negative effects; many of the programs also seem to have positive effects on some nonacademic outcomes.

Third, few studies measure intermediate steps. This is problematic because it means that we do not know whether schools that fail to see improved outcomes are those that do not successfully improve intermediate outcomes, or whether the underlying conceptual model is missing another step. To state the reverse, what are successful schools doing that accounts for their success? Answering this question is increasingly urgent to define the most successful implementation strategies.

Finally, although there are numerous null (nonsignificant) findings seen in the results, these programs increasingly share a common conceptual model that builds from improved academic and nonacademic supports—in terms of both variety and intensity—to improved academic outcomes. This builds on everything we know about theoretical developmental science in terms of how children grow and learn, allowing us to strongly conclude that these models are promising. In a 2016 policy brief, the City Connects team reported that customized, comprehensive, coordinated, and continuous programs can successfully improve children's outcomes.[xcvii]

This chapter has focused on findings from the rigorous, quantitative evaluations, while the next (Chapter 5) discusses findings about implementation. An understanding of the essential parts of implementation is critical to the field, so that replication can continue and heterogeneity of outcomes can be better understood.

Chapter 5: Implementation Evaluations

By Hannah Lantos, Rebecca Jones, and Kristin Anderson Moore

Chapter Overview

As reviewed in Chapter 4, which reported on rigorous outcome evaluations of ISS models, enrollment in an ISS school is not consistently associated with the intended outcomes across different studies. Although the model is well-aligned with child development research and theory and despite some promising findings, there is a need to better understand program implementation. ISS implementation method in a school can differentiate successful programs from unsuccessful ones (and everything in between). The quality, intensity, and duration of program services is regularly found to be a critical factor in the effectiveness of early childhood[xcviii] and after school programs,[xcix] but implementation is remarkably under-studied in ISS evaluations.

This chapter is structured around two sections. First, we will expand on the quantitative findings to explore the implementation lessons from several of the rigorously evaluated programs discussed in the previous chapter. To understand the factors associated with better implementation, we reviewed the five implementation (process) evaluations done in conjunction with the outcome evaluations. Specifically, we identified which evaluations included assessments of implementation quality and fidelity—either descriptively or through analyses linking them to outcomes—with the goal of identifying what happened, what worked, and what seemed to matter most. These findings provide guidance on what should be replicated in future program and evaluation development.

Second, we will share new qualitative findings from interviews that Child Trends conducted with principals implementing ISS models in their schools. These interviews focused on the five core components of ISS models described earlier in this report. They first explored whether these components still accurately capture what happens on the ground, and then examined how principals and teachers implement these new processes in the day-to-day life of the school, what has been successful, and what remains challenging.

This chapter focuses on answering the following research questions and highlighting where questions remain:

- What indicators of high-quality implementation are associated with better student outcomes?

- What are the critical components of the ISS model? Are the five components identified in the 2014 review still appropriate? Have any new elements been identified?
- How is each ISS component implemented? Are there critical aspects of each component? Are there elements of the components for which there is no evidence?
- Are there additional implementation elements that have not yet been incorporated into evaluations—elements frequently highlighted as critical for schools that future evaluations should assess?

As stated previously, our conclusions in this updated 2017 report echo those in the earlier review: we still conclude that the evidence in support of integrated student supports is promising, but not conclusive. With this chapter, our caution stems primarily from the fact that the concrete recipe for success is as yet undefined. Additionally, we do not yet know who will fully operationalize the ISS framework to scale rapidly in an effective manner, or how they might do so. The implementation studies add to our understanding, but more studies will be necessary to explore the intricacies of implementation.

Implementation Findings from the Evaluation Studies

The six implementation evaluations described here were part of rigorous outcome evaluations. Two types of evaluations were conducted: (1) explorations into whether and how components of the model were implemented (process evaluations); and (2) how completely (with fidelity) or well (with quality) schools implemented a set of predefined components.

These evaluations often helped program implementers and evaluators place their understanding of evaluation results in context. For instance, if there was no initial community support for the program, we might understand a lack of findings within this context and explore why there was no support. Additionally, if there are no statistically significant gains for participants, analyzing the extent to which an intervention was implemented with fidelity and quality can help explain the lack of significant improvement. Was it because the program was not implemented well, or that it may need to be changed because, even with excellent implementation, it did not positively impact outcomes?

We describe the five studies in chronological order below. We start by describing their different approaches and conclusions, and end this first section by discussing what we learned and which key research questions remain.

Implementation Evaluations

Comer's School Development Program (SDP)

Study: Cook, T., Habib, F.N., Phillips, M., & Settersten, R. A., Shagle, S. C., & Degirmencioglu, S.M. (1999). Comer's School Development Program in Prince George's County, Maryland: A Theory-Based Evaluation. *American Educational Research Journal*, *36 (3), 543–597.*

This was a large-scale evaluation of the Comer School Development Program (SDP), which included 23 of the 25 middle schools in Prince George's County, MD, with data from over 12,000 students, 2,000 staff members, and 1,000 parents. Researchers evaluated implementation during a four-year randomized control trial (RCT). Data were collected from an annual staff implementation questionnaire, a student school climate survey completed at the end of seventh grade, annual interviews with facilitators and the district Comer coordinator, interviews with school principals in the first and last year of the study, and a telephone survey of parents to gauge their perceptions of school engagement.[1] There are only two years of student data for each student because middle schools at this time in Prince George's County only had seventh and eighth graders. The researchers explored three questions. First, was implementation higher in treatment schools? Second, did implementation improve over time? Third, did children in treatment schools do better than those in control schools?

First, SDP schools were able to actually implement the model, a question assessed through a quality measure developed by the researchers. This means that components of the model were seen more in the SDP schools than in the comparison schools. While this is to be expected because non-SDP schools did not use the model, the difference was quite small.[2] Researchers concluded that the variation was due to the role of the program facilitator (a role that only existed in program schools) and the perceived extent of parent and community involvement, which was higher in program schools. The authors comment on the perceived parental involvement, describing it as more symbolic because it did not translate into increased parental homework help or parent-child communication.

Second, the researchers also found that implementation quality improved over time, particularly from years 2 to 3. However, quality did not improve more in program schools than in control schools. Researchers found no correlation between the quality of the implementation and a number of school characteristics, including school size, average student socioeconomic status (SES), free lunch, and absenteeism. However, there was a negative correlation between quality and the percentage of black students; and a positive correlation between quality and tenured staff that also increased over time, suggesting that students who were already particularly vulnerable (either racial minorities or in schools with high teacher turnover) were less likely to attend a school with high-quality implementation.

Finally, the researchers also linked implementation quality to outcomes, with somewhat complex findings. There was no evidence that being in a Comer school with higher quality was linked to better academic or nonacademic outcomes; however, some relationships were found when analyses were run

[1] The implementation questionnaire assessed staff perceptions of 1) the School Planning and Management Team, 2) the Social Service Team, 3) the Parent Teacher Association, 4) the school improvement plan, 5) the communication between teams, 6) the use of child development knowledge throughout the school, 7) whether decisions were made by consensus, 8) the commitment level of team members to improving the school, 9) the degree to which all members of the school community were included in decisions, and 10) the inclusion of cultural and racial groups. The items were analyzed both separately and collectively.

[2] The difference between SDP schools and non-SDP schools was 0.15 units. This was the average for all four years over all 11 items measured on a five-point scale. The authors did not comment on statistical significance.

at the individual level using an implementation scale that measured how "Comer-like" a school was (essentially how much even non-SDP schools met some Comer criteria). Specifically, being an SDP school did not affect school climate or student outcomes, but schools with "Comer-like" qualities were found to have some positive effects. Specifically, there was a positive correlation with Comer components and changes in social behavior, psychological adjustment, and attendance. These schools also had slightly more negative math scores.

Communities in Schools (CIS)

Study: ICF International. (2008). Communities in Schools National Evaluation Volume 1: Results from the Quasi-Experimental Study, Natural Variation Study, and Typology Study. ICF International: Fairfax, VA.

This was a large-scale evaluation of the Communities in Schools (CIS) program. Researchers evaluated a two-year period of implementation of the CIS program in schools in seven states (Florida, Georgia, North Carolina, Michigan, Pennsylvania, Texas, and Washington). Data were collected from two surveys of sites—one focused on the comparison between CIS and non-CIS schools and one focused on what created that variation in outcomes across only CIS schools. These surveys included questions about needs assessment, referrals, services, and monitoring/adjustment. The one that also focused on non-CIS schools was administered earlier and included a domain on planning. The first survey was meant to be short and filled out by as many sites as possible (1,894 schools eventually filled it out). The second survey was only offered to 576 eligible CIS schools (368 completed it). The researchers in this study sought to answer two primary research questions. First, does quality implementation explain the variation between high- and low-performing CIS schools? Second, do high-quality CIS schools perform even better relative to non-CIS schools than the full sample?

Researchers developed a scoring rubric for each survey that identified key components of the CIS model and scored schools from 0–5 based on whether they incorporated each component (and/or how intensively they incorporated it). This scoring was based on conversations with school-based staff to identify "tipping points" for success. For instance, if a school conducted a needs assessment once per year, it received 3 points; if it conducted one less than once a year, it received 1 point; and if it did so more than once a year, it received 5 points. Some questions were binary (0 versus 5) while others ranged from 0–5. Researchers also identified eight different service domains, including 22 different types of services (such as mentoring, case management, or pregnancy prevention), and assessed the number of hours spent working in each domain.

Researchers found that just under half (47.6 percent) of CIS sites were "high implementers," compared with 52.4 percent classified as "partial implementers" (sites that scored fewer than 70 points on the second survey).[3] Overall, there was a positive association between high implementation and outcomes (high implementers had better outcomes than partial implementers, specifically in promoting power, graduation rates, and fourth- and eighth-grade reading and math). Researchers also found a small

[3] Appendix C of Volume 1, page 140. You have to scroll all the way down to the actual typology report.

positive association with attendance at the elementary school level, and a slightly more positive association at the high school level. In contrast, they found that sites classified as partial implementers had better tenth grade math scores, middle school attendance, and tenth grade reading scores than high implementers. Additionally, CIS has created a list of eight key service domains—the buckets in which student supports fall: maintaining family and peer relationships, academics, case management, behavior, after school, career, public service, and health. The high-implementing CIS schools put more hours into the eight service domains than their lower-performing peers.

This evaluation also categorized schools with better implementation scores as high implementers. Relative to non-CIS schools, high implementers performed even better than the overall sample of implementing schools. The patterns were similar to those mentioned above, but in a few instances the partial implementers actually did worse than non-CIS comparison schools. In terms of promoting power; graduation rates; elementary school attendance; fourth, eighth, and tenth grade math; and fourth and eighth grade reading, the partial implementers actually did worse than non-CIS comparison schools overall—suggesting that high-quality implementation is one of the key drivers of successful CIS schools.

Diplomas Now

Studies: 1) Corrin, W., Sepanik, S., Gray, A., Fernandez, F., Briggs, A., & Wang. K. K. (2014). *Laying Tracks to Graduation: The First Year of Implementing Diplomas Now*. MDRC: New York, NY.

2) Corrin, W., Sepanik, S., Rosen, R. & Shane, A. (2016). Addressing Early Warning Indicators: Interim Impact Findings from the Investing in Innovation (i3) Evaluation of Diplomas Now. MDRC: New York, NY.

The Diplomas Now (DN) study was part of an RCT conducted by MDRC to study the implementation and preliminary results of Diplomas Now (DN) nationally. In the first year of the study, there were 22 schools (12 DN and 10 control schools); an additional 40 schools (20 DN and 20 control schools) were recruited and added to the study in the second year for a final study sample of 62 schools (32 DN and 30 control schools). The researchers had one primary research question and one sub-research question. This is an interim report (the final version will be available within the next two years), and the researchers were most interested in describing exactly what implementation looked like—were these schools able to implement DN with fidelity? Second, they conducted a preliminary exploration of the link between fidelity and outcomes, noting that it should be interpreted cautiously as an interim report.

The DN model has four pillars of success: teacher teams and small learning communities, curriculum and instruction with professional development, tiered student supports, and a "can-do" culture and climate. The researchers identified more than 100 components that they felt represented full implementation of these pillars, which they used as a measure of fidelity to the model. Program staff surveys; school administration and teacher surveys;[4] and interviews and focus groups of school staff, parents, and

[4] There were 94 administrator respondents from 31 schools and 742 responses from sixth- and ninth-grade teachers at 32 schools.

students[5] were all used to identify whether each site implemented each of the 100 components. A score of zero indicated that there was no or low implementation, whereas a score of one indicated implementation with fidelity to what was intended by the model.

During both years of the evaluation, the average fidelity score of DN schools was just over 60, meaning that most schools were implementing over half of the components with high fidelity. Schools were most successful in hiring staff to implement the model, using data to identify students in need of additional supports, and coordinating interventions for individuals and small groups of students. Many schools were less successful at offering peer coaching to teachers or involving parents and community members in their initiatives.[6]

As this is an interim report, the authors appear hesitant to delve deeply into linking specific implementation findings to outcomes in answer to their second question. With longer follow-up times, they anticipate being able to explore implementation in more depth. However, they highlight two patterns. First, second-year DN schools were more likely to engage in activities that fell under each of the four pillars, suggesting that the model was being implemented and differentiating DN schools from non-DN schools. Second, sixth graders in DN schools were more likely to have better academic outcomes compared to those in non-DN schools. No significant differences were found for ninth graders.

City Year

Study: Meredith, J. & Anderson, L. M. (2015). Analysis of the Impacts of City Year's Whole School Whole Child Model on Partner Schools' Performance. Policy Studies Associates, Inc.: Washington, DC.

The City Year implementation study was a component of the bigger evaluation of the program itself, including an outcome evaluation. A total of 327 schools participated in the quasi-experimental evaluation (143 elementary schools, 79 middle schools, and 81 high schools). Some schools started the program in the 2011–2012 school year and the rest were added over the next two years (for the 2012–2013 and 2013–2014 school years). The researchers were interested in understanding whether schools were able to implement the program with fidelity, and whether certain components of implementation were associated with improved outcomes. Their analyses linking implementation to outcomes used three specific items to focus exploration: implementation quality (from the overall implementation index), the ratio of AmeriCorps members to students, and when schools began their partnership with City Year (which measured the length of time the model was implemented in each school).

---

[5] Forty-nine interviews with school-based staff (school transformation facilitators, City Year program managers and team leaders, CIS site coordinators, instructional coaches, school administrators, and school counselors) and district-based staff (Diplomas Now instructional facilitators, field managers, school and student support services facilitators, Implementation Support Team representatives, and school district leaders); twenty-eight focus groups were conducted with parents, students, teachers, and City Year corps members for a total of 173 participants.

[6] Corrin et al., 2014; Corrin et al., 2016

The implementation indices were created using survey data from City Year program managers at the end of the year. The survey measured 39 different indicators of implementation. Each indicator was given a point value (the authors do not state clearly whether this was one point per item or categorized differently), and sites were given scores calculated as their percentage of total available points. From these, the evaluators created several indices, including an overall fidelity index and specific indices measuring the math, English language arts, attendance, and behavior and social-emotional learning components of the model. Implementation scores varied across sites, with San Jose having the lowest overall average percent (47) and Chicago having the highest (82). Twelve sites with scores above 60 were categorized as "high implementing" sites. Scores also varied within sites, with Washington, DC having the most variation in scores (from each of 12 schools in the city). Chicago and Philadelphia both had narrow bands of variation. There was no correlation between the number of partner schools per site or site size and implementation scores. These results suggested that schools and cities were able to implement the model with fidelity, although some cities clearly struggled to do so consistently.

The researchers then explored whether the three variables identified above were related to outcomes. Their findings are complex. First, the authors studied whether schools that began their partnership before 2011–2012 saw more academic improvements. They found that schools that partnered with City Year later saw more improvements on math assessments. Second, researchers found that high implementing schools[7] were more likely to show improvement on ELA assessments, but that these results did not translate for math assessments. Third, ELA scores improved at the elementary, middle, and high school levels if there were three or more AmeriCorps (corps) members for every 100 students. For math scores, improvements were seen at the middle school level and marginal improvements were seen at the elementary school level with the same corps-to-student ratios. This study makes it clear that higher implementation quality and better student-to-staff ratios were important to improved outcomes. The finding about length of program time is interesting and deserves more exploration into potential causes.

City Connects

Study: City Connects. (2012). *The Impact of City Connects: Progress Report 2012.* Boston College Center for Optimized Student Support: Boston, MA.

City Connects mentioned a fidelity measure in a previous report, but fully presented the process of development and its components in this 2012 annual report. The development of this measure involved four steps: reviewing the practice manual to define the critical components of the intervention, selecting key components of practice essential for other programs to have fidelity to the model, determining the facets of each component, and creating indicators for each facet. City Connects ultimately developed seven key components, each with four to eight facets. These components were as follows: whole class review, individual student review, community partnerships, family partnerships, health and wellness,

[7] High implementation sites, as defined in the chart on page 16, are those that scored an average of 68 percent on the implementation index.

opening of school, and close of school. We will not describe each facet and indicator (as these can be found in the City Connects report), but this very intentional process was meant to be transparent and very easily operationalized on the ground. City Connects notes that the fidelity of its model across schools is quite high, with schools scoring 80 to 100 percent on the fidelity scores for each component. City Connects does not, however, test whether schools with higher fidelity have better student outcomes.

Talent Development

Study: Kemple, J. J., Herlihy, C. M., & Smith, T. J. (2005). Making Progress Toward Graduation: Evidence from the Talent Development High School Model. MDRC: New York, NY.

