U.S. Department of Education
June 2018

The Investing in Innovation Fund: Summary of 67 Evaluations

Final Report

Beth Boulay
Barbara Goodson
Abt Associates

Rob Olsen
Rob Olsen, LLC

Rachel McCormick
Catherine Darrow
Michael Frye
Katherine Gan
Eleanor Harvill
Maureen Sarna
Abt Associates

INSTITUTE OF EDUCATION SCIENCES

NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE

This page has been left blank for double–sided copying.

The Investing in Innovation Fund: Summary of 67 Evaluations

Final Report

June 2018

Beth Boulay
Barbara Goodson
Abt Associates

Rob Olsen
Rob Olsen, LLC

Rachel McCormick
Catherine Darrow
Michael Frye
Katherine Gan
Eleanor Harvill
Maureen Sarna
Abt Associates

Tracy Rimdzius
Project Officer
Institute of Education Sciences

NCEE 2018-4013
U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

U.S. Department of Education

Betsy DeVos

Secretary

Institute of Education Sciences

Mark Schneider

Director

National Center for Education Evaluation and Regional Assistance

June 2018

This report was prepared for the Institute of Education Sciences under Contract ED-IES-10-C-0064. The project officer is Tracy Rimdzius in the National Center for Education Evaluation and Regional Assistance.

IES evaluation reports present objective information on the conditions of implementation and impacts of the programs being evaluated. IES evaluation reports do not include conclusions or recommendations or views with regard to actions policymakers or practitioners should take in light of the findings in the report.

This report is available on the Institute of Education Sciences website at http://ies.ed.gov/ncee.

Alternate Formats: Upon request, this report is available in alternate formats, such as Braille, large print, audiotape, or computer diskette. For more information, please contact the Department's Alternate Format Center at 202-260-9895 or 202-205-8113.

Table of Contents

List of Exhibits

Acknowledgements

We gratefully acknowledge the contributions of many individuals in conducting the study and producing the report.

We greatly appreciate the efforts of the evaluators that participated in the data collection for this study, as well as the grantees who implemented the interventions described in this report. Their contributions were essential to this report.

We are fortunate to have the advice of our expert technical working group members: Thomas Cook, Northwestern University; David Cordray, Vanderbilt University; Neal Finkelstein, WestEd; Lawrence Hedges, Northwestern University; Carolyn Hill, Georgetown University; David Francis, University of Houston; Robert Granger, William T. Grant Foundation; Brian Jacob, University of Michigan; Christopher Lonigan, Florida State University; Vivian Tseng, William T. Grant Foundation; and Russ Whitehurst, The Brookings Institution.

The report and the technical assistance described therein would not have been possible without the contributions of many staff at Abt Associates, American Institutes for Research (AIR), ANALYTICA, Century Analytics, Mike Puma, LLC, and Westat. At Abt, we thank the following current and former staff: Sarah Ballinger, Michelle Blocklin, Erin Bumgarner, Linda Caswell, Sarah Costelloe, Samuel Dastrup, RJ de la Cruz, Hayley Didriksen, Carter Epstein, Judy Geyer, Joseph Harkness, Yeqin He, Kerry Hofer, Marina Kosareva, Jeff Lambart, Katheleen Linton, Daniel Litwok, Erin Miles, Haleigh Miller, Hannah Miller, Sean Morris, Marc Moss, Shawn Moulton, Audra Nakas, Hiren Nisar, Krista Olsen, Amanda Parsad, Marissa Personette, Allan Porowski, Cristofer Price, Radha Roy, Sarah Sahni, Alex Silverman, Djaniele Taylor, Joe Taylor, Fatih Unlu, Jessica Walker, Sandra Wilson, and Anne Wolf. At AIR, we thank Xinsheng (Cindy) Cai, Lisa Hoogstra, Burhan Ogut, and Kwang Yoon. At ANALYTICA, we thank Herb Turner, Michele Munoz-Miller, and Mackson Ncube. At Century Analytics, we thank Bruce Randel. At Chesapeake Research Associates/Mike Puma, LLC, we thank Geoffrey Borman, Mike Puma, and Bob St. Pierre. At Westat, we thank Eva Chen, Jill Feldman, Jen Hamilton, Jill Lammert, and Allison Meisch.

1. The Investing in Innovation (i3) Fund

The Investing in Innovation (i3) Fund was established by the American Recovery and Reinvestment Act of 2009, which called on the U.S. Department of Education (ED) to provide student achievement awards to local education agencies and nonprofits with a record of improving student achievement.[1] ED set priorities for the program and specified them in the program regulations and the invitation for applications.[2,3] The program awards competitive grants to implement and evaluate educational interventions.

The i3 program aligns the amount of funding awarded to the prior evidence on and the expected implementation scale of the proposed intervention. The smallest, or "Development," grants support developing and testing interventions with limited or no prior evidence. Interventions with moderate evidence of effectiveness can receive a larger "Validation" grant, to implement and test the intervention in a broader population or in new contexts. The largest, or "Scale-up," grants support interventions with strong prior evidence of effectiveness to be implemented and tested on a much larger scale. The differential funding and expectations for the three types of grants means that, in effect, i3 operates like three grant programs in one.[4]

ED's Office of Innovation and Improvement (OII) administers the i3 program. The i3 program requires grantees to conduct an independent evaluation to build evidence and identify effective educational interventions. OII recognized that credible evidence is needed to identify effective interventions but is challenging to produce. To address this challenge, OII asked the Institute of Education Sciences (IES) to provide support for planning and conducting strong evaluations.

IES contracted with Abt Associates Inc. and its partners (the Abt Team) to conduct the i3 Technical Assistance and Evaluation Project. The Abt Team's two key activities were (1) providing technical assistance to support the design and conduct of strong evaluations, and (2) assessing the strength of and summarizing the findings from the i3 evaluations. This project built on prior work supporting grants with a requirement to conduct rigorous evaluations.[5]

---

1 American Recovery and Reinvestment Act of 2009, Section 14007, Title XIV (Public Law 111-5). http://www2.ed.gov/policy/gen/leg/recovery/statutory/stabilization-fund.pdf

2 U.S. Department of Education. (2010, March 12). Office of Innovation and Improvement; Investing in Innovation Fund; Final Rule and Notice. *Federal Register, 75(48):* 12004-12071. https://www.gpo.gov/fdsys/pkg/FR-2010-03-12/pdf/2010-5147.pdf

3 U.S. Department of Education. (2010, March 12). Office of Innovation and Improvement; Overview Information: Investing in Innovation Fund; Notice Inviting Applications for New Awards for Fiscal Year (FY) 2010. *Federal Register, 75*(48): 12072-12086. https://www.gpo.gov/fdsys/pkg/FR-2010-03-12/pdf/2010-5139.pdf

4 In competitions after Fiscal Year 2010, the i3 program issued multiple Notices Inviting Applications for the three tiers, which is consistent with the characterization of the program as three programs in one.

5 Boulay, B., Goodson, B., Frye, M., Blocklin, M., and Price, C. (2015). *Summary of research generated by Striving Readers on the effectiveness of interventions for struggling adolescent readers* (NCEE 2016-4001). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/pubs/20164001/

This first report of the i3 Technical Assistance and Evaluation Project summarizes the strength of and findings from the 67 i3 evaluations that were completed by May 2017. In addition, this report summarizes the extent to which the i3 program met its goal of contributing to the evidence base on effective educational interventions. Below, we describe the i3 program in more detail.

1.1 The i3 Program Logic Model

The Abt Team developed a logic model to describe the program and provide an organizing framework for this report. The i3 program logic model in Exhibit 1.1, therefore, is our interpretation of the key aspects of the i3 program. We describe the i3 program as based on a theory of action that includes a set of *priorities*, which lead to the implementation of the *key program activities*, in order to meet the intended *short-* and *long-term goals* of the program.

1.1.1 i3 Program Priorities

We identified three i3 program priorities from the program regulations: (1) fund tiers of grants, (2) require independent evaluations, and (3) fund a broad portfolio of educational interventions. We describe the implementation of the i3 program priorities in more detail in Chapter 2.

Fund tiers of grants. The i3 program awards its largest grants to the interventions with the strongest prior evidence of effectiveness and smaller grants to programs with little or no prior evidence. To date, the i3 program has awarded more than $1.4 billion to 172 grantees in three tiers:

- $340 million (24 percent) to 11 Scale-up grants,
- $702 million (49 percent) to 46 Validation grants, and
- $386 million (27 percent) to 115 Development grants.

Require independent evaluations. The i3 program requires that each grantee fund an independent evaluation that estimates the impact of the i3-funded intervention as implemented at the proposed level of scale. ED provided additional details about expectations for these evaluations in the program's performance measures and selection criteria.[6] The grantees were also required to cooperate with ED's technical assistance and to make the results of their evaluation publicly available.

Fund a broad portfolio of interventions. Unlike many ED grant programs that focus on a specific content area, such as reading, i3 funds a broad portfolio of educational interventions that serve high-need students in kindergarten through grade 12. Each year, the i3 program selects absolute priorities. In FY 2010, for example, the absolute priorities were to fund innovations that (1) supported effective teachers and principals, (2) improved the use of data, (3) complemented the implementation of high standards and high-quality assessments, and (4) turned around persistently low-performing schools. Selecting absolute priorities allows the i3 program to focus within the overall broad scope of education issues for high-need elementary and secondary students.

---

[6] U.S. Department of Education (2010), pp. 12080-12086.

1.1.2 Key Program Activities

Three of the six key program activities are related to the required evaluations and are shown in the white boxes in Exhibit 1.1.[7] The Abt Team conducted these three evaluation-related program activities under contract to IES, as summarized below and in more detail in Chapter 3.

Adopt or develop criteria to assess the strength of the i3 evaluations. A short-term goal of the i3 program is to produce strong evaluations. Assessments of whether the evaluations meet the short-term goal of i3 rest on establishing clear criteria against which to assess the evaluations. The Abt Team adopted existing criteria and developed additional criteria to assess the strength of the i3 evaluations. These criteria formed the basis for (1) the Abt Team's periodic review and assessment of the design and conduct of the i3 evaluations, and (2) the provision of technical assistance to evaluators.

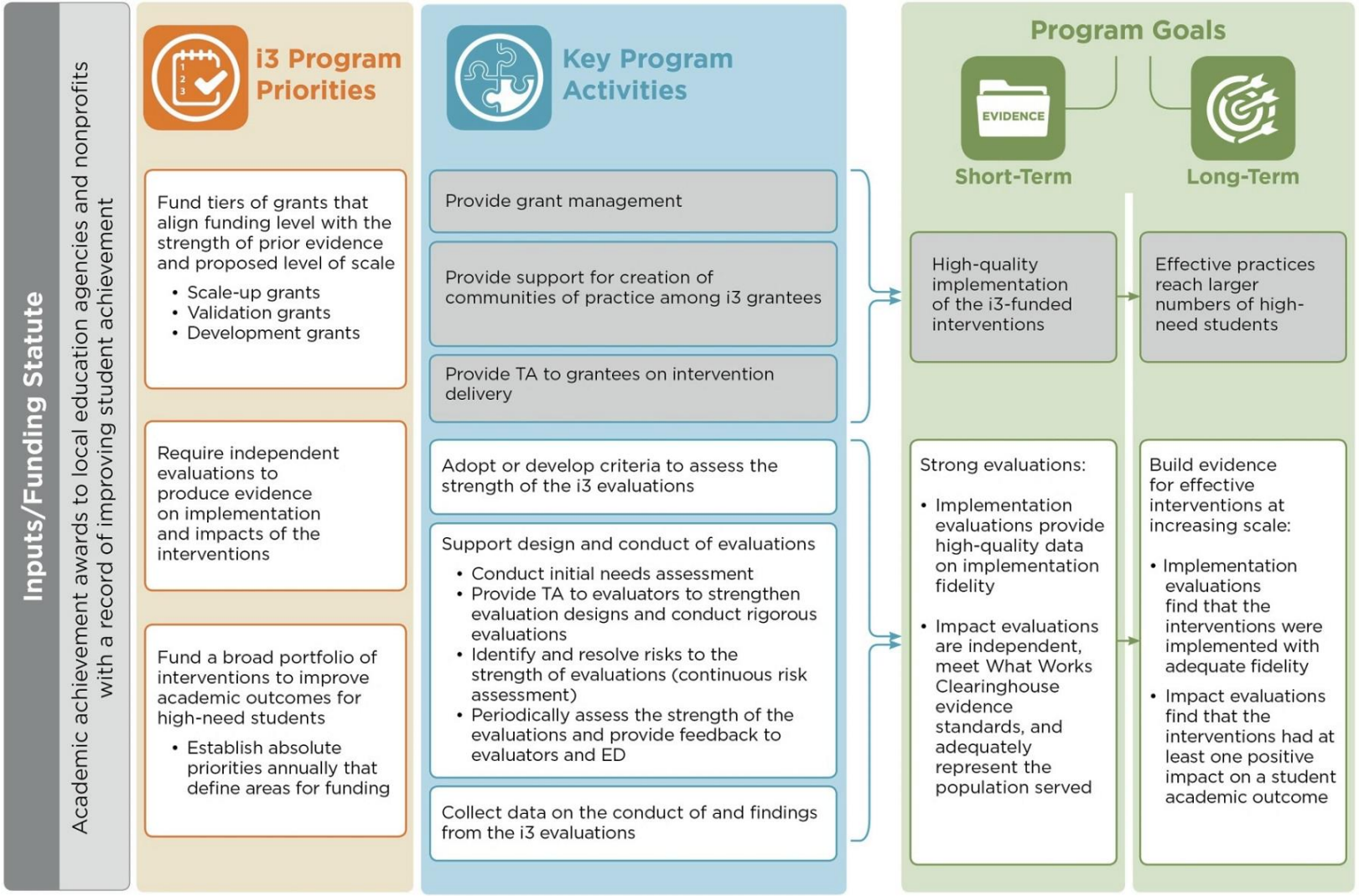
Support design and conduct of evaluations. The Abt Team provided comprehensive technical assistance to ensure that the i3 evaluators used the strongest possible evaluation designs to determine the impacts of the i3 interventions and provide high-quality implementation data. The Abt Team provided ongoing technical assistance to help evaluators identify and mitigate risks to the strength of their evaluations. Finally, Abt Team members, separate from those delivering technical assistance, periodically reviewed the strength of the evaluations. Abt Team reviewers assessed whether the proposed evaluation designs would meet the criteria for high-quality implementation evaluations and strong impact evaluations. Their assessments informed the technical assistance and were communicated to both the evaluators and ED.

Collect data on the conduct of and findings from the i3 evaluations. The Abt Team conducted a systematic data collection of information needed to (1) summarize the findings from the completed evaluations, and (2) put the findings in the context of the strength of the evaluations that produced them. Evaluators participated in the data collection near the end of the grant period. We focused our data collection efforts on findings that were the result of analyses that the evaluators had prespecified; that is, analyses they had committed in advance to conducting and reporting to the Abt Team. The prespecified analyses were the focus of both our technical assistance and our periodic reviews of the strength of the evaluations, and they are the focus of this report's summary of findings.

---

[7] The three key program activities shown in the shaded boxes in the middle panel of Exhibit 1.1 were conducted by i3 program staff or their contractors: provide grant management, provide support for creation of communities of practice, and provide technical assistance on the delivery of the interventions. Summarizing these activities is beyond the scope of this report.

Exhibit 1.1: i3 Program Logic Model



Note: The Key Program Activities in the shaded boxes were provided by OII and their contractors. Those activities are expected to lead to the Program Goals in the shaded boxes. Describing the activities and goals in the shaded boxes is beyond the scope of this report.

1.1.3 i3 Program Goals

The i3 program aims to build the evidence base on effective educational interventions.[8] The short-term goal of the i3 program is therefore strong evaluations that provide high-quality data on implementation fidelity and independent impact evaluations that meet What Works Clearinghouse™ (WWC) Standards[9] and adequately represent the population served by the intervention. The long-term goal of i3 is to identify effective interventions that can be implemented with adequate fidelity at increasing scale.

1.2 Roadmap to this Report

The remainder of this report describes the implementation of the i3 program priorities and key program activities and summarizes the findings from the 67 i3 evaluations completed by May 2017 to assess the extent to which the i3 program achieved its goals.

Chapter 2 describes the implementation of the i3 program priorities, and describes the interventions that ED funded grantees to implement and evaluate.

Chapter 3 describes the implementation of the three i3 key program activities that were related to the i3 evaluations.

Chapter 4 describes the extent to which the i3 program achieved its short-term goal by describing the strength of the i3 evaluations that have been completed by May 2017. It also describes the extent to which the i3 program met its long-term goal of identifying effective interventions that can be implemented with adequate fidelity (as defined by the grantees and evaluators).

Chapter 5 discusses the extent to which the four Scale-up grantees met their self-identified goals for scaling their interventions.

Finally, Chapter 6 offers lessons learned for other grant programs considering a rigorous evaluation requirement.

---

[8] The program goals shown in the shaded boxes in the right-hand panel of Exhibit 1.1 are the expected outcomes of the key program activities delivered by the i3 program staff and their contractors; summarizing these goals is beyond the scope of this report.

[9] It is important to note that only reviews conducted under a WWC contract can produce an official WWC rating for a study. Thus, the evidence assessments we conducted for this report are not a substitute for a WWC review, and we can only give an evaluation an unofficial WWC rating (see Section 4.1.3 for more details).

2. Implementation of the i3 Program Priorities

The i3 program can be characterized as having three priorities (see Exhibit 1.1): (1) fund tiers of grants, (2) require independent evaluations of the i3-funded interventions, and (3) fund a broad portfolio of interventions that aim to improve academic outcomes for high-need students. In this chapter, we describe the implementation of the three i3 program priorities in the sample of 67 i3 grants completed by May 2017. The i3 program funded grants through Fiscal Year (FY) 2016, some of which will not end until December 2021. While this sample of 67 grants might not be representative of all i3 grants, the findings in this report are informative about the i3 program in its early years.[10]

2.1 Fund Tiers of Grants

The 67 grants included in this report were awarded a total of $679 million, or 49 percent of the total i3 funds awarded to date. This includes:

- **$195 million (29 percent of the $679 million) to four Scale-up grants, which could receive up to $50 million each.** These constitute four of the 11 Scale-up grants (36 percent) awarded by the i3 program to date.
- **$292 million (43 percent) to 15 Validation grants, which could receive up to $25 million each in FY 2010 and up to $15 million each in FY 2011 and FY 2012.** These constitute 15 of the 46 Validation grants (33 percent) awarded by the i3 program to date.
- **$192 million (28 percent) to 48 Development grants, which could receive up to $5 million each in FY 2010 and up to $3 million each in FY 2011 and FY 2012.** These constitute 48 of the 115 Development grants (42 percent) awarded by the i3 program to date.

By design, the i3 program awarded a large share of the funds (almost one-third) to a small number of Scale-up grantees that proposed to implement interventions supported by strong prior evidence on a large scale. Many more Development grants were awarded a similar share of the funds to further develop and test the effectiveness of interventions with little or no prior evidence of effectiveness.

2.2 Require Independent Evaluations

The i3 program requires that each grantee propose and conduct an independent evaluation of the i3-funded intervention. To meet this requirement, the grantees named a separate evaluation team in their applications. The grantees tended to hire external organizations to conduct their evaluations, thereby laying the groundwork for independence between those implementing the intervention and those evaluating its effectiveness. The Abt Team's review of the grant applications determined that for all 67 grants, the proposed evaluator was affiliated with a different organization than the organization that would be responsible for developing and implementing the intervention.

[10] ED plans to release updated versions of the tables and figures included in this report to summarize the strength of and findings from i3 evaluations completed after May 2017.

The evaluators hired by the 67 grantees in this report include a range of organizations, from individual consultants to large national research firms. Among the grantees, 24 chose national research firms as their evaluation partner, 22 chose smaller research firms or individual consultants, and 21 grantees chose to work with a university partner.

Of the $679 million awarded to these grantees, $99 million (15 percent) was spent on evaluation. On average, grantees spent 16 percent of their grant award on evaluation, with the evaluation budgets ranging from 3 to 38 percent of the grants. On average, Scale-up grantees spent approximately $4 million on evaluation (8 percent of the grant award), Validation grantees spent approximately $3.5 million on evaluation (17 percent of the grant award), and Development grantees spent approximately $645,000 on evaluation (16 percent of the grant award).

The grantees are required to make the results of their evaluations publicly available. As of March 2018, 47 of all 172 grants funded (27 percent) have produced public reports. Of the 67 completed evaluations in this report, 46 (69 percent) have produced public reports.[11]

2.3 Fund a Broad Portfolio of Interventions

The i3 program established a set of absolute priorities that describe educational barriers to the achievement of high-need students. The 67 completed grants included in this report identified in their applications the priority that their proposed intervention addresses. Because the grants in this report span the FY 2010, FY 2011, and FY 2012 funding years, the absolute priorities that the grantees identified include absolute priorities from each of the three years (Exhibit 2.1).

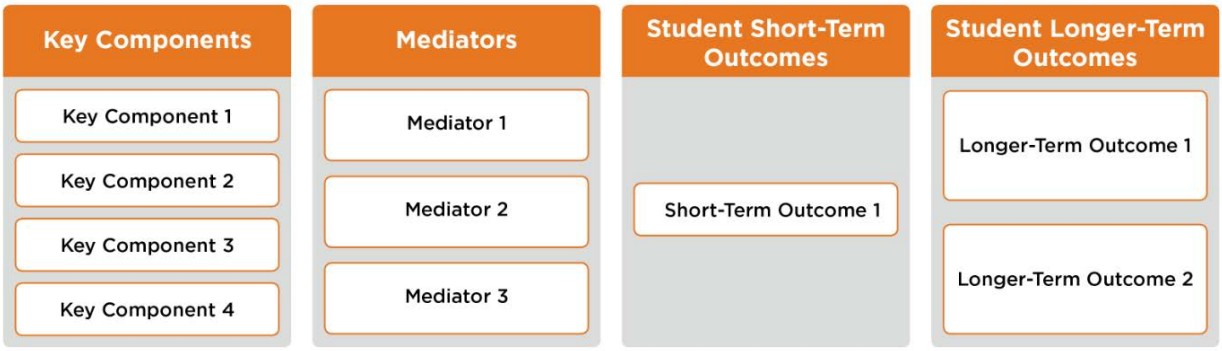
Exhibit 2.1: i3 Program Absolute Priorities Identified by the 67 Completed Grants

Absolute Priority	Grant Year	Scale-up Grants (n=4)	Validation Grants (n=15)	Development Grants (n=48)	All Grants (n=67)
Innovations that Improve the Effectiveness and Distribution of Effective Teachers or Principals	FY 2010, FY 2011	2	3	10	15
Innovations that Turn Around Persistently Low-Performing Schools	FY 2010, FY 2011, FY 2012	2	3	10	15
Innovations that Complement the Implementation of High Standards and High-Quality Assessments	FY 2010, FY 2011	0	5	14	19
Innovations that Improve the Use of Data	FY 2010	0	2	6	8
Innovations that Promote Science, Technology, Engineering, and Mathematics (STEM) Education	FY 2011	0	1	4	5
Improving Achievement and High School Graduation Rates in Rural Educational Agencies	FY 2011, FY 2012	0	1	2	3
Improving School Engagement, School Environment, and School Safety and Improving Family and Community Engagement	FY 2012	0	0	2	2

[11] The i3 program lists the available reports here: https://innovation.ed.gov/what-we-do/innovation/investing-in-innovation-i3/awards/

Each grantee provided a narrative description of the intervention proposed in their applications. However, only 39 percent included a detailed logic model depicting the theory by which the intervention is expected to improve student academic outcomes. Only 15 percent described existing methods for assessing fidelity of implementation. As part of our technical assistance (see Section 3.2 below), the Abt Team worked with evaluators to develop an intervention logic model (or to revise and expand an initial logic model) that shows how the intervention's ***key components*** (i.e., the primary activities conducted by those implementing the intervention) are expected to change ***mediators*** (i.e., behaviors, processes, and skills) through which the intervention was expected to ultimately have an impact on ***short-term student outcomes*** and ***longer-term student outcomes***. The framework for the logic models was provided by the Abt Team and took the form shown in Exhibit 2.2. The number of boxes in each column reflects the most common number of each element across the 67 intervention logic models.[12]

Exhibit 2.2: Generic i3 Intervention Logic Model



The logic models developed by the grantees and evaluators include a large number of different key components and mediators, as well as different short- and longer-term student outcomes. The logic models for the 67 interventions summarized in this report specify 329 key components, 414 mediators, 151 short-term student outcomes, and 221 longer-term student outcomes.

The Abt Team coded each element of the logic models to summarize the variety of approaches the interventions took to improve student academic achievement. The coding scheme included 11 categories of key components, 3 categories each of mediators and short-term student outcomes, and 4 categories of longer-term student outcomes. The following sections summarize the Abt Team's coding of the 67 intervention logic models.

2.3.1 Key Components Implemented by the Intervention Developers

The i3 program funded a broad portfolio of interventions, so it is not surprising that the logic models included a range of both *individual* key components and *combinations* of key components. Most interventions specified four key components in their logic models; all but one of the interventions had at least two key components and nine had eight or more.[13] The number of key components ranged from 1 to 24. Nine of the 67 interventions had 8 or more key components.

---

[12] Equal numbers of grants included three or four mediators (nine grants each). The generic logic model displays the lower number of elements.

[13] One Scale-up intervention, KIPP, characterized its intervention with a single key component. See Appendix A.

The most common category of key components was Provide Professional Development, defined as direct training to school staff to improve practice (Exhibit 2.3). Eighty-five percent of the interventions included at least one professional development activity as a key component of the intervention. Of those, 29 interventions (51 percent) provided professional development solely to teachers, and 23 interventions (40 percent) provided professional development to both teachers and other school staff. A smaller number of interventions (5 interventions, 9 percent) provided professional development solely to other school staff (e.g., counselors, administrators) or non-school staff (e.g., mentors, local professionals).

Exhibit 2.3: Key Components Identified in the i3 Intervention Logic Models

Coaching (Provide Coaching) was categorized separately, as another mode of professional development, and was included in 22 percent of interventions. The only other category of key component that was included by a majority of interventions was Developing and Instituting New Curriculum and Materials into classrooms (60 percent of interventions).

There was a wide variety of specific program activities within each of the 11 categories of key components. For example, Institute Structural Changes includes activities such as introduction of flexible student scheduling, extending the school year, and establishing a student technology center in the school. Provide Services Targeting Individualized Learning includes activities such as individual student mentoring and individualized case management for students, especially for high-need students. (See Appendix A for descriptions of all 67 interventions and the key components named in their logic models.)

2.3.2 Mediators through which Interventions are Expected to Improve Short- and Longer-Term Student Outcomes

The logic models also describe the mediators that are the behaviors, processes, and skills targeted by each intervention's key components. Mediators are expected, in turn, to lead to improved student outcomes. The Abt Team coded the 414 mediators into one of three categories: (1) changes in school structures or processes (School Mediators), (2) changes in classrooms (Classroom Mediators), or (3) changes in families (Family Mediators).

Examples of school mediators include:

- development of schoolwide norms and processes for continuous improvement;
- use of a facilitative leadership approach by the principal;
- joint teacher and administrator analysis of formative assessments to support data-driven decision-making;
- development of school action plans; and
- increased learning time through extension of the school day or school year (e.g., summer instructional activities).

Examples of classroom mediators include:

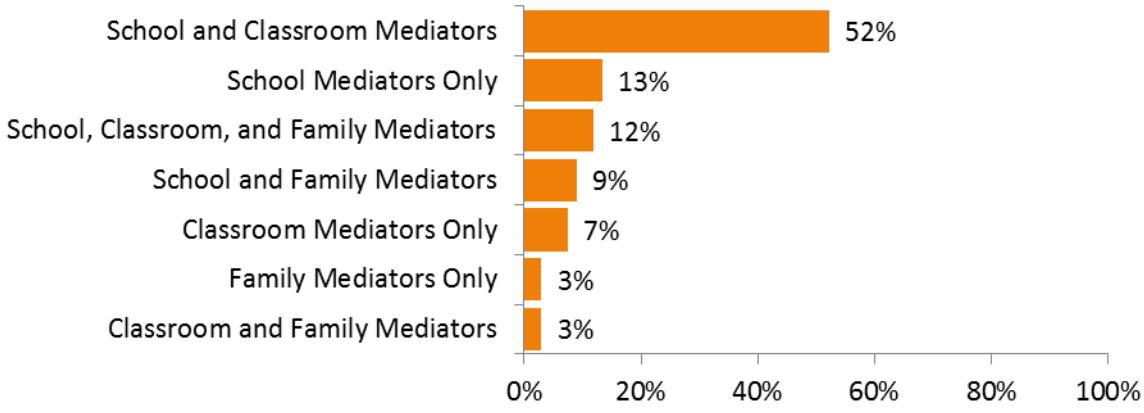
- integration of technology into science instruction;
- use of inquiry-based approaches for teaching content area skills;
- increased individualization of classroom instruction and better classroom management; and
- integration of the arts into reading and writing through a variety of art forms and literary genres.

Examples of family mediators include:

- increasing support for the child's education at home;
- increasing family involvement in school activities; and
- developing family-school partnerships.

Three-quarters (76 percent) of the intervention logic models specified mediators in more than one of the three categories (Exhibit 2.4). Slightly more than half of the intervention logic models included both school and classroom mediators (but not family mediators).

Exhibit 2.4: Mediators Identified in the i3 Intervention Logic Models



2.3.3 Short- and Longer-Term Student Outcomes

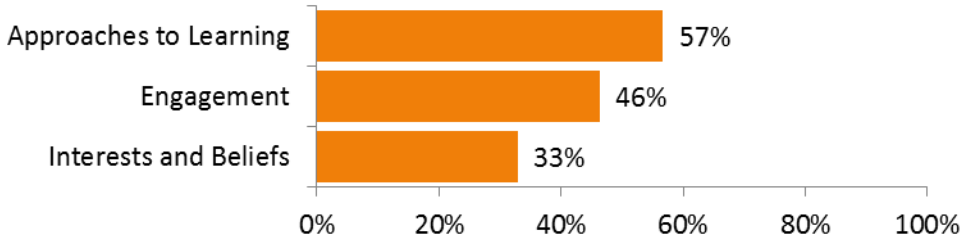
While all of the i3 interventions aimed to improve student academic outcomes, most also specified short-term (or intermediate) outcomes on the path to longer-term achievement and attainment outcomes (54

interventions, or 81 percent). The Abt Team coded the 151 short-term student outcomes in the 67 logic models into three categories:

- Approaches to Learning, including 21st-century skills, use of strategies/methods for learning and analysis, inquiry skills, and effective study habits;
- Engagement in school, including greater interest in school, engagement in school activities, more time on task, and better school attendance; and
- Interests and Beliefs, including higher college and career aspirations and greater confidence and sense of self-efficacy, and positive relations with other members of the school community.

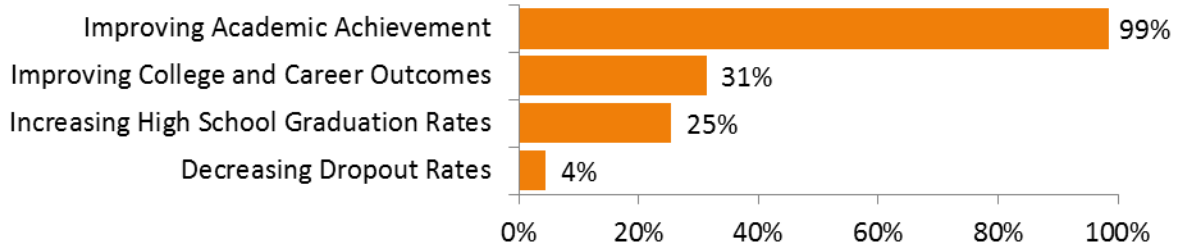
Approaches to Learning and Engagement were each identified as short-term student outcomes in approximately half of the 67 logic models (57 percent and 46 percent, respectively; Exhibit 2.5).

Exhibit 2.5: Short-term Student Outcomes Identified in the i3 Intervention Logic Models



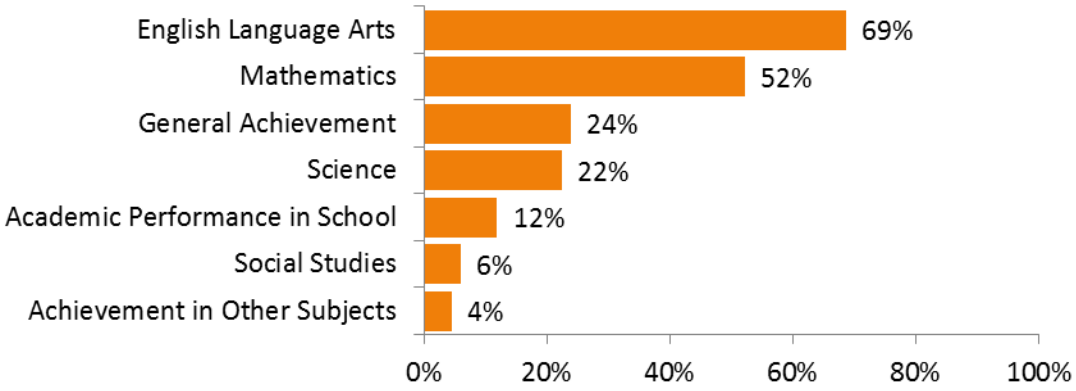
All of the i3-funded interventions were expected to aim to improve longer-term student achievement and attainment outcomes. The i3 intervention logic models included 221 longer-term outcomes that the Abt Team coded into one of the four domains of interest to the i3 program: (1) Improving Academic Achievement (including overall achievement or closing achievement gaps), (2) Decreasing Dropout Rates, (3) Increasing High School Graduation Rates (including indicators in earlier grades of being on track for graduating), and (4) Improving College and Career Outcomes (including college enrollment and readiness, completion rates, and labor market outcomes). As shown in Exhibit 2.6, all 67 interventions included Academic Achievement as a longer-term student outcome in their logic models. Around one-third of the interventions included Improving College and Career Outcomes, and one quarter included Increasing High School Graduation Rates.

Exhibit 2.6: Longer-term Student Outcomes Identified in the i3 Intervention Logic Models



Within the Academic Achievement category, the 67 interventions included outcomes in the core academic subjects (English Language Arts, mathematics, science, and social studies), as well as in other subjects such as art (Exhibit 2.7).

Exhibit 2.7: Types of Student Achievement Outcomes Identified in the i3 Intervention Logic Models



3. Implementation of the Three Key Program Activities Related to Evaluation

The federal investment in the i3 grants is substantial; consequently, ED aims to learn as much as possible from the i3 evaluations. ED expects the i3 grants to sponsor strong evaluations to build evidence on effective interventions. In addition to the funds awarded to grantees, ED invested approximately $30,000 per grant per year, or $10 million total, to support the comprehensive technical assistance provided to the 67 evaluations summarized in this report.[14] This level of investment in direct support for evaluation is unusual among grant programs and reflects ED's high expectations for what would be learned from the i3 program. It is also important to put the strength of the completed i3 evaluations (see Chapter 4) in the context of these substantial supports.

The three key program activities related to the i3 evaluations are (1) adopt or develop criteria to assess the strength of the i3 evaluations, (2) support the design and conduct of the i3 evaluations, and (3) collect data on the conduct of and findings from the i3 evaluations. This chapter summarizes how the Abt Team conducted these three activities.

3.1 Adopt or Develop Criteria to Assess the Strength of the i3 Evaluations

The Abt Team adopted or developed criteria to assess the strength of i3 evaluations. These criteria guided technical assistance we provided the evaluators. They also were used to provide periodic assessments of the strength of the evaluations to the grantees and ED. More specifically, the Abt Team:

- Adopted the WWC Evidence Standards (version 3.0) to assess the strength of the impact evaluations and developed a review protocol to ensure the consistent application of those standards to all i3 impact evaluations (see Appendix B), including those that do not fall under existing WWC review protocols;[15, 16]

---

[14] This cost estimate does not include collecting evaluation data from the grantees, the third key activity described in this section.

[15] It is important to note that the reviews reported here differ from an official WWC review in two important ways. First, the data available to the Abt Team differ from the data available to the WWC. The Abt Team collected the information necessary for the reviews directly from the evaluators and limited the summary of impact findings to those that were prespecified. The WWC limits its reviews to publicly available reports of evaluation findings. Second, only reviews conducted under a WWC contract can produce an official WWC rating for a study. Thus, the evidence reviews we conducted for this report are not a substitute for a WWC review, and we can only give an evaluation an unofficial WWC rating. For these reasons, the official WWC rating of the i3 evaluations can differ from the ratings reported here. We include our unofficial ratings in this report because meeting these standards is a key goal of the i3 program and because it is important to place this summary of evaluation findings in the context of the strength of the evaluations.

[16] The Abt Team also developed criteria to assess whether evaluations funded by Development grants provided evidence on the intervention's promise. In the FY 2010–FY 2014 competitions, impact evaluations supported by Development grants were expected to provide evidence on the promise of the intervention for improving student outcomes. Though evidence of promise can support further development and testing of the intervention, it does not support conclusions that the intervention caused the observed effects on student outcomes, making

- Developed criteria for assessing the *strength* of implementation data and performance feedback, including whether an implementation evaluation provided information on key features of the program for future replication or testing in the form of a comprehensive logic model and periodically measured fidelity of implementation of the key components of the intervention model;
- Developed criteria for whether an impact evaluation was conducted *independent* of those responsible for developing or implementing the i3 intervention; and
- Developed criteria for whether an impact evaluation sample was *representative* of those served by each i3 intervention.