The Talent Development evaluation, conducted by MDRC, includes a total of eight schools in Philadelphia that initiated program implementation between 1998 and 2003. Two schools began the planning year during the 1998–1999 school year, and either one or two additional schools were added each year, with the last two schools beginning their planning year during the 2002–2003 school year.[8]

Unlike the other implementation studies included here, this study is included because of the detail provided on the program's start in Philadelphia. The authors note that the relationship between Talent Development and the Philadelphia school district was not a formal relationship, but was received positively by the superintendent, who was familiar with Talent Development from his tenure in Baltimore, Maryland, where the program began. The report does not analyze the measures used to evaluate implementation or the relationships between implementation and outcomes, but it does provide some context that might be important to understanding implementation and subsequent outcomes. For example, because the program was not formally sanctioned by the school district and was not presented as a reform model of choice, reception from school leaders and the subsequent decision to implement the program varied. This contextual information sheds light on some early political and implementation choices that can lead to more successful program completion.

Conclusions from Implementation Studies

Looking at overall implementation results, schools with higher implementation scores seem to also have better outcomes. Comer, CIS, and City Connects found this to be true in terms of student-staff ratios, fidelity to the model, and successful implementation of more core model components. This may reflect buy-in, willingness to problem-solve because of a belief in the value of the program, strong leadership, or successful identification of an approach that works when fully implemented with quality. However, without knowing whether a program was implemented well, it is difficult to understand null or negative findings. Moreover, for programs with positive outcomes, it is not clear which components account for program success. Additionally, as this research agenda continues to advance and gain depth, teasing

[8] The sixth school to begin implementation began to close after only the first year of school implementation.

apart which components of models are most important for program success will allow schools to invest limited resources in the most effective programs, processes, systems, and people.

Three additional findings from these evaluations can be highlighted. First, there must be flexibility in the model so that it can meet schools' needs, which was highlighted in the CIS and Talent Development studies. In schools implementing Talent Development, school staff used the first year to plan and observe other schools and then develop a specific plan for their school's needs. This allowed administrators to think about what processes would work in their specific school environment. The context also provides a backdrop for understanding other characteristics—some of which are more difficult to measure, such as initial reception in schools—that help describe variation in implementation and why a certain site may have been more (or less) successful at implementing the model. Understanding context may allow program implementers to more quickly troubleshoot issues and navigate personalities, conflicts, or competing needs for resources.

Second, after understanding context, planning, and politics, it is important to identify which model elements must be implemented and scaled. The Diplomas Now study focused on capturing the core components of the model, which included detailing the four pillars for success and developing specific indicators from these specific pillars. CIS has developed a similar rubric. This not only provided a clear set of indicators, but held schools specifically accountable to the model in a way that was transparent. It also allowed researchers to specifically identify where the model fails and to add resources to support those components. Identifying where a school falls short can assist researchers and program staff in identifying key, core components of each model associated with greater success.

Finally, different programs measure implementation differently. Some use a quality measure while others measure fidelity to a clearly described model. It is important to measure the association between either quality or fidelity, or both, and outcomes. Although the Diplomas Now study describes in detail how implementation was measured, the authors cautiously note that length of implementation and the age of students are associated with differential outcomes. We note these since they were reinforced by the City Year study, which did link implementation to outcomes. Generally, higher implementation resulted in greater gains (with some exceptions). Programmatic staff know that turning program data into variables that are useable in quantitative models can be challenging. This can limit the nuance or depth of factors related to implementation that are studied and linked to outcomes. Ratios, length of time, and indices of multiple indicators leave questions about *why* these are important factors that need to be studied further. The City Year authors do not provide a theoretical explanation for why different implementation factors might be related to improved outcomes, but many are self-explanatory. For example, a higher AmeriCorps-to-student ratio makes logical sense, as students likely received more one-on-one attention when more AmeriCorps members were present. On the other hand, the fact that schools implementing City Year later performed better was initially confusing, as length of time implementing a program could theoretically be correlated with improved outcomes. Ultimately, we understood this finding in light of the fact that the later year was when the program scaled up, suggesting that this was when there were more resources or a better, more fluid program to implement. We also hypothesize that this finding might be capturing other important factors like model refinement over time, knowledge about which additional resources were needed to effectively implement the

model, or learning from challenges in other schools in previous years. Implementation researchers must remember that creating binary or categorical variables that can be included in regression models still needs to be explained and explored once results are found. More information about these studies is included in the appendix.

Ideally, an implementation evaluation would encompass the three critical aspects—flexibility, understanding of context, and linking quality to outcomes statistically—that were highlighted in these evaluations. Talent Development provided background for context and to explain the introduction of a model to a school or district. Diplomas Now detailed the key aspects of the model and a method for measuring implementation. City Year analyzed the relationship between implementation and outcome measures. SDP Comer Schools focused on staff and student perceptions. Finally, CIS emphasized the variation that begins to occur when programs are scaled. All provide insight into how one can measure interpersonal indicators or logistical components.

There are, however, several limitations to these studies. We faced several challenges in drawing conclusions as we analyzed patterns across the five studies. First, most of the implementation evaluations were conducted (or at least planned) well before the 2014 report, so it is not surprising that the five elements identified as characteristic of ISS models (shown in Figure 1) were not systematically covered. Since these elements were identified on the basis of conversations with programs and vetted with stakeholders, it would be helpful to obtain information (or organize the information obtained) around these five elements in future implementation evaluations. This will provide some structure to compare across studies moving forward.

Second, while it is clear that program staff and evaluators understand that high-quality implementation or fidelity to a specific program model yields better outcomes, few studies included an implementation component, forcing our conclusions in this chapter to be based on a small sample of just five studies. Third, each evaluation was quite different, which limited our ability to compare across implementation evaluations.

Fourth, there appears to be little consensus on what specific aspects of implementation must be measured. While the five ISS model components shown in Figure 1 were incorporated into these models, the specifics of their implementation remain unclear. Like the variation in outcomes measured in the last chapter, programs used different measures of quality or fidelity, and evaluators linked them to outcomes in different ways—again limiting our ability to conclude that any one component was essential.

Fifth, most studies relied on schools' self-reports indicating whether they were implementing certain components of the model. This is more a measure of whether administrators at each school think they are implementing such components than an outsider's more objective assessment of quality. Finally, few studies measured fidelity from multiple perspectives—knowing whether students think certain programmatic components were implemented may be an important measure of implementation quality.

Principal Scan

In the first section of this chapter, we included a list of research questions and began to answer the first two. Specifically, we described what existing implementation studies have found, and identified indicators of high-quality implementation found to be associated with better student outcomes in the literature. However, because so few studies made the linkage between implementation quality and outcomes, conclusions remain tentative.

In this second section of the implementation chapter, we share insights about ISS implementation from more than 20 conversations with principals and program staff that aim to answer the last two questions: 1) how are the five core components implemented, and 2) were we missing any critical components? These conversations were structured such that we could explore whether the core components identified in 2014 continue to capture what these programs do, whether principals would add anything essential now, how each component is implemented, whether principals feel that any key components are most important for improved outcomes, and which challenges to implementing an ISS model are biggest. These interviews have allowed us to develop a better understanding of how schools and organizations implement ISS models in schools on a day-to-day basis. Not only do these five core components still very much resonate, but educators highlighted some skills and resources needed to do this work on the ground, as well as some challenges.

Many principals spoke of various integrated student supports in a way that encompassed other initiatives that were often complements of ISS. Specifically, they spoke about Multi-Tier Systems of Supports (MTSS)—including Positive Behavior Interventions and Supports (PBIS)—and implementing restorative justice practices. Our conversations confirmed that ISS is an approach rather than an "add-on," or new initiative to replace the work that schools were already doing. Interviews also highlighted that ISS incorporates other practices and school resources for implementation. Insights into principals' successes and challenges related to implementation, buy-in, or systems change may be useful to other schools rolling out similar or slightly tweaked ISS models, and to researchers designing future evaluations. Some themes heard in these conversations are shared below.

Methods

Interviews were all structured similarly. Each conversation varied from 30 to 60 minutes, with most lasting around 50 minutes. Interviews were designed to understand the specific model implemented in each school and to explore how each school implemented the five core components. Many principals had not seen or heard of the five core components, but it was clear that many had incorporated these components into their models—either intentionally (if the model was a nationally implemented model like CIS) or unintentionally (if the principal had created their own model). As the interviews developed, we also asked some of the later principals whether our observed patterns made sense. The five core components are defined again in Table 11, below.

Most of the schools were urban, but three principals were in rural or smaller communities. There were five elementary schools, eight middle schools, two K–8 schools, and four high schools; we talked to

principals in most regions of the United States with the exception of the Southwest. Twelve schools were in the West while seven were in the East. Most schools were affiliated with CIS, but there were also schools from City Connects and the Children's Aid Society in New York, and some with no programmatic affiliation.

Table 11: Five Core Components of the Integrated Student Supports Model

Needs Assessment	Community Partnerships	Coordination of Supports	Integration within Schools	Data Collection and Tracking
A comprehensive needs assessment is conducted at the student and family level, but may also be necessary at the program, class, school, and/or community levels. The assessment identifies existing strengths, challenges, and gaps in services.	Establishing, fostering, and sustaining relationships with existing community organizations is the only way to support students in all the necessary ways. ISS models emphasize community partnerships to help schools and families provide a full array of resources and supports to students.	Coordinating supports requires creating a system where all student supports are provided as planned and followed up on, in terms of whether issues were addressed or additional supports may be needed.	The key to integration is making sure that all adults in the school understand the resources available to support students, as well as the processes needed to support students in accessing such resources.	To ensure that students actually receive identified services and assess whether they result in the desired improvements in outcomes, it is essential to collect data about services received, fidelity of implementation, and outcomes. If outcomes do not improve, data collection and tracking allow for quick modification or a re-direct to other services.

Thematic Findings

Needs assessment

Principals provided a range of responses that pertain to how schools currently conduct needs assessments in schools and communities. When this line of questioning began, we used the term "needs assessment" to mean the preparatory work that goes into starting ISS model implementation in a school. We envisioned this including an assessment of the types of needs that students have relative to the types of resources (community connections or school counselors, for example) available, to define the school's remaining needs.

In their answers, principals highlighted two types of needs assessments. The first happens when a school, principal, or school district is beginning to implement an ISS program. The second type happens at the beginning of each school year and identifies needs for each student in that year. Although many principals identified this as a needs assessment before prompts about programmatic needs assessments, the type aligns more with the fifth component, Data Collection and Tracking. As such, the five components do not represent a fixed sequence. Rather, they are generally, but not inevitably, implemented in the order that they appear in the ISS model shown in Figure 2.

The initial needs assessment often happens during the planning stages of the program. Many principals noted the importance of conducting some level of assessment to understand and address needed supports for students; however, many noted that the process was only a process in name, with little regularity. Schools fall into two categories: those with a principal-led assessment or those that partnered with an outside organization.

Schools with principal-led needs assessments tended to focus mostly on collecting attendance, behavior, and course performance data and developing a plan—sometimes with community support—to address needs in these areas. Some principals mentioned that these efforts were guided by other models, such as PBIS (a type of MTSS). Others shared that their needs assessment was ad hoc, gaining specificity over time as the model developed.

Very few people conducted a needs assessment that included the local community perspective, and family perspectives were also less systematic. Principals receive training on tracking student outcomes, but are not systematically trained to think about systems change, community involvement, or how to design a needs assessment. This reality sheds light on why some initial needs assessments may be less organized and evolved over time.

Needs assessments were sometimes conducted by schools in partnership with outside organizations or research centers (e.g., universities), including City Connects, CIS, or the University of Washington. These tended to be more systematic and broad, although principals sometimes knew less about the process when someone from the outside had conducted it. For example, schools that partnered with CIS receive guidance from the national organization and local representatives, but because this work is done primarily by the CIS site coordinator, schools were unable to detail many of the logistical layers of conducting a needs assessment. However, principals indicated that a needs assessment is conducted at the beginning of the program, and noted that data are also collected at the beginning and end of each school year.

Community partnerships

As with implementing a needs assessment, there was a fair amount of variability—in terms of both how and to what extent school leaders have been successful—around establishing and sustaining community partnerships. An integral part of implementing ISS is the establishment of partnerships with local organizations and businesses in the community that provide services, but which may not traditionally occur within the school building or during the school day. These include services such as emergency food provision, housing assistance, afterschool care or programming, or the provision of eye exams,

washing machines, or mental health evaluations. Typical relationships between schools and their communities in the past have come in the form of local business sponsorship for school events (e.g., sports or talent shows); however, ISS calls for a more intensive, sustained, and purposeful partnership. One such model, community schools,[9] focuses on clearly distinguishing how both the physical school building and the partnerships are key components of the model. Other ISS models may not fall under the umbrella of community schools, but all must integrate resources within the community and have a certain level of cultural competence to effectively use those resources with the community they serve.

Developing relationships with the community is not necessarily a skillset intentionally imparted to principals. Some focus on these partnerships, but managing teachers and students and their families is a full-time job. Being a community leader, liaison, and representative requires more work, more time, and more skills. It is not surprising that those schools which are most successful in this area have principals who are natural leaders, and hire a full-time staff member to lead partnership development and maintenance. Many principals outsource this role to their school coordinator, who becomes the key liaison to outside organizations. The principal then becomes engaged only when issues arise or successes are celebrated.

Based on the interviews, many schools that have been successful have partnered with an organization that facilitates connections with community-based organizations, and/or are located in a large city with numerous organizations that provide services. The most common partnership across all schools was one with a mental health provider; in many cases, this seemed like a gateway to other partnerships. Other partnerships include the Salvation Army, Boys and Girls Club, libraries, churches, soup kitchens or food banks, area colleges and universities, and community action agencies though the Community Action Partnership.

Identifying community organizations is more difficult in smaller communities where community-based organizations may be less prevalent than in larger cities. It is essential to think creatively about whom to partner with (and how to partner) in these smaller communities that do not have the number of options readily available in larger cities. In contrast, in larger metropolitan areas, the challenge is to first develop a comprehensive list of what is available and what services are provided by whom, where, and when; and then to prioritize which services are most needed and which partnerships will be most beneficial.

Many school leaders recognized this as an area of weakness and identified community partnerships as an area of needed growth. One leader learned that building connections was initially time-consuming, while another indicated that the focus on local organizations was particularly relevant to ground students in their own community because of the high percentage of students who will eventually remain

[9] The Coalition for Community Schools defines a community school as "both a place and a set of partnerships between the school and other community resources. Its integrated focus on academics, health, and social services, youth and community development and community engagement leads to improved student learning, stronger families and healthier communities… Schools become centers of the community and are open to everyone – all day, every day, evenings, and weekends."
http://www.communityschools.org/aboutschools/what_is_a_community_school.aspx

there. This school leader also stated that it was important to find organizations that provide students with skills beyond the classroom (life or work skills). These include experiences such as swimming lessons that could lead to lifeguard certification, glassblowing to engage students in both arts and school, and physical education through a local bicycle store to teach students how to build a bicycle. Another school leader noted the importance of being intentional about partnerships and identifying the need they fill; if this is not done, there can be rapid partnership burnout because the time required to maintain partnerships is not sustainable. Finally, one principal mentioned a focus on making sure that partnerships are mutually beneficial, perhaps by having students volunteer with partner organizations so that some effort to maintain the partnership goes both ways. For example, at one school, students volunteer at the food pantry that also provides emergency food bags to students in need.

Some organizations facilitate the development of community partnerships by creating a database of partners within a community, organized by service. This resource helps the school point person identify partners based on the specific needs identified during the needs assessment process.

Coordinating services

For coordination, we were interested in learning more about how schools prevent students from falling through the cracks, and how students with multiple needs (that may be reported by different teachers at different times) are case managed successfully. Generally, most school leaders indicated that there were regularly scheduled meetings, weekly or biweekly, where a team of guidance counselors, social workers, administrators, and/or behavioral/intervention specialists meets to discuss students identified as needing additional services; those students are typically identified as tier two or three when the language of MTSS is used. During these meetings, the team discusses student progress and may identify students to flag for additional support. Each meeting focuses on a select number of students specified by timing (six- to eight-week intervals), or the group meeting prioritizes a student, or students, due to concerns.

There seem to be two challenges in this category, one identified by principals and one by Child Trends' researchers. The challenge identified by school leaders is the occasional lack of resources for kids with the most need (the students in tiers two or three). Identifying these students and integrating processes throughout the school is helpful, but if schools lack resources to provide services, the success of the model reaches its limits. Additionally, in the 2014 report, coordinating services was intended to capture the kinds of communication that happen across service providers, both inside *and* outside the school building. Child Trends researches noted that, given a few strong community partnerships that provide high-quality services to students, coordinating services begins to overlap significantly with integration of processes in the school. To get to a place where coordination will stand on its own requires stronger community partnerships that work closely with schools to serve students.

Similar to developing community partnerships, many principals hire school coordinators to be a full-time staff member in charge of coordination. This person is often in charge of both coordination and partnerships because—if they know all the players—it allows them to also ensure that everyone understands what is going on and the nature of their role. This is especially relevant because bringing

community partners into school meetings can be challenging, as meetings sit within the structure of the school in terms of times and necessary staff . Additionally, integrating outside agencies brings up issues of privacy and data sharing, so Child Trends did not learn of a school that regularly brought community partners to these meetings. However, thinking about coordination in full would require a system for communication across services, including services outside the school building. Accommodating the very different schedules of partners in the community can be hard, so it is essential to have a point person to make these linkages even when in-person meetings are challenging.

Integrating supports within schools

There was some overlap in how schools coordinated and integrated services. Child Trends was interested in how teachers learn new processes in the system. Specifically, how are referrals made, how was this communicated to teachers, what data are teachers expected to collect or report, are they expected to engage in certain interventions before they reach out for support, and when did the new system start to feel like a system with processes that worked? Did teachers ever begin to feel like the new systems made their work easier and served children better?

Principals commented on a number of these questions. One theme highlighted in several conversations was the shift in school culture. One organization leader stated that implementing ISS was not just a thing that was happening, but instead that it was a process. School leaders were able to articulate the integration as a shift in the way teachers and school leaders operated. For example, teachers knew they could rely on school leaders and personnel to share information, and were watchful for changes in student behavior—whether improvements to meet goals or declines that might require intervention. Schools that were able to hire personnel specifically for ISS infused that person into the day-to-day processes of the school. This person established relationships with administrators, teachers, and community organizations. With teachers in particular, this person was considered someone to rely on, according to school and organization leaders.

The process also required training teachers to use a new system and informing them about what new services were available. The referral system was a new process as well. Although it was not clear from conversations how intentional these processes were, school leaders worked to implement systems that worked for their schools—whether completing a referral form or emailing and/or text messaging the contact person.

Data collection and tracking

As noted in the interviews, principals tended to focus first and foremost on the ongoing collection and use of data. This focus on ongoing needs assessment and tracking may be explained by several factors. First, some principals noted they had not been at the school when the program was initially implemented, highlighting why they chose to focus on ongoing needs assessment. Second, identifying student needs is ongoing and more salient when principals are asked how they manage their ISS model. Third, many principals have been trained to look at micro-student data (information about individuals), and have never been trained to do macro-level systems analyses (aggregated data). Thus, it should perhaps be unsurprising that their organizational focus leans more toward the former than the latter.

Schools indicating that they had not conducted a needs assessment at the onset of implementing new programs did often have a process for periodically analyzing student data to assess (and reassess) student needs and identify supports. Many have systems for assessing students at the beginning of each school year, others have student support teams that meet on an ongoing basis to identify and address student needs, and others have coordinators (hired themselves or through outside organizations) who regularly check in on teachers, students, and families.

One CIS high school works closely with the local CIS middle school to identify students for additional supports before they begin the transition into high school. These students are then placed on the caseload for the site coordinator. As necessary, guidance counselors continue to identify and refer students to the coordinator. This was perhaps one area where rural school districts have an easier time than urban districts, because this type of communication across schools could be facilitated more simply in smaller districts. City Connects has full-time staff at each school to serve as the program coordinator, utilize the needs assessment to gain understanding of what services are necessary in the community, and be a vehicle for teacher buy-in. Because the data for needs assessment and tracking are collected with community partners, principals, families, and teachers, organizations can get a sense of the varying perspectives of all stakeholders. In addition to providing an opportunity for teachers' voices to be heard, the needs assessment also provides an opportunity to identify both weaknesses and strengths. Communities in Schools also does end-of-year needs assessments, which are used to assess outcomes and progress toward the goals outlined at the beginning-of-year needs assessment.

Overall, almost all school leaders collect attendance, behavior, and course performance data (the ABCs). These data seem to be consistently collected, and either school-chosen or district-mandated data collection systems are used to house and track them. Several school leaders mentioned using SWIS—a data collection system—to collect these data, and other methods include Skyward, Excel sheets, and Google Documents. Many principals noted that they keep regular tabs (even daily) on these data. A few schools receive data directly from the school district based on information collected at the school level, including attendance, discipline, and grades. Sometimes this includes early warning indicators to identify students of concern who have either failed a number of classes, missed a lot of school, or received a concerning number of behavioral citations.

In addition, many schools also collect data on services needed and received in various systems—including, for example, which students need additional food sources, which would benefit from behavioral health supports, or interactions with partner organizations in the community (such as afterschool programming or tutoring). However, few schools collect data on students' nonacademic outcomes, nor do they systematically assess the linkages between nonacademic supports and academic outcomes. Under ESSA, states have begun to think about what data they may require schools to report, to measure these issues through the required "fifth indicator." The majority of states are using chronic absenteeism, which—although it may be the symptom of other underlying problems—may leave a lot of specific information unknown, such as why a student is not attending school regularly.[c] Principals themselves seemed to understand the need to understand more of these specifics to address them. However, they did not necessarily have a system to collect and follow up on that data other than when problems arose and students were discussed in regularly scheduled support meetings.

Finally, some school leaders expressed concerns about sharing data. Specifically, teachers would like to access data on their students, but there are concerns about sharing some student-level data. How can schools/teachers utilize this new resource and knowledge base in a way that is confidential and protective? School leaders are trying to find ways to give teachers access to the data they need. One principal has overcome challenges by requesting that students and parents provide consent at the beginning of the year, so that all necessary school staff can have access to student data. Overall, school leaders are collecting attendance, behavior, and course performance data, and one principal discussed the importance of using the data collected to support requests for additional resources from the district.

Other themes

Five emerging themes were identified from conversations with school leaders and key personnel:

- The importance of having a unified school vision
- Hiring committed staff
- Distributing leadership responsibilities
- Finding and/or being creative with funding
- Considering the potential strengths and challenges of smaller communities

First, principals highlighted the importance of having a school vision and a student-focused sense of common purpose. This seemed to be the driving focus for many school leaders and their staff. It is particularly important when considering the overall changes that must occur to implement strong practices in schools. Almost every individual in the school has a new responsibility to, at the bare minimum, observe changes in student behavior to address consistent and emerging needs. This is also foundational in understanding how unmet needs can contribute to student outcomes and how meeting those needs can help improve outcomes. A critical lesson learned from these conversations is the important role of teachers in supporting implementation of ISS models, in addition to the role of trust among all staff in committing to a common vision and that all parties will follow through with their roles.

Second, there is a need to hire staff who are also committed to the vision and the students. This includes teachers who are equally committed to rethinking the way they engage with students as the school leadership team, and school leadership that extends beyond the traditional administration. In addition to assistant principals, principals appreciated having other leaders take on new initiatives. Staff must understand their common mission of not giving up on kids. One principal noted that everyone in his school knows that they are willing to “do whatever it takes” to serve children. Several principals mentioned that staff must love and care about children and have a mantra of “Never give up on kids.”

Third, funding is an important area where school leaders felt they managed to be successful, but there was still room for growth. School leaders detailed their process of managing school funds and obtaining additional funding from various sources, including federal grants and cost-sharing with partners. One principal said that she hadn’t realized the extent to which running a school is like running a business, and that one must be comfortable with moving money around. One school leader partners with organizations that help address this issue by raising funds themselves to support their work with the

school. Although principals and teachers did not mention the term "braided funding streams," they often mentioned wanting more flexibility about their use of funding, or access to other types of funds. This most often came up around access to healthcare (usually behavioral health) or housing.

Finally, because many principals we spoke with were in urban communities, concerns remain about how ISS might work in rural communities where there are likely fewer options for partnerships, thus presenting a challenge to a core component. With the information available, based on limited interviews in smaller communities, we note that one principal spoke positively about the ability of schools in the district to work together and align services from kindergarten through 12th grade. Specifically, a middle school principal in the rural Midwest spoke of monthly meetings with the other two principals in the district and with the superintendent. These monthly meetings were rooted in decision making across the schools.

Conclusions from the Principal Scan

There are three conclusions to highlight from conversations with principals. First, a good classroom teacher or principal is one who understands when their students need more supports. They can often state exactly what's going on with different students—in class, in the hall, and at home. They know when a student is hungry or sad. They serve are a resource in this process, and empowering them to create a system that works for them and their school is essential.

Second, the five core components of ISS are not only relevant but are truly linked; they require teachers and administrators to potentially do a job they were not trained to do. When one component is missing, the ISS models likely will not work as well as they could. It is helpful to have a model to build from (e.g., City Connects or CIS) because principals are not typically trained to develop community partnerships, design a school-wide needs assessment, integrate new systems into the school, or collect and analyze data—and teachers rarely are. This underscores the importance of hiring a specific person for these jobs. Additionally, principals, teachers, and counselors already have lengthy lists of responsibilities. Many perform these tasks to the best of their abilities and sometimes work well beyond the normal school hours. To implement ISS at or near full implementation represents substantial work for one person. However, it seems important to have a person dedicated to keeping track of community partnerships, systematizing the needs assessment and data collection, and making sure that teachers, parents, and other administrators (as well as students) understand processes and where to turn for help. This person can also train staff members to do these new tasks and feel comfortable with them.

Finally, the fact that schools are not collecting data on nonacademic outcomes or linking data on services and academic outcomes is problematic for two reasons. First, it means that few schools are able to say whether existing nonacademic supports result in better nonacademic outcomes. They are unable to say definitively whether hunger has decreased or health has improved, for example, and whether it matters. The model depends on these nonacademic needs being met—ISS programs are theorized to work because they meet nonacademic needs that otherwise represent barriers to student success. When we do not know whether nonacademic needs have been met, we cannot know whether the program does not work because the model is wrong or because the intermediate step has not been completed. Second, because very few schools and districts use data to connect nonacademic supports to academic outcomes, it is hard to determine whether supports make a difference at a school population level in terms of academic outcomes. For individual students, it can be easier to track improvement, as many schools have weekly or bi-weekly meetings to share information about specific

students they are concerned about and can follow up with. Notes from these meetings are often kept on paper or in a running document online, making it difficult (or impossible) to link the supports and discussion to systematic improvement of outcomes. Also, many students who are struggling, but are of borderline concern or do not act out, may be missed by these meetings. Thus, because many schools are still in the beginning stages of implementing ISS, this is an area where school leaders are still deciding what data to collect, where and how to collect it, and how to share it. The area is ripe for support.

Discussion and Next Steps

In 2014, Child Trends concluded that ISS implementation seemed to be consistent with what researchers had learned about early childhood programs and after-school programs: high-quality implementation is associated with more positive outcomes, while low-quality implementation has the same effect as no program at all.[10] In addition, extended exposure to ISS programs over years seems to be associated with more positive outcomes, suggesting that dosage is also an important factor.[11]

The findings described here do not refute this: high-quality implementation seems essential for positive outcomes. Where outcomes are linked to implementation, we see that poor-quality implementation is often similar to receiving no services. However, it remains unclear what constitutes high-quality implementation. Having concrete information would be valuable to schools seeking to implement an ISS model.

However, this review has emphasized that it can be challenging to include nuanced, complex understandings of implementation rigor into outcomes evaluations. Operationalizing variables inevitably makes them lose some of their depth; however, it is increasingly imperative to clearly define a program model and explore its essential components for outcome improvement—including a rigorous measurement of any mediating variables.

We have also learned the extent of the resources needed to effectively run these programs. For example, in most conversations, principals made it clear that they need a full-time support staff member to help them make these models function effectively. Without this support, models would never get off the ground or be sustained. Considerable financial and time resources are needed to hire someone trained for this position and equipped with the skills to be effective, and data are needed to justify their continued expenditure if ISS models continue to be used.

---

[10] Durlak, J. A. (2010). "The importance of doing well in whatever you do: A commentary on the special section, "Implementation research in early childhood education." *Early Childhood Research Quarterly*, *25*(3), 348-357; Durlak, J. A., & DuPre, E. P. (2008). "Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation." *American Journal of Community* Psychology, *41*(3-4), 327-350; Redd, Z., C. Boccanfuso, et al. (2012). *Expanding Time for Learning Both Inside and Outside the Classroom: A Review of the Evidence Base*. Child Trends, Commissioned by the Wallace Foundation.

[11] Walsh, M. E., Madaus, G. F., Raczek, A. E., Dearing, E., Foley, C., An, C., Lee-St. John, T. J., & Beaton, A. (2014). "A New Model for Student Support in High-Poverty Urban Elementary Schools: Effects on Elementary and Middle School Academic Outcomes." *American Educational Research Journal*, *51*(4), 704-737.

Finally, teachers, principals, and other administrators require patience and supports as they learn to incorporate resources and develop skills for which they have not been trained. There will be a learning curve and—in supporting some of the hardest-to-reach and most vulnerable young people—the balance between allowing teachers and administrators to develop needed skills while not allowing programs that do not work to continue will need to be regularly assessed.

For all three reasons, it is essential that implementation research be a key component of all studies as the research agenda continues to develop in the area of integrated student supports. The ISS model aligns well with child development research and theoretical literature. Nevertheless, evaluations continue to find inconclusive and inconsistent results. This suggests the need to better understand what is going on as programs are rolled out on the ground, and the need to quickly identify what works when principals and staff implement integrated student support models.

Chapter 6: Benefit-Cost Studies

By Jon Belford, Kristin Anderson Moore, and Hannah Lantos

Chapter Overview

Estimating the benefits and costs of social interventions is a recent, rapidly evolving field. These benefit-cost analyses (also referred to as cost-benefit analyses) are an important tool for policymakers, organizations, philanthropists, and other decision-makers in deciding which interventions are the best public investments, given finite financial resources. Several other education interventions have demonstrated that their long-term benefits outweigh program costs, including class size reduction, teacher bonuses in hard-to-staff schools, and early childhood education.[ci] For example, the benefits of early childhood education programs have been found to outweigh costs from anywhere between 2-to-1 for universal pre-Kindergarten in Tulsa, Oklahoma, to between 8.5-to-1 and 16-to-1 for the small, targeted Perry Preschool program.[cii] While it is difficult to directly compare benefit-cost results between studies—since benefits and costs are often calculated differently and benefits can be estimated for different stakeholders (e.g., individuals, taxpayers, or society)—understanding the economic returns of ISS programs will help policymakers decide whether these programs are worth implementing compared to other education interventions.

Four benefit-cost studies of ISS models have been identified to date: City Connects, Communities in Schools, the Children's Aid Society, and Elev8 Oakland. Some benefit-cost work was also done on Harlem Children's Zone in 2008, but the analysis was incomplete and did not utilize the program's more recent and notable impact estimates.[12]

While ISS program benefit-cost analyses share common elements, they differ in the types of benefits and costs that are estimated, as well as the methods used and assumptions made in predicting the economic returns of ISS programs. An important difference between the older benefit-cost analyses and the newer one on City Connects is that the latter includes estimates for community resources under the program. All other benefit-cost analyses assume that community services exist and that students simply utilize them; therefore, they do not include estimates for such costs in their analyses. This is particularly challenging in places where it is assumed that demand will increase with effective support systems

[12] Child Trends researchers were told that HCZ plans on releasing updated benefit-cost estimates in the near future, but release dates were not planned yet.

(which is the assumption in the City Connects paper), or in rural areas where services may not exist or may be too far away to utilize regularly. However, all four studies find that ISS interventions have positive returns to society. The benefits take some time to accrue, but are large enough that they greatly outweigh the costs. The evidence suggests that, for every $1 invested in an ISS model, society will likely gain somewhere between $3 and $15 in benefits.[13]

Benefit-Cost Studies

City Connects. Researchers at the Center for Benefit-Cost Studies in Education (CBCSE) at Columbia University's Teachers College prepared "A Benefit-Cost Analysis of City Connects."[ciii] This study takes several approaches to provide a range of estimated costs of the City Connects program relative to its benefits. It first estimates the costs of City Connects for two school sites during the 2013–2014 school year. The researchers used the ingredients method to identify costs. This method includes all resources used in program implementation, such as personnel, facilities, equipment and materials, other program inputs, and in-kind supports, such as volunteer time. Resources expended are also broken down by stakeholder to determine which costs were incurred by the program, the school, and the parents. All resources are then matched to 2013 Boston prices. The resulting cost estimates represent the opportunity costs of using resources for City Connects that could be used for the best alternative intervention.

The authors then calculate three different cost models: one that includes only the direct costs of City Connects, another that adds the cost of community partner services, and a last one that includes a partial cost for these services. This is unique among these studies, as the remaining benefit-cost analyses assume that community services will be present at no additional cost—a particularly challenging assumption in low-resourced and/or rural areas, or in settings where children with higher needs increase demand beyond current supply. The cost estimate that includes partial costs of community partner services is used by the authors because it is the median estimate, which averages $4,570 per student. Including the costs of community partner services is important because students in City Connects may receive more services from community partners than other students, potentially increasing the resources needed in these other organizations. The study then uses estimates from previous research on the decreased high school dropout rate and increased sixth- to eighth-grade math and ELA test scores of City Connects participants, relative to students from similar schools, to identify program benefits.[civ] The future earnings, health status, crime, and welfare participation of program participants and similar students are then predicted over their lifetimes, based on their predicted educational attainment or achievement; these estimates are then converted into monetary benefits. The average benefits from reduced high school dropouts and increased academic achievement are used as the final benefit estimate.

[13] These benefit-cost estimates were estimated in dollars for different years and different locations. To accurately compare the ratios, they would need to be converted into comparable dollars using the same year and location. However, due to limited information in the analyses, we are unable to calculate the dollars to be comparable.

After discounting future benefits back to kindergarten to account for the time value of money, the study finds that the present value of the social benefits of City Connects participation averages around $13,850 per student. Sensitivity analyses are also conducted to examine the lower and upper bounds of the benefit-cost predictions using different benefit and cost estimates. Overall, the benefit-cost ratio suggests that, for every $1 invested in City Connects, society will gain around $3 in benefits. Additionally, it is estimated that City Connects would hit the break-even point even if the program was only half as effective at increasing educational attainment and academic achievement.

Communities in Schools. "The Economic Impact of Communities in Schools" is a benefit-cost analysis of CIS completed by Economic Modeling Specialists Inc.[cv] The program's impact estimates rely on high school dropout and graduation rate data from a quasi-experimental study that uses CIS performance management data, combined with data from other studies on educational persistence after high school. These estimates are used to predict the benefits of increased disposable income and tax revenue over a 53-year period. Also, because education is correlated with better social outcomes, the authors estimate the monetary benefits of improved health (through less smoking and alcohol use), as well as reduced crime, unemployment, and welfare utilization. These benefits are estimated for both the individual and society.

The costs of CIS are calculated as the amount spent annually by the program to coordinate and provide student services in schools, and include the opportunity costs of students not joining the labor force and schools hosting the intervention. However, the study does not include the cost of community supports that may be provided to CIS participants at higher levels, relative to other students. Both costs and benefits are then discounted to estimate their present value. Finally, sensitivity analyses are conducted to determine how estimates are affected when assumptions change, including the discount rate, length of students' careers, and program impact estimates. The benefit-cost ratio for CIS is estimated to be 11.6. In other words, there is an estimated return of $11.60 for every $1 invested. The investment is estimated to reach a break-even point after nine years.