The Abt Team developed definitions for the *strength* of an implementation evaluation, *independence*, and *representativeness*. The definitions were operationalized into criteria that were communicated to grantees and evaluators and that guided our delivery of TA aimed at strengthening the evaluations. The criteria represent the Abt Team's best judgment in developing a framework for assessing the evaluations along these dimensions.

The task of developing criteria for implementation evaluations was especially challenging. The field of implementation science has generated a range of approaches and best practices for measuring and summarizing the implementation of an intervention.[17] However, we are unaware of any attempts to codify best practices into a set of criteria that assesses whether a particular evaluation was well conducted. As indicated earlier, most of the grantees had blueprints for implementing their interventions but had not developed well-specified logic models or methods of measuring fidelity of implementation.[18] The Abt Team developed criteria that allowed grantees substantial latitude to specify what the key components of their interventions were, how the implementation of these components should be measured, and what qualified as adequate fidelity of each component. The Abt Team did not include measurement of implementation in the comparison group in the criteria because this was not one of ED's expectations, although we encouraged them to do so as good practice.

---

the evidence less useful to those considering adopting an intervention. In the FY 2015 competition and beyond, the selection criteria for Development grant awards included proposing an evaluation that would meet WWC Standards with Reservations and not just providing evidence on the promise of the intervention.

[17] Century, J., Cassata, A., Rudnick, M., & Freeman, C. (2012). Measuring enactment of innovations and the factors that affect implementation and sustainability: Moving toward common language and shared conceptual understanding. *Journal of Behavioral Health Services & Research, 39*(4), 343–361. doi: 10.1007/s11414-012-9287-x

Meyers, D. C., Durlak, J. A., & Wandersman, A. (2012). The quality implementation framework: A synthesis of critical steps in the implementation process. *American Journal of Community Psychology, 50,* 462–480.

Durlak, J. A. (2015). What everyone should know about implementation. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.), *Handbook of social and emotional learning*: *Research and practice* (pp. 395–405). New York: Guilford.

Abry, T., Hulleman, C. S., & Rimm-Kaufman, S. E. (2014). Using indices of fidelity to intervention core components to identify program active ingredients. *American Journal of Evaluation,* published online November 26. doi: 10.1177/1098214014557009

[18] Starting in the FY 2013 competition, ED added a selection factor on the extent to which the evaluation plan clearly articulates the logic model and measurable fidelity threshold.

The Abt Team's criteria assess whether the evaluation included:

- a logic model with key components of the intervention, at least one mediator, and at least one student outcome;
- a set of measurable indicators of implementation fidelity of the key components, a way to combine the indicators into a score for each key component, and a threshold that defines adequate fidelity; and
- the periodic measurement of fidelity.

The criteria assess only the *presence or absence* of these elements of the implementation evaluation, not their validity or quality. The Abt Team did not assess whether the key components accurately describe how the intervention is delivered, nor whether they constitute a full representation of the intervention model.

Each grantee/evaluator team established their own thresholds for adequate fidelity of each of the key components in their logic model. The Abt Team provided support, but ultimately the thresholds were established by each evaluator/grantee team and reflect their definition of what it means to implement the funded intervention with adequate fidelity. This resulted in substantial variation across grantees in the way they defined adequate fidelity.

Our criteria for implementation evaluations were developed in light of the initial needs assessment of the grantee plans (described in Section 3.2.1), which revealed that none of the evaluation plans included logic models or plans to systematically measure fidelity. We designed the criteria to incentivize evaluators to meet minimal standards for the presence of these elements, which would represent a substantial improvement over the implementation evaluations as proposed. Though there is a growing consensus in the field that evaluators should develop logic models and measure fidelity of implementation, there is disagreement about what constitutes valid measurement of and thresholds for adequate fidelity. Therefore, we developed criteria that reflect minimum standards for what is generally accepted in the field.

3.2 Support the Design and Conduct of the i3 Evaluations

Our technical assistance to evaluators included multiple activities with the common goal of supporting the i3 evaluators to design and conduct strong impact and implementation evaluations. All of the technical assistance provided to i3 evaluators was aligned with WWC Standards and the additional criteria developed by the Abt Team.

The technical assistance included the following series of activities:

- An initial assessment by the Abt Team of the quality of the evaluations proposed in the successful i3 applications;
- Comprehensive, customized technical assistance for each evaluator, guided by the initial needs assessment and aimed at strengthening the design and conduct of the evaluations;
- Ongoing identification and resolution of risks to the strength of the evaluations while they are being conducted; and
- Periodic assessments of the strength of the evaluations that are shared with both evaluators and ED.

The technical assistance activities are described in more detail below.

3.2.1 Conduct an Initial Assessment of the Need for Support

Each successful application for i3 funding described a plan for an external evaluation of the proposed intervention. The Abt Team conducted a systematic review of these evaluation plans to make an initial assessment of the evaluation's technical assistance needs. Our reviews revealed that none of the successful applications provided sufficient information to assess whether it was designed to meet WWC Standards. For example, most of the applications proposing quasi-experimental designs (QEDs; more than 80 percent) did not discuss how baseline equivalence would be established, which is necessary to meet WWC Standards. In addition, many of the proposed designs had significant weaknesses that needed to be addressed for the evaluations to be strong. For example, some proposed a randomized controlled trial (RCT) in which evaluators planned to exclude students from the analysis based on indicators measured after random assignment, such as attendance.

Our review also concluded that none of the successful applications provided enough information to assess the strength of the implementation evaluation. Only 39 percent of the applications included a logic model for the intervention, and only 15 percent described methods for assessing fidelity of implementation.

3.2.2 Provide Technical Assistance to Strengthen Design and Conduct of i3 Evaluations

The Abt Team used the initial assessment of the funded applications to develop a customized technical assistance plan for each of the i3 grantees' evaluators. The plan included activities for the initial 12-month design phase, during which the evaluators finalized the designs for their impact and implementation evaluations.

We delivered the technical assistance primarily through ***one-on-one calls*** between the evaluator and a technical assistance provider who was a WWC-certified reviewer and experienced in conducting rigorous evaluations in schools. The technical assistance provider served as the primary point of contact with the evaluator throughout the grant period, from design through analysis and reporting of results. The Abt Team also provided ***standardized tools and templates*** to help the evaluators track the progress of their evaluations (including key design decisions and milestones) and report their findings; and ***group technical assistance*** (including webinars and in-person sessions at the annual i3 Project Directors Meetings).

Across the 67 grants summarized in this report, the typical i3 evaluator received:

- 72 calls with a technical assistance provider—approximately monthly during the five-year grant period;
- 4 standardized tools/templates to document and track the progress of the evaluation and the analysis and reporting of findings for both the impact and implementation evaluations;
- 5 webinars on the criteria for and the conduct of strong evaluations, delivered during the first year of the evaluations; and 29 evaluation-related sessions at annual i3 Project Directors Meetings (2011–2015).

Each of the three technical assistance modes is described in more detail below.

One-on-One Evaluation Technical Assistance

One-on-one evaluation technical assistance was ongoing, frequent, and customized. Throughout the grant period, each evaluator had regular telephone calls with an evaluation technical assistance provider. The calls occurred twice a month during the first year of the grants, when evaluators were revising and

finalizing their evaluation plans, and monthly for the remainder of the grant period. The technical assistance provider customized the call agendas to the context, resources, key research interests, and needs of each evaluation. During the design year, one-on-one technical assistance focused on supporting evaluators in designing evaluations that could meet WWC Standards.

Technical assistance providers tracked and documented their one-on-one technical assistance with evaluators, which addressed a wide range of topics related to evaluation design. The topics most frequently addressed covered how different methodological decisions might affect assessments of the strength of the evaluation. For randomized studies, the most frequent issues were related to (1) the conduct of cluster RCTs, including how to determine whether a sample included joiners and ramifications for the evaluation's evidence rating; (2) allowable sample exclusions post-randomization; and (3) issues related to calculating the amount of sample loss (attrition) at the student and school levels. For QEDs, the most frequent issues were (1) selecting the best matching procedures; (2) selecting baseline measures for outcomes for which there were no natural pretests; and (3) establishing baseline equivalence in a way that was consistent with the WWC Standards. In addition, evaluators frequently asked for assistance calculating power for specific designs, particularly for cluster designs and designs using interrupted time series.

The most common issues affecting the assessment of the implementation evaluations were (1) developing a fidelity measure that measured *all* of the key components of the intervention that were shown in the logic model; (2) establishing thresholds at the sample level for each key component; and (3) defining a relevant sample and schedule for measuring fidelity and reporting of findings that aligned with ED's expectations.

Defining fidelity measures and establishing thresholds for adequate fidelity proved particularly challenging for the grantees and evaluators, because they had limited, if any, objective data on the extent to which different levels of the key components were necessary to achieve outcomes. Further, many evaluators recognized that the presence or amount of a key component was not by itself sufficient, and they wanted to measure the quality of the delivery of the key component. However, operationalizing and measuring "quality" of implementation for inclusion in fidelity measures were beyond the scope of the implementation evaluations. When included, quality was most often measured through participant self-report, such as teacher perceptions of the usefulness of the training.

The following examples are the Abt Team's descriptions of the approaches evaluators took to developing measures of implementation fidelity and setting their own thresholds for adequate fidelity.

> The Children's Literacy Initiative used its Validation grant to implement its early literacy model in 38 elementary schools. The logic model for the intervention included five key components. The grantee defined the same threshold for fidelity of implementation for each of the components as 80 percent or more of the treatment schools fully implementing the component. Full implementation was defined in terms of a set of indicators for each component. As an example of how the grantee defined fidelity of implementation, consider its key component of "teacher professional development": for this component, *full implementation* for an individual teacher was defined as attendance at the summer training institute and attendance at three seminars during the year. At the school level, *full implementation* was defined as at least 90 percent of teachers demonstrating full implementation. At the sample level, the *threshold for fidelity of implementation* was at least 80 percent of schools demonstrating full implementation.

School of One leverages a *learning progression* based on the discrete set of skills students must master and research-based evidence on the relationships among those skills. The model integrates available data about each student, administers an additional diagnostic instrument, and based on that information, creates a *playlist* for each student specifying a unique set of skills to focus on over a period of time. The *instructional content* comes from publishers, software providers, and other educational groups across nine instructional modalities (life instruction, live reinforcement of prior lessons, live tutoring, small group collaboration, independent practice, virtual computerized instruction, virtual live instruction, virtual live tutoring and homework). A *learning algorithm* captures and analyzes the data from each lesson and recommends to teachers a unique daily schedule for each student that they can adjust as necessary. At the end of each day, students take a unique assessment (the *playlist update*) to measure mastery of the skill they studied. Fidelity of the model includes three key components: These three key components are: (1) Operational Aspects (space and training for teachers), (2) Required Resources (computers, bandwidth, written materials) and (3) Sufficient Time (instructional time for students and planning and professional development time for staff). The three key components are defined by 12 indicators, differentially weighted in the score for the key components by virtue of being scored on a one-point, two-point, three-point, or five-point scale. A total Component Implementation Score is calculated for each key component, and, based on the score, a school is categorized as Low, Medium, or High implementation, based on thresholds established for each component. For all three key components, fidelity of implementation of the component at the program level is "70% of schools in the implementation sample had a score indicating High implementation of the indicators for that component". For example, high fidelity for the key component of *Required Resources* is defined by meeting the threshold on three indicators (a) at least 75% of school days with a sufficient number of computers to accommodate all student schedules; (b) % of school days with sufficient broadband and network functionality, and (c) % of school days with sufficient supply of student forms/ materials prepared. High fidelity on the summary Component Implementation Score across the three indicators is a score of 3 and fidelity at the program level is defined as at least 4 of the 6 schools receive a score of 3.

PTA Comunitario seeks to establish successful partnerships between community-based organizations and schools with minority and low-income families in order to increase student achievement, particularly of students who are low-income, Hispanic, and/or limited English proficient (LEP). The project seeks to establish PTA Comunitarios; establish a partnership between PTA Comunitarios and the local school district through training and technical assistance; and facilitate collaborative educational leadership projects for PTA Comunitarios and partner schools. The model incorporates four key components: (1) Community Organizing and Mobilization; (2) Initiation and Relationship-Setting; (3) Taking Collaborative Action for Student Success, Education Leadership Projects; and (4) Instituting Sustainable Connections. These four components are defined by nine indicators, each scored on a two-point or a three-point scale. The total score across indicators for each key component is converted into a Component Implementation Score (Low, Moderate, or High). Depending on the component, fidelity of implementation of the component at the program level was defined as "80% of the 24 PTA Comunitarios (partnered with 24 schools) had a Component implementation Score = High". For example, fidelity for the key component of *Collaborative Educational Leadership Projects* is defined by (a) whether or not parents in the Comunitario develop a project that addresses a need identified by the school (0 or 1); (b) whether at least 80% of parents in the Comunitario participate at least monthly in the project (0 or 1); (c) whether at least 80% of parents in the Comunitario rate the project as effective (score on a 5-point scale) in addressing the school need (score of 0 if rating – 0, score of 1 if rating = 1-3, score of 2 if rating - 4 – 5). Adequate fidelity on the summary Component Implementation Score across the three indicators is a score of 4, and fidelity at the program level is defined as at least 20 of the 24 PTA Comunitarios receiving a score of 4.

Standardized Tools and Templates

The Abt team developed four standardized tools and templates for the i3 evaluators: (1) the Study Design Template to fully describe impact and implementation evaluation design plans, with embedded supports and links to relevant WWC Standards and i3-specific criteria; (2) the Contrast Tool to document each of the planned impact analyses, including research questions, samples, outcomes and baseline measures, and timing of the analysis; (3) the Fidelity Measure Tool to describe the methodology for measuring the fidelity of implementation of the key components of the intervention's logic model; and (4) Reporting Templates that include all the information necessary to review the strength of and summarize the findings from the evaluations to guide evaluators in drafting their public reports.

Group Technical Assistance

The Abt Team provided group technical assistance to evaluators on technical issues that were common across the i3 evaluations, using webinars and conference presentations. This group technical assistance comprised more than 100 presentations focused on many of the same topics as the one-on-one technical assistance—WWC Standards, i3-specific criteria for providing evidence on the promise of the intervention, i3 criteria for implementation evaluations, and evaluation design topics such as multivariate matching methods and calculating statistical power.

3.2.3 Identify and Resolve Risks to the Strength of the Evaluations

The Abt team identified potential risks to the strength of the evaluations that arose during design and during the conduct of the evaluations. The Abt Team developed both a categorization of different types of risks and a process for identifying and monitoring the risks. Risks included the following:

- Issues that could affect the strength of the impact evaluation (e.g., non-random movement of evaluation participants after random assignment, or a QED without acceptable baseline measures);
- Issues that could affect the interpretation of impact findings (e.g., low power to detect effects of the intervention);
- Issues that could affect whether the implementation evaluations could provide high-quality data (e.g., absence of a systematic measure of the fidelity of a key component or a threshold for adequate fidelity); and
- Issues that threatened the independence of the evaluation (e.g., the intervention developer planned to collect student outcome data).

The Abt Team reported these risks to ED along with recommended actions to resolve the risk, as applicable. During the conduct of the 67 evaluations included in this report, the Abt Team reported 136 risks to ED. Of the 67 evaluations summarized in this report, 46 evaluations (69 percent) had at least one risk reported. Of these 136 risks, 65 (48 percent) were resolved by the evaluator and 71 (52 percent) were documented as permanent.

3.2.4 Periodic Assessments of the Strength of the Evaluations and Reports to the Evaluators and ED

As described in Section 3.2.1, the evaluation plans included in the applications were not detailed enough to assess the strength of the evaluation designs or to plan technical assistance activities. Therefore, the

Abt Team asked evaluators to submit a complete evaluation plan late in the first grant year (October of the year following grant award).[19] We provided evaluators with templates for both impact and implementation evaluation designs (those described in Section 3.2.2) to ensure the plans contained all of the elements necessary to assess the strength of evaluations.

Abt Team members who were all WWC-certified reviewers, but separate from those providing technical assistance, conducted systematic reviews of the strength of each evaluator's impact and implementation evaluation designs.

This review informed evaluators:

- in advance how their evaluations would be reviewed and the consequences of different design decisions for their potential evidence rating; and
- about aspects of their designs that represented risks to the potential for their studies to meet WWC Standards.

The Abt Team communicated the results of our reviews to evaluators in detailed feedback memos, recommending revisions to the designs that could address these risks. We encouraged them to work with their technical assistance provider to revise their evaluation plans to address the issues and resubmit their revised evaluation plans for a second (or third) round of review. Our re-reviews assessed the implications of the changes for the potential evidence rating. Evaluators could submit changes to their plans until they began their data collection.

Once the evaluation plans were finalized, the Abt Team's review of the evaluation plans revealed that:

- 54 of the 67 impact evaluations (81 percent) were designed with the potential to Unofficially Meet WWC Standards with or without Reservations.
- All 19 Scale-up and Validation evaluations and 35 of the 48 Development evaluations (73 percent) were designed with the potential to Unofficially Meet WWC Standards with or without Reservations.
- 65 of the 67 implementation evaluations (97 percent) were designed to provide high-quality data on implementation fidelity.

3.3 Collect Data on the Conduct of and Findings from the i3 Evaluations

The Abt Team was tasked with summarizing the findings from the completed i3 evaluations, and for putting the findings in the context of the strength of the evaluations (see Chapter 4). Therefore, we conducted a systematic data collection of information about the conduct of and findings from the 67 completed evaluations.

In this section, we discuss how we (1) collaborated with evaluators to **prioritize** the impact evaluation findings, (2) identified the findings from **prespecified analyses** from among all the impact findings reported, and (3) **collected** the findings of the impact and implementation evaluations.

[19] ED added the submission of a revised evaluation plan as a grant requirement in the cooperative agreements in FY 2011 and beyond.

3.3.1 Prioritizing the Impact Evaluation Findings

The i3 impact evaluation plans specified by the evaluators were often ambitious. In addition to their plans to estimate the impact of the intervention on student academic outcomes, most evaluators (48 of 67, or 72 percent) also included other analyses to, for example, estimate the impact of the intervention on teacher and school outcomes (see Section 2.3.2 for examples) or on other student outcomes (see Section 2.3.3 for examples).

Evaluations that plan to estimate impacts of the intervention on many outcomes are at increased risk for producing a false positive or negative impact. That is, the more impacts estimated by the evaluation, the higher the likelihood that any one analysis will yield a statistically significant finding simply by chance. The Abt Team provided guidance to evaluators on how to minimize this risk by prioritizing their impact analyses and prespecifying a subset of analyses that they would use to assess whether or not the intervention was effective (their "confirmatory" analyses).[20] Our guidance was aimed at helping evaluators understand how standard methods for accounting for the risk of conducting multiple analyses would affect the interpretation of their findings, while at the same time acknowledging that the evaluators were ultimately free to conduct and report all analyses in their publicly available reports.