Children's Aid Society. "Measuring Social Return on Investment for Community Schools – A Practical Guide" was completed in 2013 by The Finance Project.[cvi] The study investigates the education, health, and other benefits to society of community school programs through a case study of Children's Aid Society (CAS) schools in New York City. CAS includes supports for students and their families, both during and outside of school, that focus on stimulating learning and development and strengthening community supports. The services offered by the program are extended school-day learning, medical and mental health services, early childhood education, and parent education.

The study utilized data for one elementary and one secondary school site, collected between 2007 and 2010 by the New York City Department of Education and CAS. Both sites were full-service– that is, they provided all available program supports. Cost data were collected retrospectively for direct program expenses, in-kind services provided, and overhead/administrative costs, mainly using budget data from CAS programs and regular-day school programs. Five schools with children who had similar demographics were used as a comparison group, and cost data for these schools were determined using New York City's Fair Student Funding Allocation to ensure that costs were similar across schools.

Over 40 outcomes were estimated for preschool children, school-aged students, families, and the school community using data from various public sources and databases. Eighteen of these outcomes were "monetized" using various financial proxies to estimate their value in 2010 New York dollars. Next, estimates of benefits that would occur regardless of participation in CAS, which the authors term "deadweight," were subtracted to retain only the benefits that can be attributed to the program. A net present value of benefits was calculated for a period of five years, which underestimates the true benefits of program participation since some of these benefits accrue over a lifetime. The benefit-cost ratio was calculated to be 10.3-to-1 for the elementary school and 14.8-to-1 for the secondary school site. In other words, an investment of $1 returns an estimated $10.30 or $14.80 to society.

Elev8 Oakland. "Oakland Community School Costs and Benefits: Making Dollars and Cents of the Research," prepared by the Bright Research Group, provides estimates of the economic return from the Elev8 Oakland program.[cvii] Elev8 Oakland coordinates school-based programs that provide an integrated system of supports for students at five Oakland middle schools. The programs are run by a local nonprofit called Safe Passages and include summer school, extended learning, health care, and family services.

The study first reports the initial investment, or cost, by Atlantic Philanthropies of $2.5 million annually in the Elev8 Oakland intervention. Next, the value of all services and funds that are leveraged by the middle school sites is calculated. Middle schools can leverage services—including school-based health centers, extended learning, and mental health services—because the sites provide a coordinator and a location for creating partnerships and providing supports. The initial investment is estimated to bring in leveraged resources and services that amount to an additional $3.3 million invested, which totals $5.8 million in costs. The investments are calculated using data from Elev8 Oakland financial records and input from Elev8 stakeholders.

Next, the long-term benefits of the combined initial and leveraged investment are predicted, which amounts to an estimated $25.7 million in benefits. These estimates are based on extrapolations from other research examining the long-term benefits of initiatives similar to Elev8 Oakland's intervention in terms of included components. The study predicts a return to society of $9.96 per $1 spent from the initial investment, but this does not account for all costs to society because it does not include the leveraged investment. However, a more accurate estimate of Elev8 Oakland's economic returns to society is provided, which includes the combined initial and leveraged investment; this benefit-cost estimate is $4.39 per $1 spent. The report also includes the estimated economic return for each Elev8 Oakland program component.

Discussion

All four studies conclude that the benefits of ISS programs outweigh their costs, and sometimes by a large amount. Additionally, the sensitivity analyses included in these studies demonstrate that, even if the estimates may be overly optimistic in some ways, there is almost certainly a positive economic return from the investments made in ISS schools. Benefit-cost studies can also underestimate benefits to society, since it is not possible to monetize all benefits that result from social programs. Regardless,

the positive results hold up even though each study analyzes different programs provided in various settings across the United States. Additionally, the positive conclusions hold up even though each study uses alternative estimation strategies for the costs and benefits of the programs. For example, each study examines different benefits, which may include higher earnings, reduced crime, or improved health, and not all studies include the costs of community partner supports.

The only analysis new to this 2017 report is the benefit-cost analysis of City Connects. The authors of this study make important contributions to estimating the social returns of ISS programs by doing three things carefully and in detail: cost estimation, benefit estimation, and sensitivity testing. First, the authors include a comprehensive analysis of City Connect's costs using the ingredients method. The authors also incorporate an estimate of community organization costs into their cost calculations. The benefit-cost analyses conducted for CIS and CAS assumed these supports and services to already exist, regardless of whether the ISS intervention was implemented, and therefore did not include their costs. This can underestimate program costs because the increased utilization of these services by ISS participants will likely require additional resources. Second, the authors provide program-specific impact and benefit estimates rather than rely on estimates from studies on similar programs, as is done in the Elev8 Oakland study. Finally, the authors conduct a robust set of sensitivity analyses, in which they examine how results change given high- and low-end cost and benefit estimates. This includes using cost estimates both with and without community supports, and benefits derived from either educational achievement (low end) or attainment (high end). Neither the CAS or Elev8 Oakland studies include comprehensive sensitivity analyses. These strengths highlight the evolution and maturation of the field, as a greater focus on the ISS approach is being translated into new and more detailed studies.

There are several ways in which benefit-cost analyses can be further enhanced to improve our understanding of the societal returns of ISS programs. None of these studies estimate the effects of programs using experimental methods in which students are randomly assigned to partake in an intervention, which can be difficult and costly to conduct. Additionally, as with most benefit-cost analyses, these studies rely on predicted future benefits using various external data sources and studies, rather than directly observing program benefits by collecting data on students over time. Estimates of the effects of high school graduation on later life outcomes generally provide a good indication of future benefits, but these estimates are imperfect. For instance, measuring the effects of high school graduation on crime reductions and health is imprecise, making it challenging to calculate the direct savings caused by participation in an ISS program. Moreover, while researchers are becoming increasingly adept at assessing the economic value of preventing crime, smoking, drug use, incarceration, welfare, and unemployment, it is not a perfect science. Even with an experimental study, assumptions must be made about how short-term impacts will alter social and economic trajectories into adulthood. The assumptions made in benefit-cost analyses affect the results of these studies. Therefore, it is extremely valuable to have four studies with differing assumptions that all agree that the benefits of ISS outweigh the costs. These findings represent an important indicator that ISS programs do provide net benefits for society.

Chapter 7: Discussion and Conclusion

By Hannah Lantos and Kristin Anderson Moore

Integrated student supports (ISS) models in schools are growing across the country. The number of schools using specific programs (such as Communities in Schools or City Connects) has grown rapidly in the last decade, but so has the number of schools in which principals do not follow any specific model, simply recognizing the importance of supporting students' nonacademic needs in a structured and systemic way. Along with this growth has come integration with other school frameworks like multi-tiered systems of support (MTSS), and a recognition of the types of supports that various schools might need. In particular, PBIS and ISS (both MTSS frameworks) can be very well-aligned to address both behavioral and out-of-school needs. In many ways, the last decade has seen schools move from an ad hoc application of integrated supports to more systematization; school leaders now have more supports, more models to build, and more evidence to support the importance of removing nonacademic barriers to learning.

While we have learned a lot about ISS models, much remains to be studied. This report has reinforced previous findings, shed light on factors that explain the difficulty of effectively implementing ISS, and highlighted where urgent research questions remain unanswered. To start, we know that these interventions have mostly null (no) or positive results. This is promising: in all studies included in this report, only two outcomes in two studies were negative. This is likely because ISS models are aligned with everything we know from research and theory about child development. For example, ISS models align well with the following theoretical models:

- Whole child
- Ecological model
- Life course perspective
- Child-centered
- Social determinants of health
- Social and emotional learning (SEL)
- Soft skills
- Positive Youth Development (PYD)

This alignment is important. As experimental research evidence is slow to build to conclusive findings, theory backs up specific models and their conceptual underpinnings align with what we know is

important for children: positive, safe settings that support their sense of self within a broader community, and which meet their basic needs while also recognizing the complex interplay of the contexts they live in.

Importantly, some of the most methodologically strong studies find positive impacts. The sweet spot is methodologically rigorous studies combined with rigorously well-implemented programs. Evaluations of City Connects, City Year, Harlem Children's Zone's Promise Academy (HCZ), and CIS in Chicago capture this sweet spot. These models found positive results more consistently than the rest of the studies. City Year, HCZ, and CIS-Chicago were RCTs, while City Connects was a QED with propensity score matching. All had rigorous methodologies and highlighted different components of implementation, which resulted in consistent findings: higher teacher-to-student ratios, fidelity to the specific model, and a focus on specific outcomes that were most important. All have also invested resources in people that continually care and show up. They use school-based coordinators to support students and link them to services—City Connects hires its own employees to do this in each school, City Year uses AmeriCorps City Year volunteers, CIS-Chicago has coordinators in each school, and HCZ's Promise Academy is built around its own model of support. All are clearly based heavily in the theories listed above, with City Connects and HCZ taking more of a public health approach to learning and City Year and CIS focusing on positive youth development and the whole child. These programs were able to not only have impact because they were good programs, but also because the design of their evaluations was rigorous and appropriate to the program and available data. As the field continues to design future studies, it is important to remember: if you are going to invest in an evaluation, it is not sufficient to be a good program.

Even though many different theories from different academic disciplines support what we see in ISS models—community integration, needs assessments, family inclusion, data to assess success, and prevention (among others)—the findings from the evaluations remain mixed, such that many evaluations do not find significant positive effects. Why is this? It seems that there are two overarching questions that remain about why we see such variation across programs and even outcomes. First, what nonacademic outcomes do we expect to improve between receiving support and improved academic outcomes? Are these outcomes improving? Essentially, the first question asks whether our own conceptual model of inputs affecting some intermediary, which then affect a student's ability to study or learn effectively, is correct. However, these studies have generally not examined the nonacademic outcomes they seek to change. These are part of the theory of change for ISS models, but evaluations do not assess them fully or consistently.

Second, if these intermediate factors are, in fact, improving, how does the quality with which programs are implemented affect outcomes? This question is focused on whether there are key components of implementation that matter all the time—either concrete (a single staffer to manage the entire program) or less concrete (a staffer who always greets children with a smile). Additionally, while we assume that quality of program implementation matters, the implementation recipe is not clear. In a time of limited budgets, schools want to know what elements are essential and which are not.

In addition to the many outcome evaluations finding positive impacts, the cost-benefit studies also identify positive returns on investment. However, these highlight a very important conclusion: schools need resources to carry out their tasks—a needs assessment, coordination, data collection, etc. School staff and principals may move forward doing this work out of necessity, but doing it well over time requires dedicated staff. In large schools, more than one staff person may be needed. Without these staff who have time to do the work as their job (and not in addition to teaching), these models are difficult to build and sustain.

Another important finding was about resources available in the community. Most ISS models assume some level of resources for which children can be referred. However, in some communities, there are insufficient resources for referrals. This is a potential Achilles heel for this model, especially in under-resourced communities such as many small, rural towns. Schools vary in the number of services and supports they offer directly to students, but most schools refer students outside the school for at least some services, such as mental health counseling. When these are not readily available in the community, the work becomes challenging.

As the field of ISS continues to develop, the integrated services model has the potential to impact the well-being of hundreds of thousands, if not millions, of children's lives in the United States. It is necessary to invest specifically in a greater understanding of the essential elements under each core component laid out by Child Trends, and how best to implement these essential elements. The children who require these additional supports are enrolled in our nation's schools right now, today; their needs are often large and immediate, and are sometimes urgent. The next step for the ISS field is an urgent one: supporting educators (including teachers, principals, parents, counselors, and other school staff who already work hard to provide America's children with the education and growth necessary to become tomorrow's engaged citizens) to know the best practices and the best ways to support children, how to implement these practices, and where to find the resources to do so.

Appendices

Appendix 1: Program Descriptions of Programs Studied

City Connects is an evidence-based practice that addresses the out-of-school barriers to learning, especially those imposed by poverty. City Connects is a defined, systematic practice that optimizes and transforms traditional school structures and processes aimed at addressing the non-academic needs of students. City Connects collaborates with each teacher and other school staff to systematically identify the strengths and needs of every student across academic, social/emotional/behavioral, health, and family domains. Supports are tracked individually in an electronic database, allowing for outcome evaluation and fidelity of implementation measurement. The program takes a public health approach to education. In a sense, the database acts as an electronic health record for education tracking both struggles and resources provided over time. City Connects, formerly Boston Connects is supported by the Boston College Center for Optimized Student Supports. City Connects is active in preK-8 schools and in one high school pilot program.

Information from: http://www.bc.edu/schools/lsoe/cityconnects/

City Year brings trained AmeriCorps members, who serve for 11 months, into high poverty schools in order to bridge the academic achievement gap. Facilitating a wide variety of school activities, AmeriCorps members provide one-on-one tutoring, and run afterschool programs. The City Year Program uses a whole child, whole school approach in order to provide individualized, one-on-one tutoring to at-risk kids. Kids are selected to receive one-on-one tutoring from a Corps member by a system of early warning indicators known as the ABCs (A stands for poor attendance, B stands for disruptive behavior, and C stands for course failure). This approach is based on research from Johns Hopkins, which has found that a student who exhibits just one of these signs, in as early as sixth grade, has a 25% chance of graduating from high school. On the other hand, a student who is on track to graduate in the tenth grade has a 75% chance of graduating from high school. City Year was founded in 1988 and serves students from third grade through ninth grade.

Information from: https://www.cityyear.org/what-we-do/our-approach

---

Comer School Development Program was created in 1968 by Dr. James P. Comer and his colleagues at the Yale Child Study Center. The School Development Program (SDP) is the first reported school intervention program in which the test scores, behavior, and attendance of poor and/or socially marginalized students improved dramatically. Also, it was the first intervention in which the application of child and adolescent development principles was used school-wide to create interactions and/or relationships that prepared students to learn and to begin to take responsibility for their own learning; and enabled teachers, school staff and administrators to support student personal development and learning. The SDP aims to facilitate student growth along six developmental pathways needed for school success: social-interactive, psycho-emotional, ethical, cognitive, linguistic, and physical.

Information from: http://medicine.yale.edu/childstudy/comer/index.aspx

Communities in Schools works within the public-school and charter systems, determining student needs and establishing relationships with local businesses, social service agencies, health care providers, and parent and volunteer organizations to provide embed needed resources within schools. Communities in Schools (CIS) aims to surround students in a community of support, empowering them to stay in school and achieve in life. The CIS network has been in operation for more than 30 years and is made up of 200 local affiliates nation-wide serving the lowest performing schools and students most vulnerable of dropping out.

Information from: http://www.communitiesinschools.org/

Diplomas Now strives to improve the academic outcomes of the most at-risk students at a given school by making sure each at-risk student's academic progress is monitored by a caring adult. Diplomas Now is a collaborative partnership between Johns Hopkins University's Talent Development Program, City Year, and Communities in Schools. The Diplomas Now team works directly with the school to develop a set of goals, based on grades, behavior, and attendance, for each of its struggling students. Furthermore, the Diplomas Now team uses Early Warning Indicators (EWI) to identify the students who are struggling most of all in school. A plan involving EWI meetings to review each student's progress is then put in place to help these students succeed. Diplomas Now was founded in 1994 and serves students in both middle and high school.

Information from: http://diplomasnow.org/about/

The Harlem Children's Zone (HCZ) seeks to end the cycle of intergenerational poverty in Harlem, New York through a model that provides both comprehensive supports to families from birth through college graduation, and programing that involve families, various social services, and health programs. In other words, HCZ seeks to provide support in three domains: education, family and community, and health. Education programming ranges from the Baby College program for new and expectant parents to a college preparatory program. Family services are centered around various community centers and family programming such as, preventative services. Health services include programming that focuses on healthy eating and physical fitness. The HCZ charter schools – the Promise Academies – are what is studied in the evaluations included in this report. The HCZ was founded in 1970 and serves children from birth through college graduation. This report includes studies exploring the impact of elementary, middle, and high school participation in HCZ.

Information about Harlem Children's Zone broadly from: http://hcz.org/our-programs/

Information about the Promise Academy Charter Schools from: http://hcz.org/our-programs/promise-academy-charter-schools/.

Say Yes to Education is a national non-profit committed to dramatically increasing high school and college graduation rates for our nation's inner-city youth. They provide comprehensive supports, including the promise of a tuition scholarship, aligned with what research indicates is needed to enable every child to achieve his or her potential. Say Yes' promise and supports begin when a child enters kindergarten and continue through college graduation. Say Yes partners with every sector of the community from government organizations, the school district, and higher education institutions to community-based organizations, businesses, and faith-based organizations to ensure a collaborative effort is made to dramatically increase high school and college graduation rates, as well as create a citywide transformation.

Information from: http://www.sayyestoeducation.org/

The Talent Development Model seeks to improve dropout rates and improve academic outcomes in low-performing schools across the United States. At the time this report was published, the Talent Development Model was being implemented at 33 schools within 12 different states. Overall, the Talent Development Model consists of five main components: small learning communities, a curricula leading to students participating in advanced English and math classes, extra help sessions for academic work, staff professional development activities, and parent involvement and community involvement, which seeks to promote both career and college readiness. The evaluation used in this report was produced by MDRC and focuses on the implementation of the Talent Development Program in Philadelphia and the progress of 20 ninth grade cohorts in the Philadelphia City Schools. The Talent Development Model was founded in 1998 and serves students in both middle school and high school.

Information from: http://www.mdrc.org/publication/talent-development-high-school-model

Appendix 2: Detailed Results Table for Academic Outcomes

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
CIS Chicago	Reading Proficiency	All students	+	4.3 (PP)	Results on page 5. This one does not specify grades.
	Math Proficiency	All students	+	3.2 (PP)	Results on page 5.
	Attendance	All students	0	N.S.	
CIS Austin	GPA	Incoming 9th graders	+	0.38 (ES)	Significant differences were only found from baseline to Y1; Y1 to Y2 and baseline to Y2 - no difference
	Math scores	Incoming 9th graders	0	N.S.	
	Reading scores	Incoming 9th graders	0	N.S.	
	Credit Completion	Incoming 9th graders	+	0.38 (ES)	Significant differences were only found from baseline to Y1; from Y1 to Y2 and baseline to Y2 there were no differences
	Attendance	Incoming 9th graders	+	0.45 (ES)	Significant differences were only found from baseline to Y1; from Y1 to

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
					Y2 and baseline to Y2 there were no differences
	Dropout rates	Incoming 9th graders	0	N.S.	
CIS Jacksonville	GPA	6th grade cohorts	0	N.S.	Two cohorts, study begins for participants during their sixth-grade year
	Reading Scores	6th grade cohorts	0	0.26 (ES)	Statistically significant only from baseline to Y1; from Y1 to Y2 and baseline to Y2 there were no differences
	Math Scores	6th grade cohorts	0	N.S.	
	Attendance	6th grade cohorts	0	N.S.	
CIS Wichita	GPA	10th grade cohorts	0	N.S.	
	Credit Completion	10th grade cohorts	+	0.47 (ES)	Page 20, there's a footnote that indicates p-value, but it's not marked in the table.
	Math Scores	10th grade cohorts - year 2	+	0.55 (ES)	Statistically significant from Y1 to Y2 when the cohort was 11th grade

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
	Reading Scores	10th grade cohorts	0	N.S.	
	Attendance	10th grade cohorts - year 2	+	0.72 (ES)	Statistically significant from Y1 to Y2 when the cohort was 11th grade
CIS Year 2 RCT	Chronic Absenteeism	Middle and High School Students	0	N.S.	
	Attendance	Elementary school students	0	N.S.	It's not clear if the improvement was significant; there was indication of positive outcomes for Elementary School students in the whole school study report but no numbers
	Credit Completion/Grades	High school students	0	N.S.	
City Connects 2016	Math Test Scores (Stanford Achievement Test)	5th grade	+	0.21 (ES)	This was for students who had ever been in a City Connects school with within-school fixed effects.

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
	Math Test Scores (Stanford Achievement Test)	5th grade	+	0.17 (ES)	This was for students who had been in only one year of a City Connects school with within-school fixed effects.
	Math Test Scores (Stanford Achievement Test)	5th grade	+	0.21 (ES)	This was for students who had been in more than one year of a City Connects school with within-school fixed effects.
	Math Test Scores (Stanford Achievement Test)	5th grade	+	0.21 (ES)	This was for students who had ever been in a City Connects school with within country of origin fixed effects.
	Math Test Scores (Stanford Achievement Test)	5th grade	+	0.16 (ES)	This was for students who had been in only one year of a City Connects school with within country of origin fixed effects.
	Math Test Scores (Stanford Achievement Test)	5th grade	+	0.22 (ES)	This was for students who had been in more than one year of a City Connects school within country of origin fixed effects.

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
	Reading Test Scores (Stanford Achievement Test)	5th grade	+	0.16 (ES)	This was for students who had ever been in a City Connects school with within-school fixed effects.
	Reading Test Scores (Stanford Achievement Test)	5th grade	+	null	This was for students who had been in only one year of a City Connects school with within-school fixed effects.
	Reading Test Scores (Stanford Achievement Test)	5th grade	+	0.17 (ES)	This was for students who had been in more than one year of a City Connects school with within-school fixed effects.
	Reading Test Scores (Stanford Achievement Test)	5th grade	+	0.17 (ES)	This was for students who had ever been in a City Connects school with within country of origin fixed effects.
	Reading Test Scores (Stanford Achievement Test)	5th grade	+	null	This was for students who had been in only one year of a City Connects school with within country of origin fixed effects.

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
	Reading Test Scores (Stanford Achievement Test)	5th grade	+	0.18 (ES)	This was for students who had been in more than one year of a City Connects school within country of origin fixed effects.
City Connects 2014	ELA report card scores	3rd grade – 7th grade	0	N.S.	Sample size too small for 8th grade
	Writing report card scores	3rd grade – 5th grade	0	N.S.	Significant results found when analyzing dosage (years in program) but not generally
	Math report card scores	3rd, 4th, 6th, 7th grades	0	N.S.	
	Math report card scores	5th grade	+	0.16 (ES)	
	Overall report card scores	6th and 7th grade	0	N.S.	
	ELA State scores	3rd grade - 5th grade	0	N.S.	
	ELA State scores	6th grade	+	0.14 (ES)	
	ELA State scores	7th grade	0	N.S.	
	Math State scores	6th grade	+	0.14 (ES)	
	Math State scores	7th grade	+	0.21 (ES)	

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
	Report Card GPAs	8th Grade	+		Discussion on page 727 and effect sizes between 0.34 and 0.54 are mentioned although 8th grade is not specified.
City Connects - Summary Report 2008 - 2009	Reading Report Card scores	3rd grade	+		Figures 11-13 illustrate these findings (p16-17) and the text states that there are statistically significant differences but does not present a statistical number.
	Reading Report Card scores	4th grade	0	N.S.	This is the only outcome that the text states is insignificant (top of page 16).
	Reading Report Card scores	5th grade	+		Figures 11-13 illustrate these findings (p16-17) and the text states that there are statistically significant differences but does not present a statistical number.
	Math report card scores	3rd grade	+		Figures 11-13 illustrate these findings (p16-17) and the text states that there are statistically significant differences but

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
					does not present a statistical number.
	Math report card scores	4th grade	+		Figures 11-13 illustrate these findings (p16-17) and the text states that there are statistically significant differences but does not present a statistical number.
	Math report card scores	5th grade	+		Figures 11-13 illustrate these findings (p16-17) and the text states that there are statistically significant differences but does not present a statistical number.
	Writing report card scores	3rd grade	+		Figures 11-13 illustrate these findings (p16-17) and the text states that there are statistically significant differences but does not present a statistical number.
	Writing report card scores	4th grade	+		Figures 11-13 illustrate these findings (p16-17) and the text states that there are statistically significant differences but does not present a statistical number.

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
	Writing report card scores	5th grade	+		Figures 11-13 illustrate these findings (p16-17) and the text states that there are statistically significant differences but does not present a statistical number.
	MCAS ELA (dose: one year)	3rd Grade	-	-0.06 (ES)	Table 3 on 19
	MCAS ELA (dosage: whole time in City Connect School)	3rd Grade	-	-0.07 (ES)	Table 3 on 19
	MCAS ELA (dose)	4th Grade	+	0.05 (ES)	Table 3 on 19
	MCAS ELA (dosage)	4th Grade	0	N.S.	Table 3 on 19
	MCAS ELA (dose)	5th Grade	0	N.S.	Table 3 on 19
	MCAS ELA (dosage)	5th Grade	0	N.S.	Table 3 on 19
	MCAS ELA (dose)	6th Grade	0	N.S.	Table 3 on 19
	MCAS ELA (dosage)	6th Grade	+	0.13 (ES)	Table 3 on 19
	MCAS ELA (dose)	7th Grade	0	N.S.	Table 3 on 19
	MCAS ELA (dosage)	7th Grade	+	0.11 (ES)	Table 3 on 19
	MCAS ELA (dose)	8th Grade	0	N.S.	Table 3 on 19

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
	MCAS ELA (dosage)	8th Grade	0	N.S.	Table 3 on 19
	MCAS Math (dose: one year)	3rd Grade	0	N.S.	Table 3 on 19
	MCAS Math (dosage: whole time in City Connect School)	3rd Grade	+	0.05 (ES)	Table 3 on 19
	MCAS Math (dose)	4th Grade	0	N.S.	Table 3 on 19
	MCAS Math (dosage)	4th Grade	0	N.S.	Table 3 on 19
	MCAS Math (dose)	5th Grade	0	N.S.	Table 3 on 19
	MCAS Math (dosage)	5th Grade	0	N.S.	Table 3 on 19
	MCAS Math (dose)	6th Grade	+	0.09 (ES)	Table 3 on 19
	MCAS Math (dosage)	6th Grade	+	0.14 (ES)	Table 3 on 19
	MCAS Math (dose)	7th Grade	+	0.1 (ES)	Table 3 on 19
	MCAS Math (dosage)	7th Grade	+	0.17 (ES)	Table 3 on 19
	MCAS Math (dose)	8th Grade	0	N.S.	Table 3 on 19

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
	MCAS Math (dosage)	8th Grade	+	0.09 (ES)	Table 3 on 19
City Connects - Annual Report 2010	Reading Report Card scores	3rd grade	+	0.5 (MD)	Graph of effect sizes on page 17 and the difference in means is 0.5 points higher for City Connects' students (on a scale that goes from 3-12).
	Reading Report Card scores	4th grade	+	0.41 (MD)	Graph of effect sizes on page 17 and the difference in means is 0.41 points higher for City Connects' students (on a scale that goes from 3-12).
	Reading Report Card scores	5th grade	+	0.3 (ES)	Reported in the text on page 17. Graph of effect sizes is also on page 17 and the difference in means is 0.4 points higher for City Connects' students (on a scale that goes from 3-12).
	Writing Report Card Scores	3rd grade	0	N.S.	Writing Report Card Scores for 3rd and 5th grade were statistically significant when unadjusted but did not

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
					remain significant in adjusted models. Difference in means is 0.4 on a scale that goes from 4-16.
	Writing Report Card Scores	4th grade	+	0.38 (MD)	Graph of effect sizes on page 17. Difference in means is 0.38 on a scale that goes from 4-16.
	Writing Report Card Scores	5th grade	0	N.S.	Writing Report Card Scores for 3rd and 5th grade were statistically significant when unadjusted but did not remain significant in adjusted models. Difference in means is 0.38 on a scale that goes from 4-16.
	Math Report Card Scores	3rd grade	+	0.37 (MD)	Graph of effect sizes on page 17. Difference in means is 0.37 on scale that goes from 3-12.
	Math Report Card Scores	4th grade	+	0.49 (MD)	Graph of effect sizes on page 17. Difference in means is 0.49 on scale that goes from 3-12.
	Math Report Card Scores	5th grade	+	0.5 (MD)	Graph of effect sizes on page 17. Difference in

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
					means is 0.5 on scale that goes from 3-12.
	ELA State Test Scores	3rd grade - 5th grade	0	N.S.	The mean difference was 0.04 points in grade 3, 0.02 points in grade 4, and 0.01 points in grade 5 (show in Table 5 on page 22.) State test score graph page 24.
	ELA State Test Scores	6th grade	+	0.15 (MD)	The mean difference was 0.15 points better for the City Connects Students (table 8 on page 23). State test score graph of effect sizes page 24. (p<.10)
	ELA State Test Scores	7th grade	+	0.16 (MD)	The mean difference is 0.16 (table 8 on page 23). State test score graph of effect sizes page 24.
	ELA State Test Scores	8th grade	+	0.17 (MD)	The mean difference is 0.17 (table 8 on page 23). State test score graph of effect sizes page 24.
	Math State Test scores	3rd Grade	+	N.S.	The 3rd grade mean difference is significant in the unadjusted models but loses significance in the adjusted models

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
					(tables 5 and 7). State test score graph of effect sizes page 24.
	Math State Test scores	4th and 5th grade	0	N.S.	State test score graph of effect sizes page 24.
	Math State Test scores	6th grade	+	0.2 (MD)	The mean difference is 0.2 (table 8 on page 23). State test score graph of effect sizes page 24.
	Math State Test scores	7th grade	+	0.2 (MD)	The mean difference is 0.2 (table 8 on page 23). State test score graph of effect sizes page 24.
	Math State Test scores	8th grade	+	0.35	The mean difference is 0.35 (table 8 on page 23). State test score graph of effect sizes page 24.
City Connects - Progress Report 2012	Overall GPA	6th grade	+	0.21 (MD)	Graph on page 21 of effect sizes.
	Overall GPA	7th grade	+	0.19 (MD)	Graph on page 21 of effect sizes.
	Overall GPA	8th grade	+	0.19 (MD)	Graph on page 21 of effect sizes.
	ELA GPA	6th grade	+	0.07 (MD)	Graph on page 21 of effect sizes.

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
	ELA GPA	7th grade	+	0.17 (MD)	Graph on page 21 of effect sizes.
	ELA GPA	8th grade	0	N.S.	Graph on page 21 of effect sizes.
	Math GPA	6th grade	+	0.12 (MD)	Graph on page 21 of effect sizes.
	Math GPA	7th grade	+	0.05 (MD)	Graph on page 21 of effect sizes.
	Math GPA	8th grade	+	0.16 (MD)	Graph on page 21 of effect sizes.
	Drop-out	High School Students	+	0.54 (OR) 2.1 (PP)	Graph on page 24. Decrease in dropout equals an improvement; the 2.1 percentage point differences means a 46% lower odds of dropping out between 8th and 12th grade
City Year	ELA Assessments 2011-2012		+	1.8 (OR)	Exhibit 10 on page 18
	Math Assessments 2011-2012		+	1.7 (OR)	Exhibit 10 on page 18
	ELA Assessments 2012-2013		+	2.0 (OR)	Exhibit 10 on page 18

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
	Math Assessments 2012-2013		+	2.9 (OR)	Exhibit 10 on page 18
	ELA Assessments 2013-2014		+	1.9 (OR)	Exhibit 10 on page 18
	Math Assessments 2013-2014		0	N.S.	Exhibit 10 on page 18
Diplomas Now	Attendance	Middle school and high school students	0	N.S.	Figure ES 1, Page ES-7 and Table 4.1 page 38
	Course Performance/Grades	Middle school and high school students	0	N.S.	Figure ES 1, Page ES-7 and Table 4.1 page 38
Harlem Children's Zone Promise Academy (Dobbie and Fryer – High school)	Woodcock Johnson Math	High School students	+	0.281 (beta)	Table 4, page 1006
	Woodcock Johnson Reading	High School students	0	N.S.	Table 4, page 1006
	State Tests Passage rate	High School students	+	1.228 (beta)	Measures the total number of exams passed. Table 4, page 1006
	State Test Scores	High School Students	+	0.293 (beta)	Average score on the living environment, global history, and

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
					integrated algebra Regents exams. Table 4, page 1006.
	College enrollment	High School students	+	N.S.	Table 4, page 1006
	Achievement Index	High School Students	+	0.279 (beta)	An index that combines all four of the individual achievement measures (math, reading, Regents passes, Regents scores). Table 4, page 1006
Harlem Children's Zone Promise Academy (Dobbie and Fryer – Middle and elementary School)	On Grade Level	Elementary School Students	0	N.S.	Table 5, page 173 – this is the linear regression coefficient for the 2SLS regression.
	Math (standardized test scores)	Elementary School Students	0	N.S.	Table 5, page 173 – this is the linear regression coefficient for the 2SLS regression.
	ELA (standardized test scores)	Elementary School Students	0	N.S.	Table 5, page 173 – this is the linear regression coefficient for the 2SLS regression.
	Absences	Elementary School Students	+	-2.412	Table 5, page 173 – this is the linear regression coefficient for the 2SLS regression. A decrease in

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
					absences is a positive overall effect.
	Math test scores	Middle school students	+	0.229 (beta)	Table 3, page 170 – this is the linear regression coefficient for the 2SLS regression. They also report results for subsets of students in Table 4 on page 172.
	ELA test scores	Middle School Students	0	N.S.	Table 3, page 170 – this is the linear regression coefficient for the 2SLS regression – which approximates causality more than other estimates.
	Absences	Middle School Students	+	-2.199	Table 3, page 170 – this is the linear regression coefficient for the 2SLS regression. A decrease in absences is a positive overall effect.
	On Grade Level	Middle School Students	0	N.S.	Table 3, page 170 – this is the linear regression coefficient for the 2SLS regression.
Talent Development	Algebra credit earned	First time 9th graders	+	24.5 (PP)	Figure ES.1 on page ES 5. Note these are the combined results for all

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
					three cohorts. Table 6 also disaggregates the results by cohort and these are not shown here.
	Course credits earned	First time 9th graders	+	0.25 (ES)	Table 6 on Page 50 also shows results for 5 or more credits and credits in English and math. This is a 0.67 difference.
	Basic academic curriculum completed	First time 9th graders	+	8.2 (PP)	Figure ES.1 on page ES 5.
	Promoted to 10th grade	First time 9th graders	+	8 (PP)	Figure ES.1 on page ES 5.
	Promoted to 11th grade	First time 9th graders	+	6.5 (PP)	Figure ES.1 on page ES 5.
	Attendance Rate	First time 9th graders	+	0.17 (ES)	Table B.1 on page 101 and Figure ES.1 on page ES 5. There are also graphs of rates per cohort in Figure 2 on page 48. In Table 6 on page 50 this is described as a 5.1 percentage point difference.

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
Say Yes	Attendance Rate	Elementary students (K-5)	0	N.S.	Page 12
	Course Grades	1st and 2nd grade	+		Students had higher grades in math, science, reading, and writing in 1st and 2nd grade. Page 13.
	Course Grades	3rd and 4th grade	0	N.S.	Page 13
	Course Grades	All years	0	N.S.	Generally differences in grades were rare. Figure 9, page 42.
	Terra Nova Scale Scores	3rd grade	+		Page 13
	Terra Nova Scale Scores	Elementary students (4th and 5th)	0	N.S.	Page 13 and Figure 6 on Page 36
Comer Chicago	Reading NCE Test Scores	High school students	+	1.38 (beta)	Statistical difference shown on graph on page 588. Table 10 on page 575 also shows that though there was no difference in means for the Comer participants there was a positive difference in slopes. This

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
					slope was 1.38 at the school level and 1.41 at the individual level and significant. This meant that Comer students were improving their scores more rapidly.
	Math NCE Test Scores	High school students	+	0.95 (beta)	Statistical difference on graph on page 587. Table 10 on page 575 also shows that though there was no difference in means for the Comer participants there was a positive difference in slopes. This slope was 0.95 at the school level and 0.91 at the individual level and significant. This meant that Comer students were improving their scores more rapidly.
Comer PG County	GPA	Middle school students	0	N.S.	Table 5 on page 571 has results for the academic outcomes across all cohorts in both grades 7 and 8 but none of the differences are statistically significant.