The Abt Team used the submission of the revised evaluation plan as an opportunity to ask the evaluators to establish, in advance, which of their findings the Abt Team would include in this report to draw conclusions about whether or not the intervention had positive impacts. We also asked them to commit to reporting the results of those prespecified analyses to us at the end of the evaluation. To be considered established in advance, the evaluator had to prespecify the analyses prior to collecting the outcome data that would be used in the analyses. Therefore, an analysis is considered prespecified if the plan for conducting it was established before the evaluator could have begun the impact analysis.

Virtually all i3 evaluators (66 of 67, or 99 percent) prespecified impact analyses prior to collecting student outcome data. One evaluator did not prespecify any impact analyses. Although this evaluator developed a plan for estimating the impacts of the intervention on student outcomes, it was not finalized prior to collecting student outcome data.

This process of prioritizing and prespecifying analyses meant that both the evaluators and the Abt Team had shared expectations about which findings would be used by the Abt Team to draw conclusions about the impacts of the intervention on student academic outcomes. The process set the expectation that evaluators would report the findings from the analyses to the Abt Team no matter the result, minimizing the extent to which evaluators suppressed null or unfavorable findings when they reported findings to us (sometimes referred to as the "file drawer" problem). In addition, the process limited the extent to which evaluators could redefine aspects of their analyses in order to find more favorable results to submit to the Abt Team for this report (sometimes referred to as "fishing").

The findings included in the main body of this report (Chapter 4) are limited to the prespecified analyses that focus on student academic outcomes, as the goal of i3 is to fund interventions that improve those

[20] The Abt Team's guidance was informed by the WWC review protocols, which establish outcome domains within which statistical adjustments should be made to minimize the risk of concluding, by chance, that the intervention had a positive or negative impact.

outcomes.[21] It is important to note that although all grantees are required to make the findings from the i3-funded evaluation publicly available, there was no requirement regarding which findings grantees should report publicly. Therefore, our summary of findings in this report might differ from other public reports of findings.

3.3.2 Identifying a Finding as the Result of a Prespecified Analysis

The evaluator defined four key elements of each prespecified analysis: (1) the treatment condition, (2) the control condition, (3) the educational level of the students, and (4) the outcome domain. Changing one of these four elements substantially redefines the impact being estimated; therefore, a finding is considered the result of a prespecified analysis if none of those four key elements had changed. The evaluators specified detailed analysis plans in advance that included, for example, equations for the regression models used to estimate impacts, an approach to including covariates in the impact model (i.e., whether or under what conditions they will be included) and plans for addressing missing data. The Abt Team used these plans to guide ongoing TA on the conduct of the evaluation but did not assess the extent to which the analyses conducted adhered to the details specified in the analysis plans.

Among the 66 evaluations (of 67) that prespecified their impact analyses, evaluators prespecified 222 analyses estimating the impacts on student academic outcomes and reported findings from 213 of them (96 percent). The vast majority of the 67 evaluators (59, or 88 percent) reported findings for all of the prespecified impact analyses; 7 evaluators (10 percent) reported findings for some but not all of their prespecified analyses. These 213 findings are the focus of the summary found in Chapter 4 and Appendix C, and information on each is listed in Appendix D.

Although the data collection focused on the findings resulting from prespecified analyses of student academic outcomes, we collected all findings that evaluators were willing to submit. Twenty-one evaluations (31 percent) prespecified 113 analyses on other outcomes and reported findings from 96 of them (85 percent; see Appendix E). Forty evaluations (60 percent) reported 210 findings from analyses estimating the impacts on student academic outcomes that were not prespecified (see Appendix F).

3.3.3 Using a Survey to Collect Information about the Conduct of and Findings from the Impact and Implementation Evaluations

The Abt Team asked the evaluators to complete an online survey to gather all of the information necessary to assess the strength of and summarize the findings from the impact evaluations. The survey also gathered the information necessary to summarize the findings from the implementation evaluations about whether the key components of the intervention were implemented with adequate fidelity.

To reduce burden on the evaluators, a WWC-certified member of the Abt Team facilitated data collection by using existing data sources to complete the survey. Specifically, if evaluators chose to use the standardized tools and templates provided by the Abt Team for tracking the progress of their evaluation (see Section 3.2.2), those resources were used to populate the survey. If the evaluators used their own reporting templates or had completed a draft report, the Abt Team used those resources to populate the survey. We then asked evaluators to verify the information populated by the data collectors, and to fill in any missing information.

---

[21] U.S. Department of Education, (2010), p. 12072. Findings from prespecified analyses of non-student outcomes and student non-academic outcomes are not included in Chapter 4, but are reported in Appendix E. Findings from analyses that were not prespecified are reported in Appendix F.

4. Did the i3-Funded Evaluations Meet the Goals of i3?

The short-term goal of the i3 program is to produce strong evaluations. The long-term goal is to build evidence on effective interventions at increasing scale. In this chapter, we summarize the strength of and findings from the 67 i3 evaluations completed by May 2017.

The short-term goal of the i3 program is met when:

- The implementation evaluation is **high quality** (Section 4.1.1);
- The impact evaluation is **independent**, meets **WWC Standards**,[22, 23] and **adequately represents** those served by each i3 intervention (Sections 4.1.2–4.1.4);

The long-term goals of the i3 program are met when:

- The implementation evaluation finds that the i3-funded intervention was implemented with adequate **fidelity** (Section 4.2.1); and
- The impact evaluation finds that the i3-funded intervention had at least one positive, statistically significant impact on a **student academic outcome** (Section 4.2.2).

Sections 4.1.5 and 4.2.3 provide cumulative assessments of the extent to which the i3 program met its goals.

4.1 Were the i3 Evaluations Strong?

The i3 program expected grantees to fund strong evaluations that included (1) a high-quality implementation evaluation and (2) an independent impact evaluation that (3) meets WWC Standards[24] and (4) adequately represents those served. In this section, we summarize whether the 67 completed i3 evaluations met these four expectations; Appendix C provides the assessments separately for each evaluation.

---

22 It is important to note that the evidence assessments we conducted only give an evaluation an unofficial WWC rating (see Section 4.1.3 for more details)

23 At the time of the report, the reviews in this report were conducted under WWC 3.0 Standards and Procedures for Group Designs, which covered randomized control trials and quasi-experimental designs; final standards for Regression Discontinuity (RD) Designs were not available under WWC 3.0 when these reviews were conducted. Therefore, results from an RD study that was conducted by Reading Recovery's evaluator, are not included in this report. Results from the RCT study that was also conducted by the evaluator of Reading Recovery are among the studies summarized in this report.

24 It is important to note that the evidence assessments we conducted only give an evaluation an unofficial WWC rating (see Section 4.1.3 for more details)

4.1.1 Were the i3 Implementation Evaluations High Quality?

ED expected the i3 evaluations to produce high-quality implementation data.[25] The Abt Team developed criteria to assess whether the i3 implementation evaluations met this expectation. An i3 implementation evaluation was considered high quality if it:

- Provided a logic model that identified the key components of the intervention, its mediators, and the student outcome domains it is intended to affect;
- Periodically measured the implementation fidelity of each of the key components on the sample that received the i3 intervention; and
- Compared the measurements to a threshold set by the developer and evaluator that indicates whether adequate fidelity was achieved, and reported the results to the Abt Team.

Sixty-five of the 67 i3 implementation evaluations were considered high quality. The remaining implementation evaluations were not considered high quality because implementation fidelity was not measured for the sample that was served by the i3 intervention. All 67 evaluations had complete logic models and periodically measured and reported whether adequate fidelity was achieved.

4.1.2 Were the i3 Impact Evaluations Independent?

All i3 grantees were required to conduct an independent evaluation.[26] The Abt Team developed criteria to assess the independence of the evaluations. We considered an evaluation independent if the evaluator reported at least one impact finding to the Abt Team for which they confirmed that neither the grantee nor the intervention developer collected the outcome data,[27] analyzed the outcome data, or reported the finding to the Abt Team.[28]

Almost all of the i3 impact evaluations reported findings that were independent. Sixty-five evaluations (97 percent) reported at least one finding that was considered independent (Exhibit 4.1). For 64 of these evaluations, all of the impact findings reported were considered independent. For one of these evaluations, some of the impact findings were not considered independent because the grantee collected data for a subset of the outcomes. For another evaluation, none of the findings were considered

---

[25] U.S. Department of Education (2010), p. 12086.

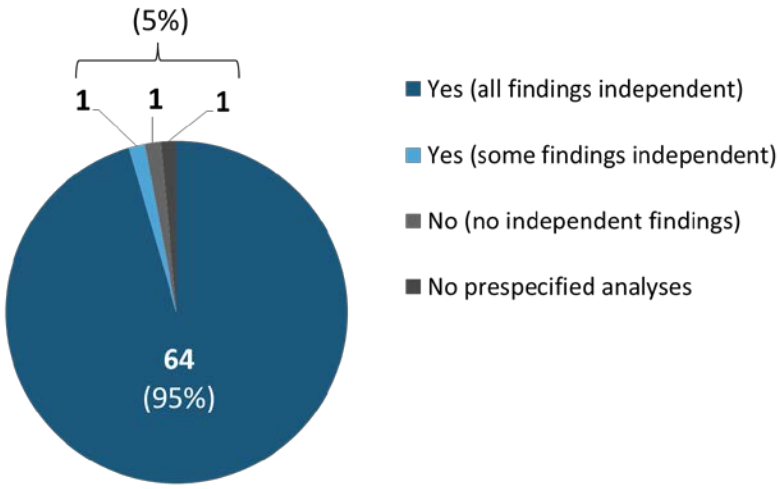
[26] U.S. Department of Education (2010), p. 12077.

[27] Achievement tests administered by states or districts are considered independent, even if the grantee is a state or local education agency.

[28] The criteria for independence could have included an assessment of whether the grantee or developer selected the comparison group in QEDs, since they may have been motivated to select comparison groups that are likely to have poorer outcomes than those receiving the intervention. However, this is covered by the WWC's baseline equivalence standards, which guard against selecting a comparison group using methods that differ substantially from those used to select the treatment group (see Section 4.1.3).

independent because the grantee collected data for all outcomes.[29] Finally, one of the evaluations did not prespecify analyses, and therefore the assessment of its independence is not included in this summary.[30]

Exhibit 4.1: Independence of the i3 Impact Evaluations



4.1.3 Were the i3 Impact Evaluations Strong?

ED expected the i3 grantees to sponsor strong impact evaluations that support conclusions about whether the intervention caused impacts on student academic outcomes. ED defines a strong impact evaluation as one that meets WWC Standards with or without Reservations.[31] To meet the standards without reservations, the WWC requires that assignment to the study groups (i.e., treatment and control) be determined by a random process—that is, the study must qualify as a randomized controlled trial (RCT)—and have tolerable levels of attrition bias, as determined by the combination of the overall rate of attrition and the difference in attrition rates between the two groups. To meet these standards with reservations, quasi-experimental designs (QEDs) and RCTs with high attrition must establish baseline equivalence of the two groups on a variable that is highly correlated with outcome. Finally, to meet WWC standards with or without reservations, the study must be free from serious confounds and use outcome

---

[29] The evaluator indicated that the grantee did not analyze the outcome data or report the findings to the Abt Team.

[30] However, from our review of the findings reported by this evaluation from analyses that were not prespecified, we conclude that those findings are independent.

[31] In the Notice Inviting Applications in 2010, ED expected evaluations funded by Development grants to provide evidence on the promise of the intervention (U.S. Department of Education (2010), p. 12075). And though evidence of promise can support further development and testing of an intervention, it does not support conclusions about whether the intervention caused any observed impacts. Starting with the 2015 Notice Inviting Applications, the selection criteria for Development grants awarded points for proposed evaluations with the potential to meet WWC Standards.

measures that have face validity and adequate reliability, that are not overaligned with the intervention, and that were collected in the same manner for the two groups.[32]

It is important to note that the Abt Team's assessment differs from an official WWC review in two important ways. First, the data available to the Abt Team differ from the data available to the WWC. The Abt Team collected the information necessary for the reviews directly from the evaluators and limited the summary of impact findings to the prespecified impact analyses (see Section 3.3). The WWC limits its reviews to publicly available reports. Second, only reviews conducted under a WWC contract can produce an official WWC rating for an evaluation. Thus, the evidence assessments we conducted for this report are not a substitute for a WWC review, and we can only give an evaluation an unofficial WWC rating. For these reasons, the official WWC rating of the i3 evaluations can differ from the ratings reported here. We include the unofficial ratings in this report because meeting the WWC Standards is a key goal of the i3 program and because it is important to place this summary of evaluation findings in the context of the strength of the evaluations.

The Abt Team used the WWC Standards that were current at the time of our assessment (Version 3.0 for this report).[33] WWC reviews are guided by a review protocol; the Abt Team developed a protocol to guide our application of the WWC Standards to the i3 evaluations.[34] The Abt Team assigned unofficial WWC ratings to the 213 findings reported by evaluators from prespecified analyses. The possible ratings were:

- Unofficially Meets WWC Standards without Reservations
- Unofficially Meets WWC Standards with Reservations
- Unofficially Does Not Meet WWC Standards
- Unofficially Ineligible for WWC Review

We summarized the ratings for each finding to assign a single rating for each evaluation. For 53 of the 67 evaluations, all of the findings for the evaluation received the same rating, which became the rating for the evaluation as a whole. For 13 of the 67 evaluations, the rating varied across findings; in these cases, we identified the most common ("modal") rating across the reported findings and assigned that rating to the evaluation.[35] One of the evaluations did not prespecify analyses, and therefore the ratings are not included in this summary.[36]

---

[32] For more details on the WWC Standards, see http://ies.ed.gov/ncee/wwc/Handbooks.

[33] See http://ies.ed.gov/ncee/wwc/Handbooks.

[34] See Appendix B for the review protocol.

[35] An alternative approach would be to identify the *highest* rating across the reported findings and assign that rating to the evaluation. Taking that approach changes the rating for three evaluations; one Scale-up grant would be reclassified from Unofficially Meets WWC Standards with Reservations to Unofficially Meets WWC Standards without Reservations, one Development grant would be reclassified from Unofficially Does Not Meet WWC Standards to Unofficially Meets WWC Standards with Reservations, and one Development grant would be reclassified from Unofficially Ineligible for WWC Review to Unofficially Does Not Meet WWC Standards.

[36] However, our review of the findings reported by this evaluation from analyses that were not prespecified concludes that the evaluation is Unofficially Ineligible for WWC Review.

Almost three-quarters of the i3 evaluations were considered strong because the evaluation received a rating of Unofficially Meets WWC Standards with or without Reservations (49 of 67, or 73 percent; see Exhibit 4.2). Of the 67 evaluations, 22 impact evaluations (33 percent) received a rating of Unofficially Meets WWC Standards without Reservations and 27 evaluations (40 percent) received a rating of Unofficially Meets WWC Standards with Reservations.

Exhibit 4.2: Modal Evidence Rating Received by Findings from i3 Impact Evaluations

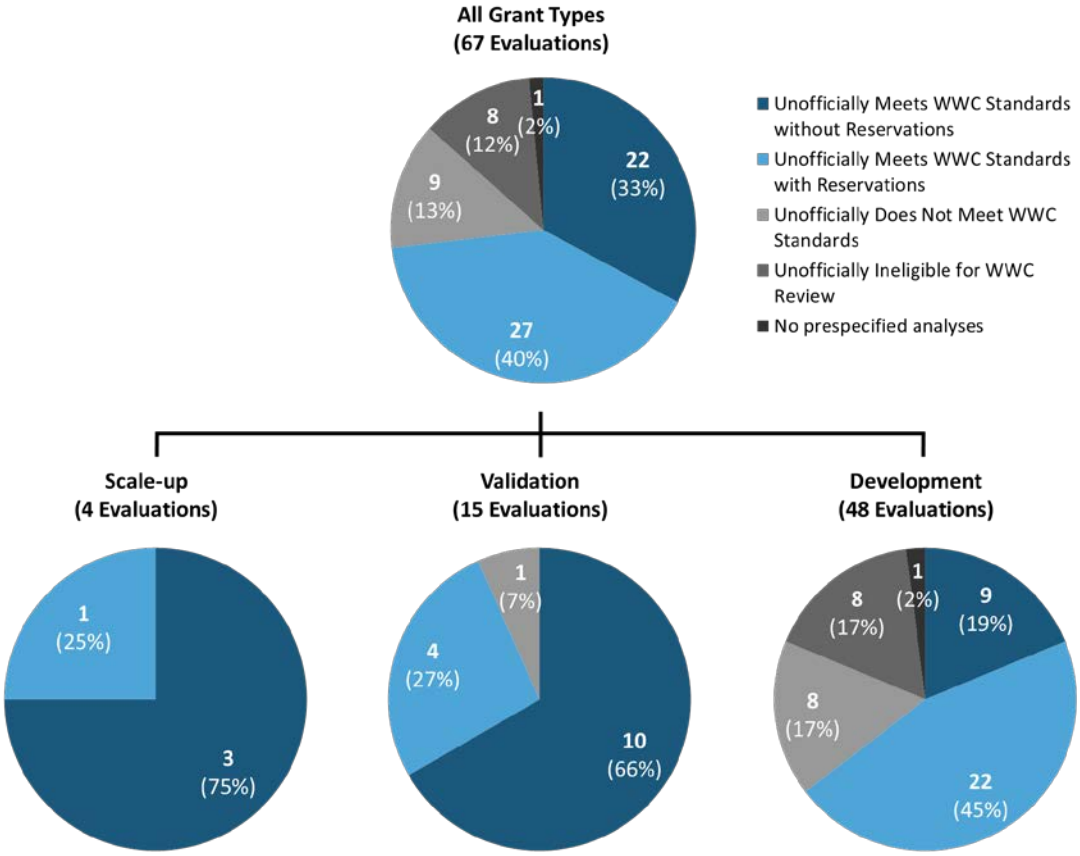


Exhibit 4.2 also shows that most grants of each type funded strong evaluations:

- All four evaluations funded by Scale-up grants received a rating of Unofficially Meets WWC Standards with or without Reservations;
- 14 of 15 evaluations funded by Validation grants (93 percent) received a rating of Unofficially Meets WWC Standards with or without Reservations; and
- 31 of 48 evaluations funded by Development grants (64 percent) received a rating of Unofficially Meets WWC Standards with or without Reservations.[37]

[37] The Abt Team reviewed the findings from 16 evaluations sponsored by Development grants that were rated Unofficially Does Not Meet WWC Standards or Unofficially Ineligible for WWC Review against criteria we developed to assess whether they "provide evidence on the promise of the intervention." All 16 evaluations received a rating of Provides Evidence on the Intervention's Promise.

The evidence ratings tend to be higher for evaluations of Scale-up and Validation grants than for evaluations of Development grants. For example, though three-quarters of Scale-up grants and two-thirds of Validation grants received the highest rating—Unofficially Meets WWC Standards without Reservations—approximately one-fifth of Development grants (19 percent) received this rating.

Eight impact evaluations, all funded by Development grants, were rated as Unofficially Ineligible for WWC Review because they conducted a pre-post comparison of outcomes for an intervention group without a comparison group that did not receive the intervention. Nine impact evaluations were rated as Unofficially Does Not Meet WWC Standards: three were QEDs that did not establish that the treatment and comparison groups were equivalent at baseline, one was an RCT with high attrition that did not establish that the treatment and control groups were equivalent at baseline, one was an RCT in which participants were non-randomly excluded after random assignment and did not establish that the treatment and control groups were equivalent at baseline, and four were QEDs in which the observed impact could not be attributed solely to the intervention (i.e., there was a confound).

4.1.4 Were the i3 Impact Evaluation Samples Adequately Representative of those Served by each i3 Intervention?

An important element of meeting the short-term goal of i3 is to learn about the average impact of each i3 intervention on the population of students served by each i3 project. The population of students ultimately served by each i3 intervention may differ from: (1) the population specified in the funded i3 grant application or (2) the population the intervention developer describes as the target population for the intervention when implemented beyond the i3 project. The Abt Team's assessment focuses on the population of students ultimately served by each i3 intervention, which differs for each i3 project.

Estimating the average impact of each i3 intervention on the population served by that intervention requires a sample that is representative of that population—and not just a sample that is convenient to include in the evaluation. Many i3 evaluations non-randomly excluded some of the schools, teachers, or students who received the i3 intervention because they were difficult or impossible to include in the evaluation (e.g., schools that received the intervention after the sample was selected, students that required one-on-one assessments) or did not take the steps necessary for inclusion (e.g., did not comply with the evaluation's assignment process or with data collection requirements). Therefore, the Abt Team assessed whether the exclusions of schools, teachers, or students that received the i3 intervention jeopardized the generalizability of the impact estimates to the population served by that intervention.

The impact evaluation sample is considered adequately representative of the population served by the i3 intervention if it includes (1) all or a random sample of the schools, teachers, and students that received the i3 intervention or (2) a non-random sample of these schools, teachers, or students that satisfied both of the criteria listed below:

1. **Excluded no more than 25 percent of the schools that received the i3 intervention.**[38, 39] The non-random exclusion of schools threatens the generalizability of the impact estimates because the

[38] An i3 evaluation conducted outside of a school setting must not non-randomly exclude more than 25 percent of the sites or localities in which the intervention was implemented.

impacts could differ between schools that were excluded and those that were included. Ideally, this criterion would document the characteristics of the excluded schools to assess whether the exclusions jeopardize the generalizability of the impact estimates. However, many evaluators did not report why schools were excluded from the sample.

2. **Excluded no more than 10 percent of teachers or students who received the i3 intervention based on variables that are known to be strongly associated with student academic outcomes.**[40] These variables include teacher experience, prior student achievement, prior academic performance (e.g., grade point average), race and ethnicity, income (e.g., eligibility for free or reduced-price lunch), English language proficiency (e.g., English learner status, language spoken at home), and special education status.

Over three-quarters of the i3 evaluations (53 of 67, or 79 percent) were considered to have estimated impacts on a sample adequately representative of those served by the i3 intervention (Exhibit 4.3). Of those:

- 29 evaluations included all schools that received the i3 intervention and did not exclude students based on factors that are strongly associated with student academic outcomes;
- 4 evaluations included all schools that received the i3 intervention and excluded less than 10 percent of the students based on factors that are strongly associated with student academic outcomes;
- 19 evaluations excluded less than 25 percent of schools that received the i3 intervention and did not exclude students based on factors that are strongly associated with student academic outcomes; and
- 1 evaluation excluded less than 25 percent of schools that received the i3 intervention and excluded less than 10 percent of the students based on factors that are strongly associated with student academic outcomes.

Thirteen impact evaluations estimated impacts on samples that were not considered to be adequately representative of the population served by the i3 intervention.[41] More specifically:

- 10 impact evaluations excluded more than 25 percent of the schools that received the intervention;
- 2 impact evaluations included all schools that received the i3 intervention and excluded more than 10 percent of the teachers who received and delivered the intervention based on teacher experience, a factor likely to be associated with student academic achievement; and

---

[39] Because any threshold is inherently arbitrary, we tested whether lowering the threshold from 25 percent to 10 percent would substantially increase the share of i3 evaluations that satisfied the sample representativeness criteria. For the results from this analysis, see Appendix G.

[40] If the impact evaluation excluded some of the schools that received the i3 intervention, the 10 percent threshold was applied to the teachers and students in the remaining schools.

[41] The assessment of one evaluation is not included in the summary because it did not prespecify any impact analyses. However, our review of the findings reported by this evaluation from analyses that were not prespecified concludes that the evaluation estimated impacts on a sample adequately representative of those served by i3.

- 1 impact evaluation included all schools that received the i3 intervention and excluded more than 10 percent of the students who received the intervention based on disability status, a factor known to be associated with student academic achievement.

The percentage of i3 impact evaluations based on adequately representative samples was smaller for Scale-up grants (25 percent) than for Validation grants (93 percent) or Development grants (79 percent; see Exhibit 4.3).

Only one of the four Scale-up evaluations included a sample that was adequately representative of the population that received the i3 intervention. The lack of representativeness raises concerns about drawing conclusions about the impacts on the population served by the i3 interventions from the remaining three Scale-up grant evaluations. The evaluation that was considered representative selected a random sample of schools that received the intervention and included 82 percent of those schools in the evaluation. The other three Scale-up evaluations excluded more than 25 percent of the schools that received the intervention. In particular:

- 1 excluded two-thirds of the schools that received the intervention because they were not sufficiently oversubscribed to support random assignment, opened after the third year of the grant, or lacked the data needed for the evaluation;
- 1 excluded all middle and high schools that received the intervention, thereby excluding the 64 percent of teachers who were in grades 6–12;[42] and
- 1 excluded more than 90 percent of schools that received the intervention at a discounted price under the i3 grant, and only included those schools that received the intervention at no cost.

All but one of the 15 evaluations of Validation grants (93 percent) were based on samples that were adequately representative of the population served by i3. The evaluation that was not based on an adequately representative sample excluded more than 25 percent of the schools that received the intervention.

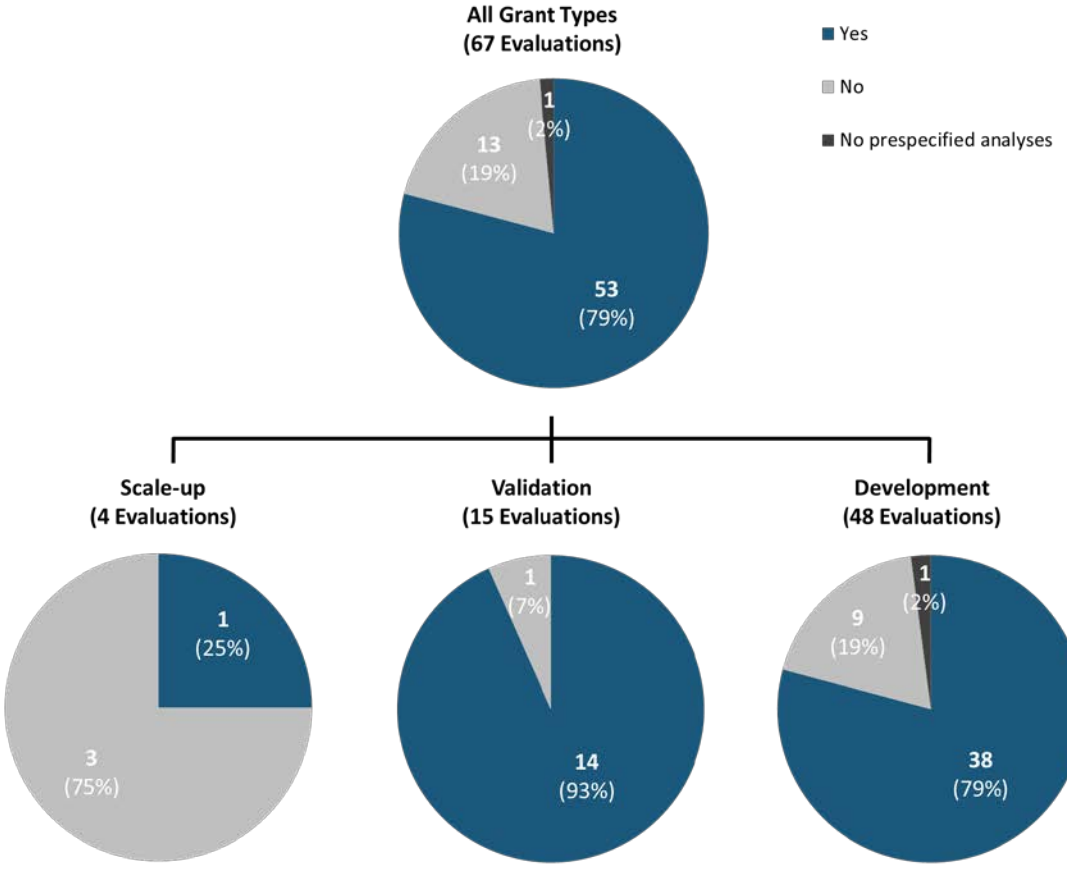
Thirty-eight of the 48 evaluations of Development grants (79 percent) were based on samples that were adequately representative of the population served by i3. Thirty-five included at least 75 percent of the schools and did not exclude students based on factors that are strongly associated with student academic outcomes (21 of these 34 evaluations included all schools). The other three included all schools and excluded less than 10 percent of the students based on factors that are strongly associated with student academic outcomes. Nine evaluations were not based on an adequately representative sample:[43] six excluded more than 25 percent of the schools, and three excluded more than 10 percent of students or teachers based on factors that are strongly associated with student academic outcomes (all three of these evaluations included all schools).

---

[42] This evaluation also excluded all teachers placed in participating schools after the first two years of the grant.

[43] See Appendix B for the review protocol; the assessment for one development grant was not included in this summary because it did not prespecify any impact analyses.

Exhibit 4.3: Representativeness of the i3 Impact Evaluations



4.1.5 Summary: Did the i3 Evaluations Meet the Short-term Goal of i3?

To meet the short-term goal of i3, the i3 evaluations were expected to produce high-quality implementation data, be independent, meet WWC standards,[44] and be adequately representative of those served by the intervention. Meeting the short-term goal of i3 requires meeting all four expectations; in this section, we summarize whether the i3 evaluations meet each expectation, given that they met the prior expectations. The assessments shown in Exhibit 4.4 build on each other, such that, for example, the number of grants shown as having an independent evaluation (second column) were also found to have a high-quality implementation evaluation (first column). See Appendix H for the assessments for each grant.

- **Almost all i3 evaluations (65 of 67, or 97 percent) conducted a high-quality implementation evaluation**. Meeting the expectations for implementation evaluations was not a common barrier to meeting the short-term goals of i3.

[44] It is important to note that the evidence assessments we conducted only give an evaluation an unofficial WWC rating (see Section 4.1.3 for more details).

- **Sixty-three i3 evaluations (94 percent) were *also* considered independent**; meeting this expectation was only a barrier to meeting the short-term goals of i3 for two grants.[45]
- **Almost three-quarters of the i3 evaluations (48 of 67, or 72 percent) *also*** included strong impact evaluations that produced findings that received a rating of Unofficially Meets WWC Standards with or without Reservations. Of the 63 grants that met the first two expectations, 15 did not meet the expectation that the findings unofficially meet WWC standards, making this expectation the most common barrier to meeting the short-term goals of i3.
- **Sixty percent of the i3 evaluations (40 of 67) *also* conducted the impact evaluation on a sample that adequately represented those served by the intervention, thereby meeting all expectations necessary to meet the short-term goal of i3.** This includes one of the four Scale-up grants, 13 of the 15 Validation grants, and 26 of the 48 Development grants.

Exhibit 4.4: Number and Percentage of Evaluations that Meet the Short-Term Goal of i3

Expectations for Meeting the Short-Term Goals of i3				Number of Evaluations Meeting All Expectations (N=67)
High-Quality Implementation Evaluation	Independent Impact Evaluation	Unofficially Meets WWC Standards	Impact Evaluation Adequately Representative of Population Served	
✓	–	–	–	65 (97 percent)
✓	✓	–	–	63 (94 percent)
✓	✓	✓	–	48 (72 percent)
✓	✓	✓	✓	40 (60 percent)

4.2 What did the i3 Evaluations Find?

Meeting the i3 program's long-term goal of building evidence on effective interventions requires identifying those interventions that can be implemented with adequate fidelity and are shown to improve student academic outcomes. In this section, we summarize the evaluations' findings; the findings from each of the i3 implementation and impact evaluations are presented in Appendix C.

4.2.1 Did the i3 Implementation Evaluations Find that the Interventions were Implemented with Adequate Fidelity?

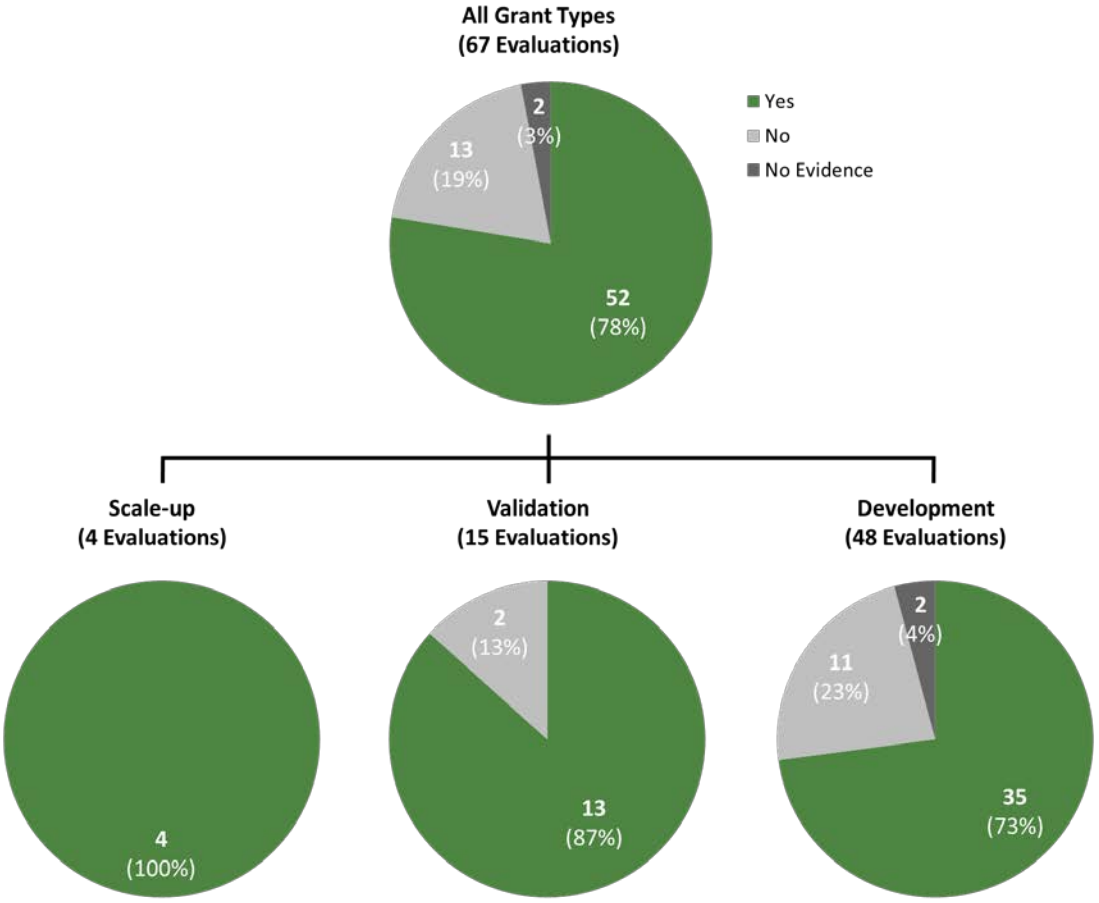
The i3 grantees were expected to implement the i3-funded interventions and serve students to improve student academic outcomes. As discussed earlier in the report, the evaluators, in consultation with the intervention developers, established thresholds for what they considered the adequate implementation of each key component named in the intervention's logic model. As a result, the fidelity measurement systems and thresholds for fidelity varied widely across grantees. (For example, whereas one fidelity measurement system might focus on the delivery of training to teachers by the program staff, another

[45] The assessments for one impact evaluation are not included in this summary because it did not prespecify any analyses.

might also include the extent of teacher participation in or satisfaction with the training as well.) Regardless of the measurement system, all i3 evaluators were expected to similarly document and report the extent to which each key component was implemented with fidelity, relative to those thresholds, in each year of measurement. The intervention was considered to have been implemented with adequate fidelity if the implementation of the majority of the intervention's key components met the specified fidelity threshold in at least 50 percent of the years of measurement.

Fifty-two of the 67 interventions (78 percent) were implemented with adequate fidelity (see Exhibit 4.5). Thirteen interventions were not implemented with adequate fidelity and two did not include a high-quality implementation evaluation. There was variation by grant type in whether or not the interventions were implemented with adequate fidelity. Seventeen of 19 Scale-up and Validation implementation evaluations (89 percent) found that the interventions were implemented with adequate fidelity compared to 35 of 48 Development implementation evaluations (73 percent). It is perhaps not surprising that Development grant interventions were less often implemented with adequate fidelity, given that the interventions were generally earlier in their development and testing at the time of the grant award.

Exhibit 4.5: Summary of Whether the i3 Implementation Evaluations Found that the Interventions Were Implemented with Adequate Fidelity



4.2.2 Did the i3 Impact Evaluations Find that the i3 Interventions Improved Student Academic Outcomes?

The short-term goal of the i3 program emphasizes the value of strong evidence, even when the results include null or negative findings. But ultimately, the long-term goal of i3 is to identify interventions that improve student academic outcomes. To conclude that the intervention improved student academic outcomes, an i3 impact evaluation must Unofficially Meet WWC Standards and find at least one positive and statistically significant impact and no negative and statistically significant impacts.[46] Exhibit 4.6 summarizes whether the i3 impact evaluations found any positive impacts on a student academic outcome.[47] Appendix C reports the summary findings separately for each i3 evaluation; for each finding reported by evaluators from prespecified analyses that Unofficially Meet WWC Standards, Appendix D includes the magnitude of the finding (i.e., the estimated effect size in standard deviation units) and whether it was statistically significant at the .05 level.