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
	Absenteeism	Middle school students	0	N.S.	Table 5 on page 571 has results for the academic outcomes across all cohorts in both grades 7 and 8 but none of the differences are statistically significant.
	Math scores	Middle school students	0	N.S.	Table 5 on page 571 has results for the academic outcomes across all cohorts in both grades 7 and 8 but none of the differences are statistically significant.
CIS - QED	Graduation	High school students	0	N.S.	There are significant increases in both CIS and comparison schools but no significant difference between the two. However, using a p-value of 0.1, the difference is significant and positive (p=0.088). We use a cutoff of P=0.05 meaning we report it as null overall. Depending on which group of comparison schools, the amount (and statistical significance) of that

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
					increase differs. Effect size - CIS: +15.58 percentage points Comparison schools: +8.08 percentage points (at the end of year 3)
	Drop-out Rate	High school students	0	N.S.	There are significant decreases in both CIS and comparison schools but no significant difference between the two (p=0.211). Effect size - CIS: -3.8 percentage points Comparison Schools: -2.3 percentage points (at the end of year 3)
	Attendance rate (average daily attendance)	High school students	0	N.S.	There are significant increases in both CIS and comparison schools but no significant difference between the two (p=0.814). Effect size - CIS: 0.88 Comparison schools: 0.76
	ELA state test scores (z scores)	High school students	0	N.S.	There are significant increases in both CIS and comparison schools but no significant difference between the two

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
					however they are significant at the 10% level (p=0.085). Effect size - CIS: 0.15 Comparison schools: 0.04
	Math state test scores (z scores)	High school students	0	N.S.	There are significant increases in both CIS and comparison schools but no significant difference between the two (p=0.322). Effect size - CIS: 0.16 Comparison schools: 0.09
	Attendance (average daily attendance)	middle school students	0	N.S.	There were increases in both CIS and comparison schools although only the comparison trend is significant. Additionally, there are no statistical differences between the two (P=0.853). Effect size - CIS: 0.41 Comparison schools: 0.35
	ELA state test scores (z scores)	middle school students	0	N.S.	There are significant increases in the comparison schools but very minimal ones in the CIS schools although there are no significant difference between the

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
					two. However, using a p-value of 0.1, the difference is significant and negative (p=0.061) meaning the comparison schools performed better. Effect size - CI: 0.00 Comparison schools: 0.11
	Math state test scores (z scores)	middle school students	0	N.S.	There is no change in CIS schools and a significant increase in comparison schools though the difference between the two is not significant (0.194). Effect size - CIS: -0.01 Comparison schools: 0.09
	Attendance (average daily attendance)	elementary school students	+	See note	There are significant decreases in both CIS and comparison schools as well as significant difference between the two (P=0.030). Effect size - CIS: .61 percentage points Comparison Schools: .2 percentage points (at the end of year 3)

Program	Outcomes	For whom?	Overall	Measure of Statistical Difference	Notes
Key for measures of statistical difference: PP= percentage point difference; ES= effect size; OR= odds ratio; MD= mean difference; beta=beta coefficient in a linear regression model					
	ELA state test scores (z-scores)	elementary school students	0	N.S.	There were significant increases for the CIS schools and not for the comparison schools though they were not statistically different from each other (P=0.591). Effect size - CIS: 0.10 Comparison schools: 0.06
	Math state test scores (z scores)	elementary school students	0	N.S.	There were non-significant decreases for the CIS schools and non-significant increases for the comparison schools though they were not statistically different from each other (P=0.412). Effect size - CIS: -0.05 Comparison schools: 0.03

Appendix 3: Detailed Results Table for non-Academic Outcomes

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
RCTs 2017	Communities in Schools: Chicago	There were no non-academic outcomes included in this study.			
	Communities in Schools: Year 2 Impact Findings	Outcomes for Students as Individuals			
		School Attachment	YES	+	Students enrolled in CIS felt more happy, safe, and a part of school and were more engaged in school (P=0.020). They also were more likely to have a positive valuation of education (P=0.001).
		Behavior Problems	YES	null	They measured behavior with attendance and suspensions but because we report attendance in the academic outcomes section, we only report on suspensions here. Case managed students were more likely to be suspended (effect size 0.11) but this difference was only marginally significant (at the 10% level, P=0.051). For high and moderate-risk students there was a significant different (effect size: 0.25 and P=0.049).
		Socio-emotional Development	YES	+	The items in the educational-attitudes scale ask about students' own perceptions of whether they do well at school, plan their work, persist with homework and schoolwork, give up easily, or have trouble figuring out answers in school. Case managed students scored higher on the scale overall (P=0.037).

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
		Student Health and Well-being	NO		
		Student Outcomes Measured in Families			
		Academic support at home	NO		
		Use of positive parenting techniques	NO		
		Presence of positive parent-child relationships	YES	+	Case managed students reported more caring relationships with adults at home (P=0.001).
		Outcomes Within the School Environment			
		Postive School Climate	YES	+	Case managed students were able to form more trusting and supportive friendships than non-case managed students (P=0.002)
		Student-Teacher and/or Staff Relationships	YES	+	Case managed students were able to form more relationships with caring adults at school (P=0.004)
	Diplomas Now	Outcomes for Students as Individuals			
		School Attachment	YES	null	No significant impact on the students' engagement with school.

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
		Behavior Measures	YES	null	No significant impact on students' school behaviors (i.e. percentage of days suspended or expelled).
		Socio-emotional Development	YES	null	No significant impact on students' self-perceptions (i.e. confidence, self-worth, effort, and persistence).
		Student Health and Well-being	NO		
		Student Outcomes Measured in Families			
		Academic support at home	NO		
		Use of positive parenting techniques	NO		
		Presence of positive parent-child relationships	YES	null	Participating in DN was associated with generally positive trends in parent and community involvement in the school but none of these were statistically significant. There is one negative (though not statistically sig.) association; the number of times per month parents volunteered in the classroom was lower in DN schools than non-DN schools.
		Outcomes Within the School Environment			
		Positive School Climate	YES	null	Positive impact on teachers' perceptions of school climate, which was marginally significant (at the 10%, P=0.096).

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
		Student-Teacher and/or Staff Relationships	YES	null & +	Students enrolled in DN were more likely to report a positive relationship with a non-teacher staff member P=0.011), but there were no differences in students perceptions of their relationship with teachers and staff as compared to the non-DN students (P=0.316).
	Harlem Children's Zone's Promise Academy (High School Outcomes)	Outcomes for Students as Individuals			
		School Attachment	NO		
		Behavior Problems	YES	null & -	Harlem Children's Zone's Promise Academy participants were not significantly less likely to engage in criminal behavior (though the trend was negative) but they were significantly less likely to engage in an index of risky behaviors (measured as: ever pregnant, ever incarcerated, self-reported drug and alcohol use, and self-reported criminal behavior).
		Socio-emotional Development	YES	null & -	Lottery winners report lower levels of grit than those who lost the lottery for Harlem Children's Zone. All three measures of non-cognitive skills (self-esteem, grit and locus of control) have a negative trend though grit is the only one that is significant.
		Student Health and Well-being	YES	null & +	Participant females were less likely to be pregnant and participant males were less likely to be incarcerated. There was little impact on self-reported health, self-reported drug and alcohol use, or self-reported criminal behavior of participating in the program.

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
		Student Outcomes Measured in Families			
		Academic support at home	NO		The study reports levels of parent engagement (to academic feedback, behavioral feedback, and regular feedback) relative to other schools in NYC but this is not an outcome - just a baseline comparison.
		Use of positive parenting techniques	NO		Harlem Children's Zone's Promise Academy does focus extensively on parenting but they do not report results of those programs in this study although they are described as very important.
		Presence of positive parent-child relationships	NO		
		Outcomes Within the School Environment			
		Positive School Climate	NO		
		Student-Teacher and/or Staff Relationships	NO		
	Harlem Children's Zone's Promise Academy (Middle School Outcomes)	There were no non-academic outcomes included in this study.			

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
RCTS 2014	Communities in Schools: Austin	Outcomes for Students as Individuals			
		School Attachment	YES	null	There were no significant differences found on the scale about community and school involvement which measured school attachment and engagement.
		Behavior Problems	YES	null	There were no significant differences found in the disciplinary referral data (for a number of different behavioral outcomes) thought the tend was down in both the treatment and control groups.
		Socio-emotional Development	YES	null	There were no significant differences found between treatment and control in terms of personal responsibility, self-worth, or future aspirations.
		Student Health and Well-being	NO		
		Student Outcomes Measured in Families			
		Academic support at home	YES	null	No significant differences were found between treatment and control in terms of relationships with parents or parental involvement.
		Use of positive parenting techniques	NO		
		Presence of positive parent-child relationships	YES	null	No significant differences were found between treatment and control in terms of relationships with parents or parental involvement.

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
		Outcomes Within the School Environment			
		Positive School Climate	NO		They discuss activities to improve school climate but do not report these as outcomes.
		Student-Teacher and/or Staff Relationships	NO		
	Communities in Schools: Jacksonville	Outcomes for Students as Individuals			
		School Attachment	YES	null	There were no significant differences found on the scale about community and school involvement which measured school attachment and engagement.
		Behavior Problems	YES	null	No significant differences were found between treatment and control in terms of behavioral referrals or suspensions (either in or out of school).
		Socio-emotional Development	YES	null	Results for personal responsibility were marginally significant (P=0.051) at the year 1 mark but not at year 2. They studied changes in other indicators but don't report significance.
		Student Health and Well-being			
		Student Outcomes Measured in Families			
		Academic support at home	YES	null	There were no significant differences in family relationships and parental involvement.

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
		Use of positive parenting techniques			
		Presence of positive parent-child relationships	YES	null	There were no significant differences in family relationships and parental involvement.
		Outcomes Within the School Environment			
		Positive School Climate	NO		
		Student-Teacher and/or Staff Relationships	NO		
	Communities in Schools: Wichita	Outcomes for Students as Individuals			
		School Attachment	YES	null	There were no significant differences found on the scale about community and school involvement which measured school attachment and engagement.
		Behavior Problems	YES	null	There were no significant differences in behavioral measures between the treatment and control schools.
		Socio-emotional Development	YES	null	There were no significant differences found between treatment and control in terms of personal responsibility, self-worth, or future aspirations.

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
		Student Health and Well-being			
		Student Outcomes Measured in Families			
		Academic support at home	YES	null	There were no significant differences in family relationships and parental involvement.
		Use of positive parenting techniques	NO		
		Presence of positive parent-child relationships	YES	null	There were no significant differences in family relationships and parental involvement.
		Outcomes Within the School Environment			
		Positive School Climate	NO		
		Student-Teacher and/or Staff Relationships	NO		
	Comer: Prince George's County, MD	Outcomes for Students as Individuals			
		School Attachment	YES	null	There was marginal significance (P=0.087) for pride in school in the Comer schools. Attachment to the school and pleasure in attending the school were not significantly associated with Comer.

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
		Behavior Problems	YES	-	Students in Comer schools were less likely to engage in petty misbehaviors than students in non-Comer schools (P=0.03) in 7th grade for cohort 1. Students in the same cohort were marginally less likely to use tobacco (P=0.07). Students from the same cohort were marginally less likely to use marijuana in 8th grade (P=0.06). Overall there were no other significant differences in misbehavior or substance use.
		Socio-emotional Development	YES	null	No significant differences in self-efficacy, satisfaction with self, or anger control in the Comer and non-Comer schools across grade or cohort. For anger control, Comer students were marginally more likely to report better anger control at the beginning of 7th grade in cohort 2 (P=0.10).
		Student Health and Well-being	YES	null	There were no significant differences found in rates of depression across Comer and non-Comer schools in any of the grades or cohorts.
		Student Outcomes Measured in Families			
		Academic support at home	YES	null	Parents were marginally more likely (P<.1) to attend social, volunteering, and/or administrative meetings at the school. No evidence that they helped more with homework.
		Use of positive parenting techniques	YES	null	No evidence that parents communicated more with their children about school.

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
		Presence of positive parent-child relationships	NO		
		Outcomes Within the School Environment			
		Positive School Climate	YES	null	No clear effects on students' or staffs' perceptions of school climate after participating in Comer. Also no evidence of there being a friendly student climate though there were positive trends in terms of safety, positive ethnic group interaction, and positive problem solving over time in both Comer and non-Comer schools. The one item that was strongly significant was the adequacy of rules about misbehavior where the Comer schools scored more highly over time (P=0.008).
		Student-Teacher and/or Staff Relationships	YES	null	None of the variables about student staff relationships were significantly higher for students in Comer schools. Both teachers encouraging better academic perforamnce and teachers caring about non-academic needs saw positive trends over time in both Comer and non-Comer schools.
QEDs 2017	Communities in Schools: 2017, Texas and North Carolina	There were no non-academic outcomes included in this study.			

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
	City Connects: 2016	There were no non-academic outcomes included in this study.			
	City Year	There were no non-academic outcomes included in this study.			
	Talent Development	There were no non-academic outcomes included in this study.			
	Say Yes	Outcomes for Students as Individuals			
		School Attachment	NO		
		Behavior Problems	YES	null	There was a marginally significant decrease in suspensions in the 2nd year of the program (P=0.093) but otherwise there was no difference in suspensions between Say Yes participants and their propensity score matched comparison group across the other four years included in the study.
		Socio-emotional Development	NO		
		Student Health and Well-being	NO		
		Student Outcomes Measured in Families			
		Academic support at home	NO		
		Use of positive	NO		

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
		parenting techniques			
		Presence of positive parent-child relationships	NO		
		Outcomes Within the School Environment			
		Positive School Climate	NO		
		Student-Teacher and/or Staff Relationships	NO		
QEDs 2014	City Connects: 2008-2009	Outcomes for Students as Individuals			
		School Attachment	NO		
		Behavior Problems	YES	+	For all kids in all grades, behavior improved after being in the City Connects program for at least a year. The overall gains were largeest for those who started in 1st or 2nd grade (up to 5th grade). This was particularly true for male students.
		Socio-emotional Development	YES	+	Effort and work ethic all increase for students the who were in City Connects.
		Student Health and Well-being	YES	+	Students in 4th and 5th grade in scored more highly on tests about unhealthy nutrition and overall well-being (P=0.000 for both).
		Student Outcomes Measured in Families			

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
		Academic support at home	NO		
		Use of positive parenting techniques	NO		
		Presence of positive parent-child relationships	NO		
		Outcomes Within the School Environment			
		Positive School Climate	NO		The authors do not report statistical differences but present qualitative findings of positive impacts on school climate.
		Student Teacher and/or Staff Relationships	NO		The authors don't report statistical differences but they do show that 74-80% of City Connects teachers report knowing more about their students strengths, needs, and what services could be useful.
	City Connects: 2010	Outcomes for Students as Individuals			
		School Attachment	NO		
		Behavior Problems	YES	+	City Connects students had better classroom behavior scores in grades 3 and 5. This did not remain significantly different in adjusted models.
		Socio-emotional Development	YES	+	City Connects students had better work habit scores in grades 3 and 5 and better work effort scores in grades 3, 4, and 5. The effot

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
					differences and the work habits for 5th graders remained significant in adjusted models.
		Student Health and Well-being	YES	+	2nd and 3rd grade students in City Connects learned more about the food pyramid, the importance of exercise and were less likely to have drunk soda the day before. They also had higher overall health knowledge. 4th and 5th graders were less likely to eat junk food before of after TV, drunk soda, eat candy, or eat french fries. They also knew more about nutrition, screen time, and physical health.
		Student Outcomes Measured in Families			
		Academic support at home	NO		
		Use of positive parenting techniques	NO		
		Presence of positive parent-child relationships	NO		
		Outcomes Within the School Environment			
		Positive School Climate	NO		
		Student Teacher	NO		They do not report statistical tests on teacher relationships with students but do

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
		and/or Staff Relationships			report the percentage of teachers in City Connects schools who can support students.
	City Connects: 2012	Outcomes for Students as Individuals			
		School Attachment	NO		
		Behavior Problems	NO		
		Socio-emotional Development	NO		
		Student Health and Well-being	NO		
		Student Outcomes Measured in Families			
		Academic support at home	NO		
		Use of positive parenting techniques	NO		
		Presence of positive parent-child relationships	NO		
		Outcomes Within the School Environment			
		Positive School Climate	NO		

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
		Student-Teacher and/or Staff Relationships	NO		The authors don't report statistical differences but they do show that 85-91% of City Connects teachers report knowing more about their students, understand the non-academic aspects of their students' lives, and better understood the dynamics in their classroom.
	Comer: Chicago	Outcomes for Students as Individuals			
		School Attachment	YES	+	There was evidence of students feel more attached to school in the Comer schools.
		Behavior Problems	YES	null	Comer students had reported lower behavior scores; however there is no indication this gap widened over time, just that they started out and stayed lower.
		Socio-emotional Development	NO		
		Student Health and Well-being	NO		
		Student Outcomes Measured in Families			
		Academic support at home	YES	null	Comer students had reported lower parental valuation of education; however there is no indication this gap widened over time, just that they started out and stayed lower.
		Use of positive parenting techniques	NO		

	Study	Outcomes	Included?	Positive, Negative, or Null Impact	Notes
		Presence of positive parent-child relationships	NO		
		Outcomes Within the School Environment			
		Positive School Climate	YES	null & +	School climate increased over time in both the Comer and non-Comer schools and there was no indication that it was occurring at a faster rate in the Comer Schools (they started out lower). Individual level scores from teachers on school climate were more reliable than school level scores and showed Comer consistently below (although not increasing in distance). Student reports on school climate were more positive for Comer schools and several of the indicators had positive and significant associations with being in a Comer school.
		Student-Teacher and/or Staff Relationships	YES	null & +	Several of the indicators about relationships between students and teachers had significant positive associations with Comer at both the individual and the school level. This included evidence of a difference in slopes such that Comer schools were improving even faster than non-Comer schools.