Twelve of the 67 impact evaluations (18 percent) found a statistically significant positive impact on at least one student academic outcome. The interventions that were awarded grants on the basis of stronger prior evidence of effectiveness were more likely to improve a student academic outcome: 50 percent and 40 percent of interventions supported by Scale-up and Validation grants, respectively, improved a student academic outcome, compared with 8 percent of interventions supported by Development grants.

There are, of course, different ways to summarize the findings to assess whether the i3 interventions improved student academic outcomes. We applied two alternative approaches, and present the findings in Appendix I. Both approaches yielded almost identical results to those shown in Exhibit 4.6. We also present an analysis of whether the results are sensitive to the inclusion of findings from analyses that were not prespecified (i.e., the findings reported in Appendix F).

Given the broad portfolio of educational interventions funded by i3, we expect the magnitude of the impacts reported by the i3 evaluations to vary by outcome domain. Exhibit 4.7 presents the average effect sizes reported in the four outcome domains in which i3 produced the most evidence: (1) 31 evaluations of student achievement in English Language Arts (ELA); (2) 23 evaluations of student achievement in

---

[46] Findings that are statistically significant at the .05 level were considered statistically significant. Hypothesis tests were adjusted for multiple comparisons when two or more findings were from analyses that shared the same outcome domain. This adjustment reduced the likelihood of finding one or more significant impact estimates purely due to chance. The Abt Team followed the process described in the review protocol in Appendix B.

[47] This approach to identifying interventions with positive effects is motivated by the WWC's approach to identifying effective interventions. The WWC concludes that an intervention has a statistically significant positive effect in a particular outcome domain if at least one finding is positive and statistically significant and none is negative and statistically significant (see the *What Works Clearinghouse, Procedures Handbook, Version 4.0*, Table IV.2). Furthermore, a tool provided by the WWC, *Find What Works,* can be used to identify studies with one or more statistically significant and positive findings across outcome domains (see https://ies.ed.gov/ncee/wwc/).

mathematics; (3) 7 evaluations of student achievement in science; and (4) 5 evaluations of educational progress or attainment.[48] Exhibit 4.7 shows that:

- **Seven of the 31 evaluations that examined impacts on ELA achievement found statistically significant effects, all of which were positive.** There were eight studies that reported negative effects, although none of these were statistically significant.
- **Five of the 23 evaluations that examined impacts on mathematics achievement found statistically significant effects, and they were mixed.** Overall, roughly half of the evaluations found positive effects and the other half found negative effects. A similar number of evaluations found statistically significant negative effects (two) as found statistically significant positive effects (three).
- **Only one of the seven i3 evaluations that examined impacts on science achievement found a statistically significant positive effect.** The remaining evaluations found null effects on science achievement.
- **Two of five (40 percent) of the i3 evaluations that estimated impacts on educational progress found significant effects, and both were positive.** The remaining three evaluations found null effects, one of which was negative.

[48] The amount of evidence in a domain was measured by the number of i3 evaluations that estimated impacts on at least one outcome in that domain. We included only analyses that Unofficially Meet WWC Standards with or without Reservations in Exhibit 4.7.

Exhibit 4.6: Summary of Whether the i3 Impact Evaluations Found Statistically Significant Positive Impacts on Student Academic Outcomes

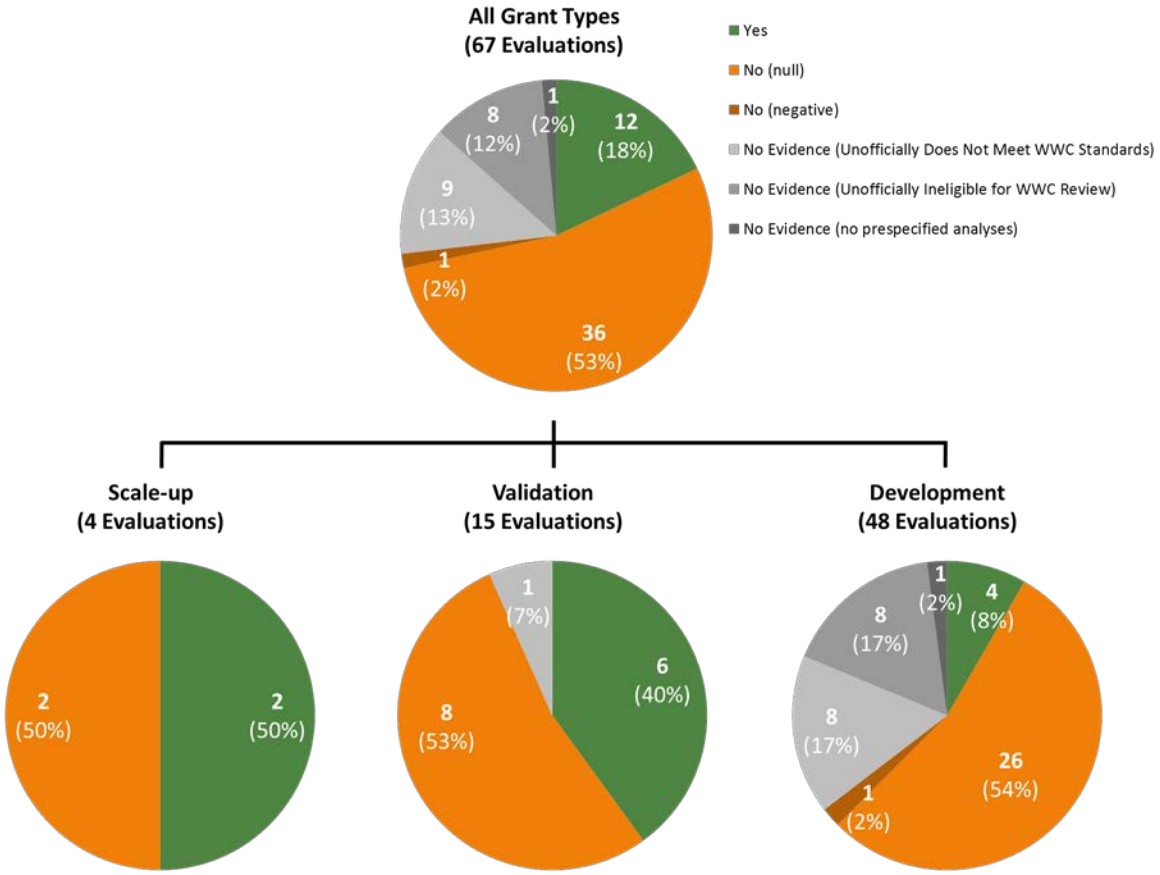


Exhibit 4.7: Effect Sizes for i3 Interventions, by Outcome Domain

4.2.3 Summary: Did the i3 Evaluations Meet the Long-term Goal of i3?

The long-term goal of i3 is to build evidence for effective interventions; this goal is met when an i3 evaluation finds that an intervention was implemented with adequate fidelity and had positive impacts on student academic outcomes. Exhibit 4.8 summarizes whether the i3 evaluations met both expectations.

- Almost 80 percent of the implementation evaluations—52 of 67—found that the intervention was implemented with adequate fidelity. This suggests that for most of the interventions, poor implementation was not an obstacle to meeting the long-term goal of i3.
- Thirteen percent of the evaluations—9 of 67—found that the intervention had been implemented with fidelity *and* had produced positive impacts, indicating that improving student academic outcomes was challenging. The nine evaluations that met the long-term goal of i3 include two Scale-up evaluations, four Validation evaluations, and three Development evaluations.

The percentage of i3 evaluations that met the long-term goal of i3 differed by grant type: 50 percent of the Scale-up and 27 percent of the Validation grants met the long-term goal of i3, compared with 6 percent of the Development grants. Perhaps this is not surprising given the high expectations for prior evidence that grant applicants needed to meet to receive a Scale-up or Validation grant. It is interesting to note that 3 of the 12 evaluations that found positive impacts also found that the intervention was not implemented with adequate fidelity. See Appendix H for the assessments for each grant.

Exhibit 4.8: Number and Percentage of Evaluations that Meet the Long-Term Goal of i3

Expectations for Meeting the Long-term Goal of I3		Number of Evaluations Meeting Expectations (N=67)
Intervention Implemented with Adequate Fidelity	Intervention Has Positive, Statistically Significant Impacts on Student Academic Outcomes	
✓	–	52 (78 Percent)
✓	✓	9 (13 Percent)

5. Did the Scale-up Grants Meet their Proposed Scale-up Goals?

A key goal of the i3 program was to bring to scale interventions with strong prior evidence of effectiveness at improving student achievement, closing achievement gaps, decreasing dropout rates, increasing high school graduation rates, or increasing college enrollment and completion rates. The i3 program selected established organizations that provided evidence of their capacity (e.g., qualified personnel, financial resources, management capacity) to scale up and meet ambitious targets for the expansion of effective practices. The targets were defined in terms of the number of students to be reached by the intervention. To achieve these targets, Scale-up grantees were expected to scale up their interventions to a state, regional, or national level. The i3 program provided substantial funding to Scale-up grantees in part due to the level of scale proposed in the application.

In this section, we describe the extent to which the Scale-up grantees:

- met their self-established "scale-up goal" (or expansion target);[49] and
- implemented their scale-up mechanism(s) with fidelity.

5.1 Capacity of the Scale-up Grantees Pre-i3 Award

The recipients of the Scale-up grants had to demonstrate organizational capacity to scale their interventions broadly. The four completed Scale-up grants were awarded to mature organizations that were all established at least 20 years ago. The grantees varied in size at the time of grant award. The scale at which each grantee was operating in the period just prior to receiving the i3 grant provides context for what it meant for that grantee organization to implement "at scale."

All of the Scale-up grantees were operating nationally just prior to the i3 award:

- In 2009–10, **KIPP** operated 82 schools serving 21,000 students in grades K–12, opening approximately 10 new schools each year;[50]
- In 2009–10, **Reading Recovery** served 73,000 grade 1 struggling readers in more than 5,000 schools in 40 states, and was declining in numbers of students and schools served in the years just prior to i3;[51]
- In 2010–11, **Success for All** was actively used in 881 schools in 48 states serving nearly 500,000 students in grades K–6;[52] and

---

[49] The assessments of whether the Scale-up grantees met their goals are based on information provided to ED in the grantees' Annual Performance Reports and their publicly available implementation reports. All other assessments in this report are based on the data provided to the Abt Team by evaluators, as described in Section 3.3.3.

[50] https://www.mathematica-mpr.com/our-publications-and-findings/publications/understanding-the-effect-of-kipp-as-it-scales-volume-ii-leadership-practices-at-kipp

[51] http://www.cpre.org/reading-recovery-evaluation-four-year-i3-scale

[52] http://www.mdrc.org/publication/scaling-success-all-model-school-reform

- In 2009–10, **Teach for America** placed 4,100 new first-year teachers in schools serving approximately 160,000 students in grades preK–12 in 27 states and the District of Columbia; it also was supporting 3,200 second-year teachers.[53]

5.2 Assessing whether Scale-up Grants met their Self-Established Scale-up Goals

The i3 program expected that the Scale-up grantees would have ambitious targets for expansion. All four Scale-up grantees established large scale-up goals, both relative to where the organizations were before the i3 grant award and in the absolute number of new implementation sites/students they intended to reach.

Three of the four grantees met their scale-up goals.

KIPP proposed to open 15 to 18 new schools each year, for a total of 75 new KIPP schools by the end of the grant period. It established a threshold of 65 schools serving at least 50,000 students as the criterion for meeting the goal (just over 85 percent of the total). These goals represented an increase of 80 percent over its pre-i3 number of schools and a doubling of the number of students served. KIPP met its scale-up goals, opening 66 new schools and reaching more than 59,000 new students.

Reading Recovery proposed that by the end of its five-year grant, it would have trained 3,675 new Reading Recovery teachers to serve more than 67,000 grade 1 struggling readers. The grantee established a threshold of 80 percent of these expansion goals as the criterion for having met its scale-up goals. Reading Recovery surpassed its goals for the number of Reading Recovery teachers and teacher leaders trained (3,747 new teachers) who worked with 62,000 struggling readers (92 percent of the goal).

Success for All (SFA) proposed to implement the model in 150 to 300 new schools each year for a total of at least 1,100 new schools that would serve 550,000 students. These goals represented a doubling of the number of SFA schools and students being served by SFA just before i3. Midway through the grant, with approval from ED, SFA revised its goals to 760 schools serving 380,000 students. SFA did not meet its original or revised scale-up goals, implementing SFA in 487 new schools and reaching 276,000 additional students.

Teach for America (TFA) proposed to place 13,500 new first-year teachers in the first two years of the grant, reaching more than 850,000 students.[54] These goals represented a 65 percent increase each year in the number of teachers placed by TFA just prior to i3. The grantee established a threshold of 75 percent of its expansion goals as the criterion for meeting its scale-up goals. TFA met its scale-up goal by placing 10,838 new teachers in the first two years of the grant.

---

[53] https://www.mathematica-mpr.com/our-publications-and-findings/publications/implementation-of-the-teach-for-america-investing-in-innovation-scaleup

[54] TFA also established a goal to train 46–47 percent of first-year teachers and 57–58 percent of second-year teachers to earn the rating of "highly effective" in the 2011–12 and 2012–13 school years. TFA met this goal (defined as reaching 75 percent of the goal) for 2011–12 but not for 2012–13.

5.3 Description of Implementation of the Scale-up Mechanisms

The Abt Team worked with each Scale-up grantee/evaluator team to develop a logic model that specified the grantee's scale-up goals and the key mechanisms through which it expected to reach them. Evaluators submitted these logic models to the Abt Team at the end of the design year, and we identified the mechanism that was central to how each grantee expected to expand to reach its scale-up goals. We assessed the extent to which the Scale-up grantees implemented the key mechanism for reaching their scale-up goals using the evaluators' public reports or their final Annual Performance Report to ED.

5.3.1 Scale-up Mechanisms

The three Scale-up grantees that met their scale-up goals also reported success in fully implementing the scale-up mechanisms for expansion. The one Scale-up grantee that did not meet its scale-up goals reported challenges in the implementation of their planned scale-up mechanisms.

KIPP. The key mechanism for scaling up was expanding the pool of trained school leaders to meet the demand for principals to lead new KIPP schools and to lead existing KIPP schools that needed new principals. To expand the supply of school leaders in the pipeline, KIPP both increased the capacity of its five KIPP leadership training programs and expanded the number of assistant principals in KIPP schools.

- KIPP provided training to 1,166 school leaders in its leadership training programs over the life of the grant, which exceeded its goal of 1,000 new KIPP-trained school leaders.
- KIPP expanded the number of assistant principal/dean positions in its schools during the grant period. KIPP created assistant principal positions in 171 of its schools, compared with the creation of 96 positions in the three years prior to the i3 grant.

Reading Recovery. The key mechanism for scaling up was to open new training centers and hire new teacher leaders at 15 of the Reading Recovery university training partners. The new teacher leaders would help recruit, train, and support the new Reading Recovery teachers in the scale-up.

- Although Reading Recovery reported that it did add 15 new training centers and hired 46 new teacher leaders across all of its training centers, it is not clear how many of the 46 teacher leaders were associated with the new training centers. Nor is it clear how many of the newly trained Reading Recovery teachers were trained through the new training centers.

Teach for America. TFA proposed three mechanisms for growing its corps of TFA teachers. First, TFA planned to scale up its recruitment efforts to create a larger recruitment pool. They proposed to achieve this increase by broadening their college recruitment base to include new types of colleges and include more graduate students and professionals at early stages of their careers. Second, TFA planned to implement new recruitment strategies to respond to an increase in attrition of TFA teachers between acceptance and matriculation. The new strategies included identifying and selecting committed candidates earlier in their college careers to increase commitment and offering transitional need-based grants to TFA teachers to help them with expenses prior to their entry into the TFA program. Third, TFA planned to increase its size, expanding into new regions of the country.

- TFA was successful at implementing its planned scale-up mechanisms. TFA increased its recruitment presence and efforts in "emerging market" schools (less selective overall but more diverse and with large new markets of prospective applicants) and in Historically Black Colleges and Universities and in schools that are members of the Hispanic Association of Colleges and Universities. The increased

recruitment targeting older students resulted in an increase in the proportion of new TFA teachers who were graduate students or post-college professionals.

- TFA reported that it was able to recruit more students earlier in their college careers than in the pre-i3 period.
- TFA expanded into 14 new regions by the second year of the i3 grant.

Success for All. The primary scale-up mechanism proposed by SFA was to significantly reduce the cost to districts and schools of adopting the Success for All program, since SFA had determined that the cost (around $150,000) posed a barrier to scaling up the model. SFA planned to achieve cost savings by moving to a district-focused model that uses local, district-embedded coaches rather than coaches from SFA. It further planned to make the district model more feasible by recruiting schools in concentrated geographic areas so that district coaches could serve more schools.

- SFA reported that it was not able to adequately implement the key mechanisms for reducing costs. SFA was able to implement the district-model in only a small number of districts; of the 29 local coaches who participated in the first year of training, only 12 continued into the second year and only 5 continued once i3 funding for the coaches ended.
- SFA did not achieve as much geographic concentration within districts or among neighboring districts as planned: 70 percent of school districts in the i3 scale-up had fewer than three SFA schools, compared with 77 percent in the period prior to the i3 grant.

5.3.2 Use of Scale-up Funds for Financial Incentives to Support Scale-up

Two of the Scale-up grantees planned to use i3 funds to offset the costs to scale up sites implementing their program. These grantees reported on the extent to which they offered these financial incentives to schools or districts. It is not clear whether and how these grantees will be able to continue after the i3 funding ends to provide financial support to sites interested in adopting the programs but for whom cost is a barrier.

Reading Recovery used its i3 grant to provide subsidies to schools interested in the program to cover one-time start-up costs of training new Reading Recovery teachers. Schools and districts that received the subsidies were expected to maintain a three-year commitment to implementing Reading Recovery.

In a fourth-year evaluation report on Reading Recovery, the grantee reported that a lower-than-expected percentage of i3-funded districts maintained the three-year time commitment to the program.[55] In order to receive i3 funds, districts, schools, and teachers were asked to make a commitment to implement Reading Recovery for three years. The report noted that many districts declined the opportunity because they were wary of assuming responsibility for funding Reading Recovery over the long term. And although willing districts signed a Memorandum of Agreement to formally document their commitment to the program, some districts did not sustain the commitment to Reading Recovery for three years. The report concluded that although the i3 grant made it possible to increase the presence of Reading Recovery in schools, it also attracted schools to the program that were not committed to funding it themselves.

---

[55] May, H., Sirinides, P., Gray, A., & Goldsworthy, H. (2016). *Reading Recovery: An evaluation of the four-year i3 scale-up. A research report.* Philadelphia, PA: University of Pennsylvania, Consortium for Policy Research in Education.

Success for All used its i3 grant to provide a partial subsidy to schools that covered start-up costs of implementing SFA, as an incentive to encourage and facilitate adoption of the intervention. Even with the subsidies, schools and districts were required to contribute two-thirds of the cost to implement the intervention (around $100,000).

In its final report,[56] SFA suggested that the higher-than-expected rate at which initially implementing sites dropped SFA (25 percent before the end of three years) might be because the financial incentive motivated some less-than-committed sites to adopt the program.

[56] Quint, J., Zhu, P., Balu, R., Rappaport, S., & DeLaurentis, M. (2015) *Scaling up the Success for All mode of school reform. Final report from the Investing in Innovation (i3) evaluation.* New York, NY: MDRC

6. Summary and Lessons Learned

In this section, we summarize the Abt Team's support to evaluators and the extent to which the strength of and findings from the i3 evaluations met the short- and long-term goals of i3. Then we identify key lessons for future grant programs considering using an evaluation requirement to produce evidence to identify effective interventions.

6.1 Summary

In addition to the $99 million that the grantees spent on evaluation, the i3 program invested approximately $30,000 per grant per year (or $10 million total) in technical assistance for the 67 evaluations summarized in this report. The support included a one-on-one relationship with a technical assistance provider who was a WWC-certified reviewer and experienced in conducting rigorous research in school settings, group technical assistance, and tools and templates for documenting the evaluation design and progress toward conducting a strong evaluation.

Almost three-quarters (73 percent) of the 67 i3 impact evaluations unofficially met the WWC evidence standards. All but two of the implementation evaluations produced high-quality data on fidelity, and almost all the i3 impact evaluations were independent. Seventy-nine percent were able to evaluate impacts on a sample that adequately represents those served by i3. **Overall, 60 percent of the i3 evaluations met the short-term goal of i3 to conduct strong evaluations of those served by i3.** This suggests that it is reasonable to expect grantees to conduct strong evaluations as part of a discretionary grant program, but that even with access to comprehensive evaluation technical assistance, meeting WWC standards *and* including a representative sample of those served is challenging.

Seventy-eight percent of the i3 implementation evaluations found that the interventions were implemented with adequate fidelity to the program models, and 18 percent of the impact evaluations found positive impacts. **Overall, nine evaluations (13 percent) met the long-term goal of i3 by finding evidence of both adequate fidelity and positive impacts on student academic outcomes.**

The findings from these evaluations can help ED in deciding which educational programs warrant additional funding and testing—and highlight the value of strong evidence, regardless of what the findings are. They also provide credible evidence to local decision makers considering whether to adopt particular interventions and useful information to policymakers contemplating the design of future discretionary grant programs.

6.2 Lessons Learned for Future Grant Programs

6.2.1 Building on Existing Evidence Standards made Expectations for and Assessments of the Impact Evaluations Clear

Future grant programs that consider using an evaluation requirement face the challenge of defining what grantees need to do to meet the requirement. The i3 program referenced the existing WWC evidence standards when setting its expectations for the impact evaluations. Because the i3 Notice Inviting Applications referred to existing standards, potential applicants could review the WWC standards to better understand what would be expected from their evaluations. The standards also served as the foundation for the Abt Team's evaluation technical assistance to help evaluators design and conduct evaluations likely to meet those standards.

The WWC standards played a central role in our ability to assess of whether the i3 grantees—and thus the i3 program as a whole—met the goal of contributing to the evidence base on which policymakers at all levels of government can make decisions about how to best educate students. Without established standards, the evaluation expectations for i3 grantees would have been less clear, and it would have been more difficult to assess whether that goal had been met.

However, it is important to note that the WWC regularly updates their standards and review processes. And while the WWC makes its current standards and review protocols available on their website, i3 evaluators often found it challenging to determine how these changes would affect the review of their evaluations. The support of the Abt Team was intended to help evaluators understand the implications of the revisions made to the WWC standards in March 2014 for their specific evaluation.

6.2.2 Developing Standards to Assess ED's Additional Expectations Provided More Information about the Strength of the Evaluations

Future grant programs may have additional expectations for evaluation that go beyond the internal validity of the impact evaluation, which is what is assessed by the WWC standards. ED had expectations of the i3 evaluations beyond whether they met WWC standards that the Abt Team operationalized into standards in several areas where existing standards were not available. The expectations included that (1) the evaluations would provide high-quality data on implementation, (2) the evaluations would be independent, and (3) that the impact evaluations would adequately represent those served by each i3 intervention.

We hypothesize that these i3-specific standards improved the i3 evaluations. For example, the Abt Team had numerous conversations with the evaluators about how to collaborate with the grantees without undermining independence. It is unclear whether evaluators and grantees would have attended to the issue of independence without that support. However, the limited time and resources that can be devoted to standards development for a single grant program, combined with the challenges of finalizing standards before evaluations begin, underscore the value of adopting or adapting established standards before funding a grant program with high expectations for evaluations.

Future grant programs might also consider stronger standards to align with higher expectations. For example, as discussed in Section 4.1.1, the standards developed to assess the i3 implementation evaluations focus on the logic model and measuring whether the key components of the intervention were implemented with adequate fidelity. Future standards could be extended to include whether the implementation evaluation measured key aspects of the experiences of those in the comparison group, to measure whether the experiences are in fact markedly different from those of the intervention group (sometimes called the "treatment-control contrast" or the "service contrast").

Finally, conducting evaluations that include a sample that adequately represents those served proved challenging; of the 9 out of 67 i3 evaluations that found both adequate implementation fidelity and positive impacts, six evaluations included a sample that was representative of those served. Future grant programs interested in whether each funded evaluation included a sample that adequately represents those served should make this expectation explicit to grantees and consider adopting standards like those developed for i3.

6.2.3 Future Grant Programs Could Find it Challenging to Assess the Strength of Evaluation Designs Submitted after Grant Award

Grant programs may want to assess the strength of the evaluations funded by the grantees as they are being designed and while they are being conducted. These assessments provide the opportunity for stakeholders, including the grant program office, the grantee, and the evaluator, to revise the evaluation plan and avoid common barriers to conducting strong evaluations. Ongoing assessments can also help the program office monitor its own progress towards funding strong evaluations across grants. Indeed, a key feature of the i3 program is that two of its short-term Government Performance and Results Act (GPRA) reporting measures require an assessment of the strength of the evaluation designs. The Abt Team, with oversight from IES, provided these assessments on an annual basis to inform ED's reporting requirements. Future grant programs interested in performance measures that include the strength of grantee evaluations will need to ensure they have the capacity to make these assessments.

Similarly, the i3 program included an assessment of the completed evaluations as part of the long-term GPRA performance measures. Future education programs could leverage the existing WWC review process for grants that are expected to produce evidence that meets WWC Evidence Standards. However, if a program had additional or varied evaluation expectations for grants, such as the expectation that they be independent or conduct a high-quality implementation evaluation, or that they would provide evidence of effectiveness aimed at informing earlier stages of program development (i.e., evidence that the intervention is promising), the program office would need to include a separate process for making those assessments.

Having stricter selection criteria around the strength of the evaluation designs included in the applications is one way future programs could begin to address this challenge. As mentioned in Section 3.2.1, the Abt Team's review of the i3 grantee applications found they did not contain enough information to determine whether the proposed evaluation designs would meet ED's expectations for strong evaluations. Over time, the i3 program has included stricter criteria regarding the quality of the project evaluation being proposed. For example, applicants are now required to include an intervention logic model in the application and to specifically discuss the magnitude of the impact on student academic outcomes that the proposed impact evaluation is designed to detect ("statistical power"). These requirements were intended to improve the evaluation designs described in submitted applications.

Further, programs should consider requiring grantees to submit revised evaluation designs early in the grant period and provide clear expectations for what should be included in those plans. The Abt Team asked evaluators in the FY 2010 cohort to submit revised evaluation plans and supported the evaluators in strengthening the plans. Starting with the FY 2011 cohort, the i3 program required the submission of revised evaluation plans and included detailed expectations for those plans in the cooperative agreements signed by each grantee in those cohorts. Requiring that grantees submit revised plans alone may not ensure the strength of the plans; the Abt Team worked closely with i3 evaluators to try to improve their plans and provided ED with detailed reviews about the extent to which the plans met ED's expectations.

6.2.4 Expecting Evaluators to Prespecify Analyses Ensured that Findings from those Analyses would be Reported to the Abt Team, Regardless of the Direction and Magnitude

Grant programs that include an evaluation requirement will want to ensure that the findings from the funded evaluations are made public and contribute to the evidence base. However, requiring that grantees produce a public report does not ensure that all findings will be reported. The i3 program emphasized to grantees and evaluators the importance of reporting all findings from the evaluations, whether they are

positive, null, or negative, but grantees may have incentives to include some findings and not others in their public reports. Further, grantees may also have incentives to conduct additional analyses if those that are originally planned are not favorable, and to highlight results from those additional analyses in their public reports when they are positive ("fishing").

The Abt Team asked the i3 evaluators to prespecify which of their evaluation findings we would include in this report summarizing the impact of the interventions on student academic outcomes. When needed, the Abt Team provided technical assistance to help evaluators narrow their planned analyses to those that would produce the most credible evidence of impacts on students. This report includes 213 impact evaluation findings submitted by evaluators, representing 96 percent of prespecified analyses; that the findings vary in their direction and magnitude could indicate that they were not selectively chosen.

It is important to note that the protections described above are unique to this report's summary of the i3 evaluation findings. The findings in this report may differ from those included in reports made publicly available by grantees and evaluators (which will be reviewed by the WWC as they are made available) because we focused this report on findings from analyses of student academic outcomes that were prespecified and reported directly to the Abt Team. The prespecification and data collection functions performed by the Abt Team were designed to provide some protection against losing potentially important findings from the impact evaluations and against including findings from analyses that were not planned in advance.

Future programs could take advantage of existing education evaluation registries, such as the one currently under construction by the Society for Research on Educational Effectiveness (SREE),[57, 58] and require that grantees register their evaluation plans. To fully recognize the benefits of the registration process, however, staff monitoring the grants would need to commit to holding grantees accountable for conducting and reporting on the evaluations *as registered*. Policymakers interested in including this type of registration and review process in future programs should not assume that evaluators will conduct and report findings as they were detailed in the evaluation plans, and will report *all* findings, regardless of whether they are favorable to the intervention, without ongoing support and oversight.

6.2.5 Testing the Interventions at Scale was a Challenge for Scale-up Grants

The purpose of the Scale-up grants was to bring interventions supported by strong prior evidence of effectiveness to scale and learn about their effectiveness at scale. As discussed in Section 5.3, three of the four Scale-up grantees met their goals for scaling up the intervention, and all four implemented their intervention on a much larger scale than it had previously been implemented. However, only one of the four impact evaluations included a sample that was representative of those served by the i3 intervention (see Section 4.1.4). This means that three of the four impact evaluations supported by i3 Scale-up grants were not able to test the effectiveness of the i3-funded intervention at scale and limits the conclusions that can be drawn from these studies.

---

[57] Maynard, R., & Spybrook, J. (March, 2016). *A registry of effectiveness studies in education*. Paper presented at the annual meeting for the Society for Research on Educational Effectiveness, Washington, DC.

[58] The WWC previously maintained a registry but transferred the database to the SREE registry, which is scheduled to launch in late 2017.

An important constraint in testing the effectiveness of the interventions at scale was that all the implementation and evaluation activities needed to be completed within the five-year grant period.[59] All four grantees planned to use the full five years of the grant to meet their ambitious scale-up goals. Further, two of the four grants (Teach For America and Success For All) funded multi-year interventions that were hypothesized to have the greatest impacts more than one year after a teacher or school begins to implement the intervention. To complete the evaluation within the five year grant period, the evaluators typically selected a sample of teachers or schools that implemented the intervention early in the grant period, before the intervention had been taken to scale. These evaluations provide evidence about the impacts of these interventions early in the scaling process, and not their effectiveness at scale. Providing evidence of the effectiveness at scale would have required a longer grant period with clear expectations to select the evaluation sample to adequately represent the schools, teachers, or students that were part of the effort to scale the intervention.

6.2.6 Relying on Administrative Data was a Challenge

Grant programs aiming to fund interventions to improve outcomes that are routinely measured (e.g., student achievement) can benefit from using that administrative data for evaluation purposes. Using outcome data that is collected for other purposes is both cost-effective and reduces burden on participants. However, states do not routinely test all subjects in all grades, and evaluators cannot estimate impacts of interventions in untested grades. This was a particular challenge for interventions aimed at improving science achievement because states typically only administer standardized science tests once in middle school and once in high school. Further, heavy reliance on administrative data to measure student outcomes is risky when states or districts change their plans for testing or collecting data. For example, when California did not field the California Standards Tests in 2014, a number of i3 evaluators had to either select a new outcome measure or measure outcomes in another year (e.g., a year earlier). Future grant programs should balance the benefits of using administrative data with the risks evaluations face if the availability of those data changes.

6.2.7 Resources for Future Programs that Require Evaluation

The Abt Team developed a number of resources to support the i3 evaluators that could be leveraged by future programs hoping to provide similar support. ED has established the following link to make resources available: http://ies.ed.gov/ncee/projects/evaluationta.asp. Examples of available resources developed as part of this project include these:

- The Evaluation Plan Template, which identifies the key components of an impact evaluation plan and provides guidance about the information typically included in each section of a plan for evaluating the effectiveness of an intervention.
- The Contrast Tool, which is an Excel spreadsheet designed to support and supplement the development on an evaluation plan. The spreadsheet documents the specific comparisons the evaluator will make between treatment and control group outcomes to test program effectiveness.
- The Fidelity Measure Tool, which documents how adequate fidelity to the intervention model will be assessed.

[59] These grants were awarded with FY 2010 funds, which expired at the end of September 2015, and therefore the i3 program could not grant any extensions to grants in the FY 2010 cohort.

Appendix A: Key Components and Measures of Implementation Fidelity for 67 i3 Interventions

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Scale-up				
KIPP				
KIPP schools seek to engage students and parents actively in the educational process, expand the time and effort students devote to their studies, reinforce students' social competencies and positive behaviors, and dramatically improve their academic achievement. KIPP schools are grounded in five core principles: high expectations, choice and commitment, more instructional time, power to lead, and focus on results.	KIPP Schools[60]	Target Leadership Structures and Supports	2	YES
Reading Recovery				
Reading Recovery is an intensive reading intervention for struggling elementary readers. The Reading Recovery program targets the lowest-achieving 15 to 20 percent of first graders in a school. These students receive 12- to 20-week cycles of daily, 30-minute, one-on-one Reading Recovery sessions. The intervention includes daily monitoring of students to ensure that instruction is responsive to changes in student achievement and needs. The intervention also includes a training and	Staff Background and Selection	Select/Evaluate Staff	4	YES
	Reading Recovery Teacher Training and Ongoing PD	Provide PD for Teachers	4	YES
	Teacher Leader and Site Capacity	Provide PD for Teachers	4	YES

[60] Unlike other grantees, KIPP defined fidelity of implementation based on a single key component that was a total of 7 indicators: KIPP-selected principal, KIPP-trained principal, KIPP-approved school design plan, school has licensing agreement to use the KIPP name, school is connected to KIPP network (e.g., through KIPP communities of practice or national retreats), school participates in school reviews within first two years, school collects and shares performance data with KIPP.

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
professional development program for teachers and teacher leaders in order to implement Reading Recovery instruction.	One-to-One Reading Recovery Lessons	Develop/Institute New Curriculum & Materials	4	YES
Success for All (SFA)				
Success for All (SFA) is a whole-school turnaround program that focuses primarily on ensuring that every child succeeds in learning to read throughout the elementary grades. The program includes professional development for all school staff, a K-6 reading program, assessments, and tutoring. SFA staff also address non-academic issues, support K-6 teachers, and work to develop leadership within schools.	Essential training	Provide PD for Teachers and Other Staff	2	NO
	Program materials	Develop/Institute New Curriculum & Materials	2	YES
	School inputs and implementation	Develop/Institute New Curriculum & Materials	2	YES
Teach for America (TFA)				
Teach For America (TFA) is a nonprofit organization that seeks to reduce educational inequities through the provision of high quality teachers to low-income schools. The program recruits college graduates and professionals with strong academic backgrounds and leadership capabilities to commit to teach for two years in low-income schools. These participants, called corps members, typically have no formal training in education but receive an intensive five-week training from TFA before beginning their first teaching job, as well as ongoing support throughout their two-year commitment.	Selection	Select/Evaluate Staff	2	YES
	Placement	Select/Evaluate Staff	2	YES
	Summer Institute (pre-service training)	Provide PD for Teachers	2	YES
	Ongoing Support	Target Leadership Structures and Supports	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Validation				
Advanced ASSET Professional Development				
Advanced ASSET Professional Development aims to improve student achievement and self-efficacy in math and science by improving teacher effectiveness. The program uses the National Science Resources Center's model for STEM education programs and uses a "teachers teaching teachers" model and two research-based approaches to professional development.	ASSET Teacher Professional Development-STRUCTURE	Provide PD for Teachers	4	YES
	ASSET Teacher Professional Development-PROCESS	Provide PD for Teachers and Other Staff	3	YES
	Development of a "Professional Learning Community" School Culture-STRUCTURE	Support Staff Collaboration	4	YES
	Development of a "Professional Learning Community" School Culture-PROCESS	Support Staff Collaboration	3	YES
The Baby Family and Child Education (FACE) Program				
The Baby Family and Child Education (FACE) program aims to improve children's school readiness through greater parent involvement prenatal to age 3. Parents as Teachers (PAT), which runs this program, provides for early detection of developmental delays and health issues, and parent education to help parents understand their role as their child's first teacher. The Baby FACE program is a targeted home-based intervention for American Indian families living on tribal reservations in primarily rural areas. The intervention provides home visits, routine health and developmental screenings for children, parent group meetings, and resource referrals as needed.	Training & Professional Development	Provide PD for Other School or Non-School Staff	1	YES
	Support & Technical Assistance	Provide Coaching for Staff	2	YES
	Service Delivery	Develop/Institute New Curriculum & Materials	2	NO