Appendix 4: Detailed Description of the Social Genome Model (SGM)

The following Appendix provides an overview of the Social Genome Model (SGM), which is used to simulate the long-term impacts of ISS programs on pages 38-39 of this report. This Appendix describes what the SGM is, how it was constructed, how the model can be used, and how the simulations work. Additional details about the model and how it operates can be found in the *Guide to the Social Genome Project* prepared by the Urban Institute.[cviii]

THE SOCIAL GENOME MODEL. The Social Genome Model, originally developed at the Brookings Institution and based at the Urban Institute, is a collaborative effort of the Brookings Institution, Child Trends, and the Urban Institute. The SGM is a microsimulation model; that is, it employs data to simulate outcomes for individual persons. It is used to empirically examine how social policies and programs can influence mobility. Currently, research on mobility usually focuses on only one intervention at a specific life stage and mainly examines short-term outcomes.[cix] However, the SGM can predict how one or more interventions that affect child and youth development during the key life stages of human capital formation can influence outcomes of well-being into adulthood. Each variable and life stage in the SGM was carefully selected based on studies of the factors that promote or hinder success at significant milestones during a person's early life. This "ecological" model of child and youth development accounts for the variety of components that influence development and is widely accepted in the field.[cx] Multiple analyses have been performed to confirm that the estimated outcomes generated through the SGM are valid.[cxi]

The SGM has been used to answer questions regarding mobility in policy briefs and research studies, including those published in peer-reviewed journals.[cxii, cxiii] The model is well suited for answering "what-if" questions about how altering certain factors during relevant life stages can change later-life outcomes. For instance, the SGM can be used to answer the question "what if we reduced school suspensions by 50 percent during middle childhood?" and determine how this would influence college completion and earnings at age 29. The model can also be used to examine how the impacts found in empirical studies of specific interventions, such as evaluations of ISS programs, can influence adolescent and adulthood outcomes.

SGM VERSIONS & LIFE STAGES. There are currently two versions of the SGM. The first iteration of the model, called SGM-79, was built by the Brookings Institution and utilizes data from the National Longitudinal Survey of Youth 1979 (NLSY79) and the Children of the National Longitudinal Survey of Youth (CNLSY). For more technical detail about how the two datasets are combined, see the *Guide to the Social Genome Project*.[cxiv] Using data from these two surveys, the model incorporates factors that promote or hinder well-being at six developmentally important life stages from birth to age 40:

- circumstances at birth (from NLSY79 and CNLSY)
- early childhood (age 5, from CNLSY)
- middle childhood (age 11, from CNLSY)
- adolescence (age 19, from NLSY79 and CNLSY)

- young adulthood (age 29, from NLSY79)
- adulthood (age 40, from NLSY79)

The second iteration of the SGM, used for the analyses in this report, was built by Child Trends using the more recent National Longitudinal Survey of Youth 1997 (NLSY97). This version, called SGM-97, is based on the initial version and includes similar factors that influence the future well-being of children and youth in six, shorter life stages from birth to age 29:

- circumstances at birth (collected retrospectively from NLSY97 respondents at ages 12 to 16)
- middle childhood (ages 12 to 13)
- early adolescence (age 15)
- adolescence (age 19)
- early transition to adulthood (age 25)
- transition to adulthood (age 29)

SGM-97 VARIABLES. Variables for the SGM-97 were selected because of their developmental significance at each life stage and their ability to predict future success. These variables are often compatible with the types of outcomes on which child and youth development interventions have impacts. For instance, many evaluations of education programs focus on how the programs impact academic achievement via math scores, and these effects can be included in the SGM using the standardized Peabody Individual Achievement Test (PIAT) math scores variable to predict adulthood outcomes like educational attainment. Missing values for variables were imputed using proximity imputation, in which values are "filled in" using data from a respondent's other interviews, and through regression imputation, in which information about a respondent's other characteristics is used to predict missing values for variables. A list of the variables in the SGM-97 can be found in Table XX.

Variables in the SGM-97

Life Stage	Variables	
Circumstances at birth	• Race • Gender • Maternal education	• Maternal age at child's birth • Maternal age at first birth • Marital status
Middle childhood	• PIAT Math score • Child does not lie/cheat	• Behavioral problems scale
Early adolescence	• Ever suspended • Delinquency index • Days per week/religious • Ever had sex	• PIAT Math score • Armed Services Vocational Aptitude Battery (ASVAB) score
Adolescence	• High school diploma (by age 19)	• Ever used other drugs • Self-esteem

Life Stage	Variables	
	• High school GPA • Ever convicted (by age 20) • Had teen birth (by age 20) • Family income (2011$, age 19) • Ever used marijuana	• Mental health scale • Days per week/religious • General health
Early transition to adulthood	• Family income (2011$) • Family income-to-needs ratio • Completed four-year degree • Lives independently • Parenthood status	• Marital status • Convicted (ages 20 to 24) • General health • Down/depressed
Transition to adulthood	• Family income (2011$) • Family income-to-needs ratio • Completed four-year degree • Lives independently • Parenthood status	• Marital status • Personal earnings (2011$) • General health • Down/depressed

SGM-97 MODEL SPECIFICATION. The SGM-97 employs a series of regressions to predict how indicators of well-being can influence later-life outcomes. Each regression predicts outcomes based on youths' characteristics and backgrounds from earlier life stages, starting at middle childhood. Ordinary least squares regressions are used for continuous dependent variables, while linear probability models are used for binary dependent variables. The equations used for the regressions at each stage are as follows:

(1) MC Outcome = $\beta_0 + \beta_1 CAB_{NLSY97} + \varepsilon$
(2) EADOL Outcome = $\beta_0 + \beta_1 CAB_{NLSY97} + \beta_2 MC_{NLSY97} + \varepsilon$
(3) ADOL Outcome = $\beta_0 + \beta_1 CAB_{NLSY97} + \beta_2 MC_{NLSY97} + \beta_3 EADOL_{NLSY97} + \varepsilon$
(4) ETTA Outcome = $\beta_0 + \beta_1 CAB_{NLSY97} + \beta_2 MC_{NLSY97} + \beta_3 EADOL_{NLSY97} + \beta_4 ADOL_{NLSY97} + \varepsilon$
(5) TTA Outcome = $\beta_0 + \beta_1 CAB_{NLSY97} + \beta_2 MC_{NLSY97} + \beta_3 EADOL_{NLSY97} + \beta_4 ADOL_{NLSY97} + \beta_5 ETTA_{NLSY97} + \varepsilon$

In these equations, CAB is the circumstances at birth life stage, MC is middle childhood, EADOL is early adolescence, ADOL is adolescence, ETTA is early transition to adulthood, and TTA is transition to

adulthood. Each regression controls for all of the variables from the previous life stages. Additionally, all regressions include NLSY97 sample weights.

SGM-97 SIMULATION PROCESS. The SGM-97 can be used to run simulations for the overall NLSY97 sample or for a specific target population, such as female respondents or children of teen mothers. The simulation process starts by estimating the coefficients of the relationships between the variables in each life stage and those in later life stages. Then, the values of each variable for respondents in the chosen population are estimated at each life stage, and are averaged across this population to determine the mean values for each variable if no intervention occurred (the baseline).

Once the baseline means have been established, the values of one or more variables at a specific life stage can be changed to simulate the influence of a program or policy intervention. For continuous variables, the change expected to result from an intervention is expressed in standardized mean difference effect sizes (Cohen's *d*), while percentage changes are used for binary variables. The decision of (a) what variables should be adjusted and (b) the magnitude to which they should be adjusted (i.e., the effect size) depends on the "what if" scenario or intervention that is being simulated and is generally based on findings from relevant research literature. For the ISS simulations, we used effect sizes from evaluations of ISS programs to alter variables in the SGM that these interventions have been found to impact, such as math scores and high school graduation rates.

For each variable in the life stages following that in which the intervention occurred, new values are estimated to simulate the effects of the intervention. These variables become the outcomes of the intervention, and their mean values can be compared to the baseline (preintervention) mean values to determine the simulated effect of the intervention on each outcome. Follow up calculations can also be performed to gain further insight into the results, such as using CPS data to predict lifetime earnings or determining if respondents are "successful" in a given life stage, based on passing certain thresholds for indicators that predict future well-being. Further details about follow-up analyses can be found in the *Guide to the Social Genome Project*.[cxv]

Appendix 5: Descriptions of MTSS and PBIS

Multi-Tiered Systems of Support

Multi-Tiered Systems of Support (MTSS) is a three-tiered framework that grew from Response-to-Intervention (RtI) -- a screening, monitoring, and decision-makin process to improve the identification of students with disabilities. MTSS starts with modifying classroom instruction and then documenting student performance to identify additional supports. The framework includes school-wide approaches at the lowest tier that are available for all students, a middle tier that aims to respond quickly to students with sudden, smaller needs, and finally, a top tier that aims to support students with the highest needs with more time and labor intensive supports as needed. MTSS relies on data and progress monitoring to determine whether the continuum of academic and behavioral supports needs adjusting, and utilizes strategies focused on students, teachers, and additional support staff primarily within the school building.

Positive Behavior Interventions and Supports

Positive Behavior Interventions and Supports (PBIS) is a type of MTSS that focuses on implementing proactive strategies to teach and promote positive behavior in youth. Based on the premise that some students will need additional supports to exhibit positive behavior, the approach is designed to provide those supports as needs are identified. The first tier includes universal approaches for all students, the second tier focuses on a select group identified through data collection, and the third tier focuses on students with the most need providing individual support. The four key elements of PBIS are outcomes, data, practices, and systems to support staff and student behavior and decision making. Ultimately, by promoting and improving positive behavior, academic achievement is expected to also improve.

Works Cited

[i] The National Child Traumatic Stress Network. (2017). *The effects of trauma on schools and learning*. Retrieved from http://www.nctsn.org/resources/audiences/school-personnel/effects-of-trauma

[ii] Trauma and Learning Policy Initiative (TLPI). (2017). *Helping traumatized children learn*. Retrieved from https://traumasensitiveschools.org/trauma-and-learning/the-problem-impact/

[iii] Magnuson, K. (2013). *Reducing the effects of poverty through early childhood interventions*. Madison, WI: Fast Focus. Institute for Research on Poverty. Retrieved from http://www.irp.wisc.edu/publications/fastfocus/pdfs/FF17-2013.pdf

[iv] National Center for Chronic Disease Prevention and Health Promotion Division of Population Health. (2017). *Heath and academic achievement*. Retrieved from https://www.cdc.gov/healthyyouth/health_and_academics/pdf/health-academic-achievement.pdf

[v] Center for School Mental Heath. (2012). *The impact of school mental health: educational, social, emotional, and behavioral outcomes*. Retrieved from http://csmh.umaryland.edu/media/SOM/Microsites/CSMH/docs/CSMH-SMH-Impact-Summary-July-2013-.pdf

[vi] Office of Policy Development and Research (PD&R) U.S. Department of Housing and Urban Development. (2014). How housing mobility affects education outcomes for low- income children. Retrieved from https://www.huduser.gov/portal/periodicals/em/fall14/highlight2.html

[vii] Sandstrom, H. & Huerta, S. (2013). *The negative effects of instability on child development: a research synthesis. low-income working families*. (Discussion Paper 3). Washington, D.C.: Urban Institute. Retrieved from http://www.urban.org/sites/default/files/publication/32706/412899-The-Negative-Effects-of-Instability-on-Child-Development-A-Research-Synthesis.PDF

[viii] Child Trends DataBank. (2014). *Food insecurity*. Retrieved from https://www.childtrends.org/wp-content/uploads/2016/12/117_Food_Insecurity.pdf

[ix] Child Trends DataBank. (2016). *Poverty*. Retrieved from https://www.childtrends.org/wp-content/uploads/2016/12/04_Poverty.pdf

[x] Walsh, M. E., Gish, J. W., Foley, C., Theodorakakis, M., & Rene, K. (2016). *Principles of effective practice for integrated student support*. Retrieved from http://www.bc.edu/content/dam/files/schools/lsoe/cityconnects/pdf/Policy%20Brief%20-%20Building%20Sustainable%20Interventions%20web.pdf

[xi] Moore, K.A., Caal, S., Carney, R., Lippman, L., Li, W., Muenks, K., Murphey, D., Princiotta, D., Ramirez, A., Rojas, A. and Ryberg, R. (2014). *Making the grade: Assessing the evidence for integrated student supports*. Bethesda, MD: Child Trends. Retrieved from https://childtrends-ciw49tixgw5lbab.stackpathdns.com/wp-content/uploads/2014/02/2014-07ISSPaper2.pdf.

[xii] RTI Action Network. (2017). What is RTI? Retrieved from http://www.rtinetwork.org/learn/what/whatisrti

[xiii] Moore, K.A., Caal, S., Carney, R., Lippman, L., Li, W., Muenks, K., Murphey, D., Princiotta, D., Ramirez, A., Rojas, A. and Ryberg, R. (2014). *Making the grade: Assessing the evidence for integrated student supports*. Bethesda, MD: Child Trends; page 1. Retrieved from https://childtrends-ciw49tixgw5lbab.stackpathdns.com/wp-content/uploads/2014/02/2014-07ISSPaper2.pdf.

[xiv] Coalition for Community Schools at the Institute for Educational Leadership. (n.d.). *Community school standards*. Retrieved from http://www.communityschools.org/assets/1/Page/Community-School%20Standards-Updatesd2017.pdf

[xv] U.S. Department of Education. (2011). *Trends in high school dropout and completion rates in the United States: 1972-2009;* page 52. Retrieved from https://nces.ed.gov/pubs2012/2012006.pdf

[xvi] National Center for Education Statistics. (2017). Public high school 4-year adjusted cohort graduation rate (ACGR) for the United States, the 50 states and the District of Columbia: School years 2010-11 to 2012-13. Retrieved from https://nces.ed.gov/ccd/tables/ACGR_2010-11_to_2012-13.asp

[xvii] National Center for Education Statistics. (2017). *High school graduation rates*. Retrieved from https://nces.ed.gov/fastfacts/display.asp?id=805

[xviii] Child Trends DataBank. (2015). *Reading proficiency*. Retrieved from https://www.childtrends.org/wp-content/uploads/2015/11/29_Reading_Proficiency.pdf

Child Trends DataBank. (2015). *Mathematics proficiency*. Retrieved from https://www.childtrends.org/wp-content/uploads/2015/11/09_Mathematics_Proficiency.pdf

[xixxix] Child Trends DataBank. (2015). *Reading proficiency*. Retrieved from https://www.childtrends.org/wp-content/uploads/2015/11/29_Reading_Proficiency.pdf

[xx] Child Trends DataBank. (2016). *Educational attainment*. Retrieved from https://www.childtrends.org/indicators/educational-attainment/

[xxi] United States Department of Labor: Bureau of Labor Statistics. (2017). *Employment projections*. Retrieved from https://www.bls.gov/emp/ep_chart_001.htm

[xxii] National Center for Education Statistics. (2017). *Public high school graduation rates*. Retrieved from https://nces.ed.gov/programs/coe/indicator_coi.asp

[xxiii] America's Promise Alliance. (2014). *Building a gradnation: progress and challenge in ending the high school dropout epidemic*. Retrieved from http://www.americaspromise.org/resource/building-gradnation-progress-and-challenge-ending-high-school-dropout-epidemic-2014

[xxiv] America's Promise Alliance. (2014). *Building a gradnation: progress and challenge in ending the high school dropout epidemic*. Retrieved from http://www.americaspromise.org/resource/building-gradnation-progress-and-challenge-ending-high-school-dropout-epidemic-2014

[xxv] Lee, J.M. (2013). *The educational experience of young men of color: a review of research, pathways and progress*. Retrieved from http://ciscentraltexas.org/wp-content/uploads/2013/08/EEYMC-ResearchReport.pdf

[xxvi] Bo Schma, J. and Brownstein, R. (2016). *The concentration of poverty in American schools*. *The Atlantic*. Retrieved from https://www.theatlantic.com/education/archive/2016/02/concentration-poverty-american-schools/471414/.

[xxvii] Child Trends DataBank. (2016). *Educational attainment*. Retrieved from https://www.childtrends.org/indicators/educational-attainment/

Pew Social Trends. (2017). *The rising cost of not going to college*. Retrieved from http://www.pewsocialtrends.org/2014/02/11/the-rising-cost-of-not-going-to-college/

[xxviii] Communities in Schools. *About us*. (2017). Retrieved from https://www.communitiesinschools.org/about/ accessed 7/18/16.

[xxix] Moore, K.A., Caal, S., Carney, R., Lippman, L., Li, W., Muenks, K., Murphey, D., Princiotta, D., Ramirez, A., Rojas, A. and Ryberg, R. (2014). *Making the grade: Assessing the evidence for integrated student supports*. Bethesda, MD: Child Trends. Retrieved from https://childtrends-ciw49tixgw5lbab.stackpathdns.com/wp-content/uploads/2014/02/2014-07ISSPaper2.pdf

[xxx] Moore, K.A., Caal, S., Carney, R., Lippman, L., Li, W., Muenks, K., Murphey, D., Princiotta, D., Ramirez, A., Rojas, A. and Ryberg, R. (2014). *Making the grade: Assessing the evidence for integrated student supports*. Bethesda, MD: Child Trends. Retrieved from https://childtrends-ciw49tixgw5lbab.stackpathdns.com/wp-content/uploads/2014/02/2014-07ISSPaper2.pdf

[xxxi] Brown, B.V. & Moore, K.A. (2003). Child and youth well-being: The social indicators field. *Handbook of Applied Developmental Science*, 437-467.

[xxxii] Moore, K.A. (1997). Criteria for indicators of child well-being. In R.M. Hauser, B.B. Brown, & W.R. Prosser (Eds.) *Indicators of Child Well-Being*. New York: Russell Sage Foundation.

[xxxiii] Phillips, D. A., & Shonkoff, J. P. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. National Academies Press.

[xxxiv] Phillips, D. A., & Shonkoff, J. P. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, D.C.: National Academies Press.

[xxxv] Elder, G. H. (1998). The life course as developmental theory. *Child Development*, *69*(1), 1-12.

[xxxvi] ASCD. (2017). *Whole school, whole community, whole child*. Retrieved from http://www.ascd.org/programs/learning-and-health/wscc-model.aspx

[xxxvii] Coalition for Community Schools at the Institute for Educational Leadership. (2017). *What is a community school?* Retrieved from http://www.communityschools.org/aboutschools/what_is_a_community_school.aspx

[xxxviii] Coalition for Community Schools at the Institute for Educational Leadership. (2017). *Community Schools: A whole child framework for school improvement*. Retrieved from http://www.communityschools.org/assets/1/AssetManager/Community-Schools-A-Whole-Child-Approach-to-School-Improvement1.pdf

[xxxix] Oakes, J., Maier, A., & Daniel, J. (2017). Community schools: An evidence-based strategy for equitable school improvement. Palo Alto, CA: *National Education Policy Center*. Retrieved from http://nepc.colorado.edu/publication/equitable-community-schools

[xl] World Health Organization. (2017). *Social determinants of health*. Retrieved from http://www.who.int/social_determinants/en/

[xli] Robert Wood Johnson Foundation. (2016). *The relationship between school attendance and health*. Health Policy Snapshot Series. Retrieved from http://www.rwjf.org/en/library/research/2016/09/the-relationship-between-school-attendance-and-health.html

[xlii]Center for Mental Health in Schools at UCLA. (2011). *Moving beyond the three tier intervention pyramid toward a comprehensive framework for student and learning supports*. Los Angeles, CA: A Center Policy Brief. Retrieved from http://smhp.psych.ucla.edu/pdfdocs/briefs/threetier.pdf

[xliii] Kolbe, L. J., Allensworth, D. D., Potts-Datema, W., & White, D. R. (2015). What have we learned from collaborative partnerships to concomitantly improve both education and health? *Journal of School Health*, *85*(11), 766-774.

[xliv] The Collaborative for Academic, Social, and Emotional Learning (CASEL). (2017). What is SEL? Retrieved from http://www.casel.org/what-is-sel/

[xlv] McKown, C. (2017). social-emotional assessment, performance, and standards. *The Future of Children*, 27(1), 157-178.

[xlvi] Moore, K. A., Lippman, L. H., & Ryberg, R. (2015). Improving outcome measures other than achievement: AERA Open, *1*(2): 1-25.

[xlvii] Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405-432.

[xlviii] Dweck, C. S. (2006). Mindset: The new psychology of success. New York: Random House.

[xlix] Princeton University. (2017). Social and Emotional Learning. *The Future of Children*, *27*(1). Retrieved from https://www.jstor.org/stable/i40176578

[l]Belfield, C., Bowden, A. B., Klapp, A., Levin, H., Shand, R., & Zander, S. (2015). The economic value of social and emotional learning. *Journal of Benefit-Cost Analysis*, *6*(3), 508-544.

[li]Frizzell, M., Braun, M., Ferguson, M., Stark Rentner, D., & Kober, N. (2017). *Building Competencies for Careers*. Washington, D.C.: Center on Education Policy. Retrieved from https://www.cep-dc.org/displayDocument.cfm?DocumentID=1478

[lii]Balcar, J. (2014). Soft skills and their wage returns: Overview of empirical literature. *Review of Economic Perspectives*, *14*(1), 3-15.

[liii] Carnevale, A. P. (2013). *21st century competencies for college and career readiness*. Broken Arrow, OK: National Career Development Association.

[liv] Eger, H., & Grossmann, V. (2004). Noncognitive abilities and within-group wage inequality. Bonn: Institute for the Study of Labour.

[lv] International Labour Organization. (2008). Skills for improved productivity, employment growth and development. Paper presented at the International Labour Conference.

[lvi] Frizzell, M., Braun, M., Ferguson, M., Stark Rentner, D., & Kober, N. (2017). *Building competencies for careers*. Washington, D.C.: Center on Education Policy. Retrieved from https://www.cep-dc.org/displayDocument.cfm?DocumentID=1478

[lvii] Lippman, L. H., Ryberg, R., Carney, R., & Moore, K. A. (2015). Workforce connections: key "soft skills" that foster youth workforce success: toward a consensus across fields. Washington, DC: Child Trends.

[lviii] Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success* (No. w20749). Cambridge, MA: National Bureau of Economic Research.

[lix] Moore, K. A., Lippman, L. H., & Ryberg, R. (2015). Improving Outcome Measures Other Than Achievement. AERA Open, *1*(2): 1-25.

[lx]Oakes, J., Maier, A., & Daniel, J. (2017). *Community schools: An evidence-based strategy for equitable school improvement*. Boulder, CO: National Education Policy Center. Retrieved from http://nepc.colorado.edu/publication/equitable-community-schools

[lxi] Oakes, J., Maier, A., & Daniel, J. (2017). Community schools: An evidence-based strategy for equitable school improvement. Boulder, CO: National Education Policy Center. Retrieved from http://nepc.colorado.edu/publication/equitable-community-schools

[lxii] United States Department of Education. (2017). Chronic absenteeism in the nation's schools. an unprecedented look at a hidden educational crisis. Retrieved from https://www2.ed.gov/datastory/chronicabsenteeism.html

[lxiii]Balfanz, R., & Byrnes, V. (2012). The importance of being in school: A report on absenteeism in the nation's public schools. *The Education Digest*, *78*(2), 4. Retrieved from https://eric.ed.gov/?id=EJ1002822

[lxiv] Rothstein, R. (2004). Class and schools: using social, economic, and educational reform to close the black-white achievement gap. Washington, DC: Economic Policy Institute.

[lxv] Oakes, J., Maier, A., & Daniel, J. (2017). Community schools: an evidence-based strategy for equitable school improvement. Boulder, CO: National Education Policy Center. Retrieved from http://nepc.colorado.edu/publication/equitable-community-schools

[lxvi] Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405-432.

[lxvii] Baker, J. Grant, s., & Morlock, L. (2008). The teacher–student relationship as a developmental context for children with internalizing or externalizing behavior problems. *School Psychology Quarterly*, *23*(1), 3-15.

[lxviii] Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, *72*(2), 625-638.

[lxix] Lacey, A., & Cornell, D. (2013). The impact of bullying climate on schoolwide academic performance. *Journal of Applied School Psychology*, *29*, 262-283.

[lxx] Cornell, D., Gregory, A., Huang, F., & Fan, X. (2013). Perceived prevalence of teasing and bullying predicts high school dropout rates. *Journal of Educational Psychology*, *105*, 138-149.

[lxxi] Office of Adolescent Health, Sample attrition in teen pregnancy prevention impact evalu. . (2014). Retrieved from https://www.hhs.gov/ash/oah/oah-initiatives/assets/attrition-ta-brief.pdf

[lxxii] Song M. & Herman, R. (2009). Practical Guide on Designing and Conducting Impact Studies in Education: Lessons Learned From the What Works Clearinghouse (Phase 1). Washington, D.C: American Institutes for Research. Retrieved from http://www.air.org/sites/default/files/downloads/report/Song__Herman_WWC_Lessons_Learned_2010_0.pdf

[lxxiii] Figlio, D. (2015). Experimental Evidence of the Effects of the Communities in Schools of Chicago Partnership Program on Student Achievement. *National Bureau of Economic Research (NBER) Working Paper.*

[lxxiv] Parise, L, Corrin, W, Granito, K et. al. (2017). Two Years of Case Management: Final Findings from the Communities in Schools Random Assignment Evaluation. MDRC, New York, NY.

[lxxv] Corrin, W, Sepanik, S, Rosen, R, and Shane, A. (2016). Addressing Early Warning Indicators: Interim Impact Findings from the Investing in Innovation (i3) Evaluation of Diplomas Now. MDRC, New York, NY.

[lxxvi] Dobbie, W and Fryer, RG. (2011). Are High-Quality Schools Enough to Increase Achievement Among the Poor? Evidence from the Harlem Children's Zone. *American Economic Journal: Applied Economics;* 3:158-187.

[lxxvii] Dobbie, W and Fryer, RD. (2015). The Medium-Term Impacts of High-Achieving Charter Schools. *Journal of Political Economy; 123 (5).*

[lxxviii] Daring, E, Walsh, ME, Sibley, E et. al. (2016). Can Community and School-Based Supports Improve the Achievement of First-Generation Immigrant Children Attending High-Poverty Schools? *Child Development;* 87(3): 883-897.

[lxxix] Walsh, ME, Madaus, GF, Raczek, AE et. al. (2014). A New Model for Student Support in High-Poverty Urban Elementary Schools: Effects on Elementary and Middle School Academic Outcomes. *American Educational Research Journal;* 51(4): 704-737.

[lxxx] Meredith, J and Anderson, LM. (2015). Analysis of the Impacts of City Year's Whole School Whole Child Model on Partner Schools' Performance. Policy Studies Associates, Inc., Washington, D.C.

[lxxxi] Somers, M-A and Haider, Z. (2017). A Quasi-Experimental Evaluation of Communities in Schools. MDRC, New York, NY.

[lxxxii] Turner, HB, Schoeneberger, J, Hartman, T, and Gold, E. (2007). Making Progress: Longitudinal Outcomes Evaluation of the Say Yes to Education Program Kindergarten-5th Grade. Research for Action and Analytica, Phoenixville, PA and Philadelphia, PA.

[lxxxiii] Kemple, JJ, Herlihy, CM, and Smith, TJ. (2005). Making Progress Toward Graduation: Evidence from the Talent Development High School Model. MDRC, New York, NY.

[lxxxiv] Cook, TD, Habib, F-N, Phillips, M et. al. (1999). Comer's School Development Program in Prince George's County, Maryland: A Theory Based Evaluation. *American Educational Research Journal;* 36(3): 543-597.

[lxxxv] ICF International. (2010). Communities in Schools National Evaluation Volume 5: Randomized Controlled Trial Study Austin, Texas. ICF International: Fairfax, VA

[lxxxvi] ICF International. (2010). Communities in Schools National Evaluation Volume 4: Randomized Controlled Trial Study Jacksonville, Florida. ICF International: Fairfax, VA

[lxxxvii] ICF International. (2010). Communities in Schools National Evaluation Volume 6: Randomized Controlled Trial Study Wichita, Kansas. ICF International: Fairfax, VA

[lxxxviii] Boston College Center for Optimized Student Support. (2012). The Impact of City Connects: Progress Report 2012. Chestnut Hill, MA.

Boston College Center for Optimized Student Support. (2010). The Impact of City Connects: Annual Report 2010. Chestnut Hill, MA.

Boston College Center for Child, Family and Community Partnerships. (2009). The Impact of Boston Connects: Summary Report 2008-2009. Chestnut Hill, MA.

[lxxxix] Cook, T. D., Murphy, R. F., & Hunt, H. D. (2000). Comer's School Development Program in Chicago: A theory-based evaluation. *American Educational Research Journal*, 37(2), 535-597.

[xc] U.S. Department of Health and Human Services. (2017). *Human subject regulations decision charts*. Retrieved from https://www.hhs.gov/ohrp/regulations-and-policy/decision-charts/index.html#c1

[xci] Balfanz, R., Herzog, L. & D. J. Mac Iver (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: early identification and effective interventions. Educational Psychologist, 42(4): 223-235.

[xcii] Bruce, M., Bridgeland, J. M., Fox, J. H., and Balfanz, R. (2011). *On track for success: the use of early warning indicator and intervention systems to build a grad nation*. Washington, DC: Civic Enterprises. Retrieved from http://new.every1graduates.org/wp-content/uploads/2012/03/on_track_for_success.pdf

[xciii] Song M. & Herman, R. (2009). Practical Guide on Designing and Conducting Impact Studies in Education: Lessons Learned From the What Works Clearinghouse (Phase 1). Washington, D.C: American Institutes for Research. Retrieved from http://www.air.org/sites/default/files/downloads/report/Song__Herman_WWC_Lessons_Learned_2010_0.pdf

[xciv] U.S. Department of Education. (2016). Chronic absenteeism in the Nation's schools: An unprecedented look at a hidden educational crisis. Retrieved from https://www2.ed.gov/datastory/chronicabsenteeism.html

[xcv] Balfanz, R., & Byrnes, V. (2012). *The importance of being in school: A report on absenteeism in the nation's public schools*. Retrieved from http://new.every1graduates.org/wp-content/uploads/2012/05/FINALChronicAbsenteeismReport_May16.pdf

[xcvi] Balfanz, R., & Byrnes, V. (2012). *The importance of being in school: A report on absenteeism in the nation's public schools*. Retrieved from http://new.every1graduates.org/wp-content/uploads/2012/05/FINALChronicAbsenteeismReport_May16.pdf

[xcvii] Walsh, ME, Gish, JW, Foley, C, Theodorakakis, M, and Rene, K. (2016). *Policy brief principles of effective practice for integrated student support*. Boston, MA: Boston College, Center for Optimized Student Support. Retrieved from http://www.bc.edu/content/dam/files/schools/lsoe/cityconnects/pdf/Policy%20Brief%20-%20Building%20Sustainable%20Interventions%20web.pdf

[xcviii] Durlak, J. A. (2010). The importance of doing well in whatever you do: A commentary on the special section, "Implementation research in early childhood education*". Early Childhood Research Quarterly*, *25*(3), 348-357. Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, *41*(3-4), 327-350.

[xcix] Russell, C. A., Mielke, M. B., & Reisner, E. R. (2009). Evidence of Program Quality and Youth Outcomes in the DYCD Out-of-School Time Initiative: Report on the Initiative's First Three Years. Washington, D.C.: Policy Studies Associates, Inc. Retrieved from http://www.wallacefoundation.org/knowledge-center/Documents/Evidence-of-Program-Quality-and-Youth-Outcomes.pdf

[c] Temkin, D. and Harper K. (2017, September 20). Some states are missing the point of ESSA's fifth indicator. [Child Trends Blog]. Retrieved from https://www.childtrends.org/states-missing-point-essas-fifth-indicator/

[ci] Jacob, B and Ludwif, J. (2008). Improving Educational Outcomes for Poor Children. *NBER Working Paper 14550*, Cambridge, MA. Retrieved from http://www.nber.org/papers/w14550.pdf.

[cii] Barik, TJ, Belford, JA, Gormley,, WT, and Anderson, S.. (2016). A Benefit-Cost Analysis of the Tulsa Universal Pre-K Program. The Merits of Universal Scholars: Benefit-Cost Evidence from the Kalamazoo Promise. Upjohn Institute Working Paper 16-261. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. Retrieved from http://research.upjohn.org/up_workingpapers/261/.

[ciii] Belfield, C., Bowden, A. B., Klapp, A., Levin, H., Shand, R., & Zander, S. (2015). The economic value of social and emotional learning. *Journal of Benefit-Cost Analysis*, *6*(3), 508-544.

[civ] Walsh, M. E., Madaus, G. F., Raczek, A. E., Dearing, E., Foley, C., An, C., Lee-St. John, T. J., & Beaton, A. (2014). A New Model for Student Support in High-Poverty Urban Elementary Schools: Effects on Elementary and Middle School Academic Outcomes. *American Educational Research Journal*, *51*(4), 704-737.

[cv] Economic Modeling Specialists Inc. – EMSI. (2012). *The Economic Impact of Communities in Schools.* Retrieved from: https://www.communitiesinschools.org/media/uploads/attachments/CIS_2.pdf

[cvi] Martinez, L., Haynes, C.D. & Silloway, T. (2013). *Measuring Social Return on Investment for Community Schools – A Practical Guide*. The Children's Aid Society and The Finance Project. Retrieved from http://www.nccs.org/sites/default/files/resource/NCCS_SROI_Guide.pdf

[cvii] DeNike, M. and B. Ohlson (2013). *Elev8: Oakland community school costs and benefits: making dollars and cents of the research*. (Bright Research Group Policy Brief Issue 1). Oakland, CA: Bright Research Group. Retrieved from http://www.brightresearchgroup.com/wp-content/uploads/2015/10/Elev8-Oakland-Community-Schools-Cost-Benefit-Analysis.pdf

[cviii] Acs, G., Steven M., and Schwabish, J.A. (2016). *Guide to the social genome project*. Washington, DC: Urban Institute. Retrieved from https://www.urban.org/sites/default/files/publication/85716/guide-to-the-social-genome-project_2.pdf

[cix] Acs, G., Steven M., and Schwabish, J.A. (2016). *Guide to the social genome project*. Washington, DC: Urban Institute. Retrieved from https://www.urban.org/sites/default/files/publication/85716/guide-to-the-social-genome-project_2.pdf

[cx] Bronfenbrenner, U. 1979. The ecology of human development: experiments by nature and *design.* Cambridge, MA: Harvard University Press.

[cxi] Acs, G., Martin, S., and Schwabish, J.A. (2016). *Guide to the social genome project*. Washington, DC: Urban Institute. Retrieved from https://www.urban.org/sites/default/files/publication/85716/guide-to-the-social-genome-project_2.pdf

[cxii] Acs, G, Martin, S. Schwabish, J.A. and Sawhill, I.V. (2016). The Social Genome Model: Estimating How policies affect outcomes, mobility and inequality across the life course. *Journal of Social Issues* 72 (4): 656–75.

[cxiii] Sawhill V., Isabel and Richard V. Reeves. (2016). Modeling equal opportunity. *The Russell Sage Foundation Journal of the Social Sciences 2*(2): 60–97

[cxiv] Acs, G., Martin, S., and Schwabish, J.A. (2016). Guide to the social genome project. Washington, DC: Urban Institute. Retrieved from https://www.urban.org/sites/default/files/publication/85716/guide-to-the-social-genome-project_2.pdf

[cxv] Acs, G., Martin, S., and Schwabish, J.A. (2016). Guide to the social genome project. Washington, DC: Urban Institute. Retrieved from https://www.urban.org/sites/default/files/publication/85716/guide-to-the-social-genome-project_2.pdf