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Children's Literacy Initiative (CLI) Program				
The Children's Literacy Initiative (CLI) program provides prekindergarten through third-grade teachers with training and coaching in effective practices for early literacy instruction. CLI's coaching and professional development aims to raise student reading achievement by helping teachers establish and maintain literacy-rich classroom environments and develop shared standards of high-quality instruction. The CLI program includes four key features: 1) Providing teachers with literacy resources, including book collections, to create a literacy-rich classroom environment; 2) Conducting professional development institutes and seminars to train teachers in strategies and techniques for literacy instruction, followed by classroom-embedded coaching to help teachers apply these strategies in the classroom; 3) Identifying one model classroom teacher per grade who receives intensive coaching and support and in whose classroom additional embedded coaching can be provided; and 4) Educating school leaders on how to leverage CLI training to sustain high-quality literacy instruction in the school.	Resources and professional development	Provide PD for Teachers	3	YES
	Coaching	Provide Coaching for Teachers	3	YES
	Model classrooms for professional development	Provide Coaching for Teachers	3	YES
	Principal and school-based coach professional development and leadership meetings	Provide Coaching for Administrators	3	NO
Collaborative Strategic Reading (CSR)				
Collaborative Strategic Reading (CSR) Colorado is a research-based intervention for cross-content-area teachers to improve discipline-specific reading comprehension. CSR Colorado is built on a foundation of reciprocal teaching and combines learning strategies, parent outreach, and teacher	Teacher knowledge of CSR	Provide PD for Teachers	2	YES
	Teacher implementation of CSR	Develop/Institute New Curriculum & Materials	2	YES
	Leadership knowledge of CSR	Provide PD for Other School or Non-School Staff	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
professional development and coaching to improve reading comprehension for students. CSR Colorado is intended to help students develop metacognitive awareness and learn specific approaches associated with enhanced reading comprehension using four strategies: (1) preview—brainstorming and predicting or setting purposes for reading; (2) click and clunk—monitoring understanding and taking steps to figure out unknown words or confusing ideas; (3) get the gist—determining main ideas and gist statements; and (4) wrap-up—generating questions and reviewing key ideas or summarizing.	Leadership implementation of CSR	Target Leadership Structures and Supports	2	YES
	District implementation of CSR parent engagement activities	Involve Parents/Community Members	2	YES
College Readiness Program (CRP)				
The College Readiness Program (CRP) provides supports to schools in the form of program management (e.g., regional director oversight, student/school-level data, multi-way performance feedback), teacher supports (e.g., lead teachers, team meetings, content leads, summer institutes, trainings), student supports (e.g., paid exam fees, equipment and supplies, open enrollment), and awards (administrator, teacher, and student incentives for meeting AP exam score goals). The goal of CRP is to create an environment where students can participate in AP courses and receive instruction that will lead to increases in the share of students who pass AP tests, therefore increasing college readiness, the central goal of the CRP program.	Awards	Develop/Institute New Curriculum & Materials	2	YES
	Program Management	Plan for and Support Assessment & Data Use	2	YES
	Student Supports	Develop/Institute New Curriculum & Materials	2	NO
	Teacher Supports	Provide PD for Teachers	2	NO

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
College-Ready Writers Program (CRWP)				
The College-Ready Writers Program (CRWP) is designed to improve the argument writing of students in grades 7 through 10 by introducing teachers to new instructional practices. Local university-based site affiliates provide professional development to English Language Arts teachers. The core program features include intensive professional development to support classroom implementation, use of curricular resources, and formative assessment to improve instruction. By changing teacher attitudes and instructional practices, CRWP aims to improve students' performance on source-based argument writing tasks and to increase students' use of writing processes to compose and refine text and to support learning in other areas.	Content of Teacher Professional Development	Provide PD for Teachers	2	YES
	Duration and Breadth of Teacher Professional Development	Provide PD for Teachers	2	YES
	Duration and Breadth of Teacher Professional Development over 2 Years	Provide PD for Teachers	1	NO
	Professional Development Strategies	Provide PD for Teachers	2	YES
Diplomas Now				
Diplomas Now seeks to turn around high-poverty high-minority middle and high schools using a whole-school reform model that aims to dramatically increase high school graduation rates. This project unites three organizations—Talent Development, City Year, and Communities in School—and includes a tiered intervention model that uses Early Warning System data to provide interventions targeted to students' needs; a staffing model that supports interdisciplinary teacher teams with common planning time; teacher professional	Strong Learning Environments (SLE)	Support Staff Collaboration	2	YES
	Professional Development and Peer Coaching (PDPC)	Provide PD for Teachers	2	NO
	Curriculum for College Readiness (CCR)	Support College/Career Readiness	2	NO
	Tiered Intervention Model (TI)	Plan for and Support Assessment & Data Use	2	YES
	Student Supports (SS)	Select/Evaluate Staff	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
development and peer coaching; a curriculum to prepare students for college readiness; additional student supports including mentoring and afterschool programming; student case management; on-site supports to facilitate implementation of the whole school organization reforms; and efforts to engage in family and community partnerships that support student success.	Student Case Management (SCM)	Provide Services Targeting Individualized Learning	2	YES
	Integrated On-Site Support (OSS)	Select/Evaluate Staff	2	YES
	Family and Community Involvement (FCI)	Involve Parents/Community Members	2	YES
	Program Staff Training and Development (PTD)	Provide PD for Other School or Non-School Staff	2	YES
Enhancing Missouri's Instructional Networked Teaching Strategies (eMINTS)				
enhancing Missouri's Instructional Networked Teaching Strategies (eMINTS) is based on four underlying research-based components: inquiry-based learning, high-quality lesson design, a community of learners, and technology integration. The program provides teachers with two years of professional development and supports that include monthly classroom visits and a suite of technology tools. In the third year, teachers have access to additional, newly created professional development that aligns with Intel's suite of web-based teaching tools to expand teachers' use of inquiry-based learning. eMINTS is hypothesized to improve student engagement and achievement in mathematics, communication arts, and 21st century skills.	Technology infrastructure	Develop/Institute New Curriculum & Materials	2	YES
	Technology use	Develop/Institute New Curriculum & Materials	2	YES
	Teacher professional development	Provide PD for Teachers	3	NO
	Administrative support	Provide PD for Other School or Non-School Staff	2	YES
	Ongoing technology support	Provide PD for Teachers	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
GO College (An Enhanced Version of Talent Search)				
GO College, an enhanced version of Talent Search, is implemented in cities and schools with previous participation in the GE Foundation Pilot. The primary goal is to develop and provide more intensive and extensive college access services for schools that serve a high proportion of low-income, first generation college-going, and minority students. GO College provides additional services through Talent Search counselors already in high schools with the objective of enhancing the college culture, increasing access to data, improving attitudes about college, and bolstering students' academic engagement and achievement to ultimately increase high school graduation and college enrollment.	Conduct focus groups and survey data needs	Plan for and Support Assessment & Data Use	2	NO
	Data sharing agreements with districts implemented	Plan for and Support Assessment & Data Use	3	NO
	Data tool developments & updates	Plan for and Support Assessment & Data Use	2	NO
	Learning Community (LC) coach training and professional development linked to data tool	Provide PD for Other School or Non-School Staff	3	YES
	Data tool usage plan for each stakeholder group	Plan for and Support Assessment & Data Use	2	YES
	Foster and monitor data tool use for all stakeholders	Plan for and Support Assessment & Data Use	3	YES
	Asset/strength based support work with students and parents	Involve Parents/Community Members	1	NO
	Embedded college coaches in schools	Select/Evaluate Staff	3	YES
	Data based goal setting with school principals and other stakeholders	Plan for and Support Assessment & Data Use	2	YES
	Provide college/university visits, and college/career awareness services	Support College/Career Readiness	3	YES
	Fostering whole school (WS) events: Career, college-going and enrichment	Support College/Career Readiness	1	YES
	Fostering math and science events	Develop/Institute New Curriculum & Materials	3	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
	Institute Summer Bridge for rising 9th graders	Develop/Institute New Curriculum & Materials	3	YES
	Academic support for taking rigorous college prep courses and success	Support College/Career Readiness	1	YES
	Tutoring and mentoring services	Provide Services Targeting Individualized Learning	1	YES
	Data use skill building and project-based data use for students	Develop/Institute New Curriculum & Materials	3	YES
	Hold regular and special stakeholder meetings	Support Staff Collaboration	3	YES
	Hold special and periodic joint meetings to create and maintain collaboration	Support Staff Collaboration	3	YES
	Distribute college-going outreach materials	Support College/Career Readiness	1	YES
	Academic year meetings for 9th graders in LCs	Develop/Institute New Curriculum & Materials	1	YES
	Academic year meetings for 10th graders in LCs	Develop/Institute New Curriculum & Materials	1	YES
	Institute supplemental summer service for rising 10th, 11th and 12th graders	Develop/Institute New Curriculum & Materials	1	YES
	Engage community to identify with brand	Involve Parents/Community Members	3	YES
	FAFSA and financial aid knowledge and support	Support College/Career Readiness	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Leadership Assistance for Science Education Reform (LASER)				
Leadership Assistance for Science Education Reform (LASER) is a professional development model to transform K-12 science education by helping state-, district-, and school-based teams create the infrastructure to support high-quality, inquiry-based science instruction. The intervention relies on a research-based curriculum, differentiated professional development, and professional learning networks.	Science Technology Concepts (STC) Curriculum Instruction	Develop/Institute New Curriculum & Materials	3	YES
	Research-Based Inquiry Instruction	Develop/Institute New Curriculum & Materials	3	YES
	Differentiated Professional Development	Provide PD for Teachers	3	NO
Northeast Tennessee College and Career Ready Consortium				
The Northeast Tennessee College and Career Ready Consortium aims to improve high school students' college and career readiness by increasing their access to, participation in, and completion of advanced courses. It seeks to achieve these goals by scaling up local promising practices to offer a wide array of advanced high school and college credit–bearing courses, particularly in mathematics, science, foreign languages, and career and technical education. The Consortium uses distance and online learning, as well as college partnerships, to increase offerings of Advanced Placement (AP), dual enrollment, and other upper-level high school courses. The i3 grant also was used to create a regional coordinating body to analyze course supply and demand in the region and determine course	Consortium management and communication	Provide PD for Teachers and Other Staff	2	YES
	College and career ready counselors team promotes a college-going culture	Develop/Institute New Curriculum & Materials	2	NO
	Instructional specialists increase quality of instruction and student participation in i3 courses	Target Leadership Structures and Supports	2	YES
	Technology coordinator/director of learning resources ensure access to distance and online courses	Target Leadership Structures and Supports	2	NO
	College partners reduce barriers to dual enrollment courses	Develop/Institute New Curriculum & Materials	2	NO

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
needs; offer professional development for teachers in an effort to improve the rigor of courses; and provide college and career counseling to encourage college access and help students with the college application process.	LEA partners share mission and resources, adopt practices in their schools, and encourage cultural change	Support Staff Collaboration	2	YES
	Resources and services provide infrastructure to expand and sustain program capacity	Develop/Institute New Curriculum & Materials	2	YES
	i3 leadership provides management and communication	Target Leadership Structures and Supports	1	YES
	College and career ready counselors team promotes a college-going culture	Develop/Institute New Curriculum & Materials	1	NO
	Learning resources team increases quality of instruction	Develop/Institute New Curriculum & Materials	1	NO
	Learning resources team increases access to courses through distance and online technology	Institute Structural Change	1	NO
	Learning resources team expands opportunities for college level courses	Institute Structural Change	1	NO
	Resources and services provide infrastructure to expand and sustain program capacity	Institute Structural Change	1	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Reading Apprenticeship (RA)				
Reading Apprenticeship (RA) is an instructional framework that helps teachers support discipline-specific literacy and learning. This intervention provided teachers with 65 hours of inquiry-based RA professional development over the course of 12 months. The professional development was designed to transform teachers' understanding of their role in adolescent literacy development and to build enduring capacity for literacy instruction in the academic disciplines. Changes in teacher attitudes and instructional approaches are hypothesized to result in changes in student attitudes and motivation while, at the same time, building skills and knowledge for subject-specific literacy tasks, strengthening students' view of themselves as readers and learners, and yielding substantial gains in student achievement. Teachers were also supported by teacher leaders who held monthly school-based meetings throughout implementation and state coordinators who provided support and resources.	Teacher Attendance at Professional Development	Provide PD for Teachers	1	NO
	Teacher Attendance at Monthly Meetings	Provide PD for Teachers	3	YES
	Teacher Leader Recruitment	Select/Evaluate Staff	3	YES
	Delivery of PD Content Adheres to Standards	Provide PD for Teachers and Other Staff	1	YES
Reading Enhances Achievement During the Summer (READS)				
The Reading Enhances Achievement During the Summer (READS) program aims to improve reading comprehension in elementary school students by fostering children's engagement with books over the summer. READS addresses the challenge of access to books at home by distributing books over the	Book Activities	Develop/Institute New Curriculum & Materials	2	YES
	Classroom Lessons	Develop/Institute New Curriculum & Materials	2	YES
	READS Family Night Event	Involve Parents/Community Members	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
summer months. Students participate in comprehension activities, including both teacher-taught lessons at the end of the school year that model comprehension activities and independent lessons during the summer months. To promote family engagement, READS hosts a 90-minute family literacy event during which parents learn about the program, and intervention staff contact parents during the summer if their children have not returned completed independent lessons to the READS program.	Summer Parent Follow-up	Involve Parents/Community Members	2	YES
StartSmart K-3 Plus Program				
The StartSmart K-3 Plus program is an extended school year intervention that provides 25 days of summer educational services to students in grades K-3 to learn core academic content. The intervention aims to turn around high poverty and persistently low-performing schools by improving student achievement outcomes.	Delivery by certified teachers	Select/Evaluate Staff	4	YES
	Focus on literacy and numeracy	Develop/Institute New Curriculum & Materials	3	YES
	Parent involvement	Involve Parents/Community Members	4	YES
	Provide 25 additional days of school in summer	Institute Structural Change	4	YES
	Provide breakfast, lunch, and transportation	Institute Structural Change	4	YES
	Teacher professional development in literacy	Provide PD for Teachers	4	YES
Virginia Initiative for Science Teaching and Achievement (VISTA)				
The Virginia Initiative for Science Teaching and Achievement (VISTA) project provides professional development for elementary and secondary science teachers and evaluates the overall effectiveness of VISTA on participating teachers and their students.	Teacher Professional Development	Provide PD for Teachers	3	YES
	Coach Professional Development	Provide PD for Other School or Non-School Staff	3	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
VISTA aims to improve science teaching, through intensive professional development, and increase student learning throughout Virginia especially in high-need (high-poverty, high minority) schools.	Principal Involvement	Provide PD for Other School or Non-School Staff	3	YES
Development				
The Achievement Network (ANet)				
The Achievement Network (ANet) is a whole-school reform model designed to train teachers and school leaders in the use of student interim assessment data to inform instruction and curriculum choices. To enable teachers and school leaders to imbed data driven decision-making in their everyday practice, ANet provides them with interim assessments and data reports, coaching, and access to peer networks.	Aligned Assessments	Plan for and Support Assessment & Data Use	2	NO
	Logistical Support	Plan for and Support Assessment & Data Use	2	NO
	Coaching	Provide Coaching for Teachers	2	NO
	Network Supports	Support Staff Collaboration	2	YES
Advanced Placement (AP) Insight				
The overall goal of the Advanced Placement (AP) Insight program is to provide research-based, classroom-tested tools such as instructional activities, curriculum resources, and formative assessments, to support teachers and students to achieve greater success in the AP classroom. Innovative formative assessments, instructional strategies and professional learning tools are designed around the Challenge Areas of the AP course to focus classroom time where it matters most. Resources for the program are primarily web-based, but there are some opportunities for in-person professional development and collaboration around the use of resources.	Online Collaboration Forums	Support Staff Collaboration	2	YES
	Online Resources	Develop/Institute New Curriculum & Materals	2	YES
	Professional Learning	Provide PD for Teachers	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Advancement via Individual Determination (AVID)				
The Baboquivari Unified School District Wisdom Project employs two key strategies—the Advancement via Individual Determination (AVID) elective course and a comprehensive school-wide college-readiness strategy. AVID strategies include training teachers and administrators to help students succeed in school and prepare for higher education. The elective class is offered to selected students in each grade (6–12) for one period per day; students work on study and organizational skills and receive academic assistance from AVID-trained teachers and tutors. The school-wide program offers college-readiness activities for all students.	AVID Implementation	Support College/Career Readiness	2	NO
	AVID Training	Provide PD for Teachers and Other Staff	2	YES
	Enrichment Activities	Support College/Career Readiness	2	YES
	Technology Standardization and Training	Develop/Institute New Curriculum & Materials	1	YES
Around the Corner (ATC)				
Around the Corner (ATC) is a program for both preschool and kindergarten students that enables teachers to show children how the world works for concepts that are not possible to illustrate in a classroom with actual objects. The program includes computer activities and videos for children. Opportunities to view videos again at home (Home Links) provide the repeated experiences with language concepts and vocabulary that are necessary for language development. An additional component of the ATC intervention is interactive, video-based professional development for teachers. ATC is designed to provide a foundation in language and literacy, mathematics, science, listening and social skills, creative expression, and positive self-esteem through a holistic, thematic approach to instruction.	Communication meetings for principals and facilitators	Target Leadership Structures and Supports	2	YES
	Distribution of ATC materials	Develop/Institute New Curriculum & Materials	2	YES
	Teacher professional development	Provide PD for Teachers	2	YES
	Train Success for All coaches on new media	Provide PD for Other School or Non-School Staff	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Arts Achieve: Impacting Student Success in the Arts				
Arts Achieve: Impacting Student Success in the Arts aims to improve student achievement in targeted arts disciplines (dance, music, theater, and visual arts) and improve students' writing and English language arts skills through targeted professional development on the use of balanced assessment (formative and summative) and the use of data to drive instruction. The intervention translates the standards and information from assessments into classroom practices that support improved arts achievement for all students. The intervention also promotes innovations in students' and arts specialists' access to content and assessment feedback through the use of technology.	Professional Development	Provide PD for Teachers and Other Staff	3	NO
	On-site Consultancies	Provide Coaching for Staff	3	YES
	Classroom Planning and Instruction	Develop/Institute New Curriculum & Materials	3	NO
	School Planning and Arts Programming	Support Staff Collaboration	3	NO
Arts for Learning (A4L) Lessons				
Arts for Learning (A4L) Lessons is a supplemental literacy program that integrates arts into the language arts curriculum to raise student achievement in reading and writing and to develop learning and life skills.	Professional Development for Classroom Teachers	Provide PD for Teachers	1	YES
	Ongoing Teacher Support	Provide PD for Teachers	2	NO
	Professional Development for Teaching Artists	Provide PD for Other School or Non-School Staff	2	YES
	A4L Lessons Delivery by Classroom Teachers	Develop/Institute New Curriculum & Materials	2	YES
	A4L Artist Residency	Develop/Institute New Curriculum & Materials	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Bay State Reading Institute (BSRI)				
The Bay State Reading Institute (BSRI) is a whole-school reform model that aims to improve literacy for students in grades K-5. BSRI provides professional development and coaching to principals, teachers, and school-based reading coaches to help them implement a research-based core curriculum, frequent student assessment, differentiated instruction, and literacy interventions. The BSRI model emphasizes collaboration among staff in planning, teaching, and using data.	Coaching of Principal	Provide Coaching for Administrators	4	YES
	Material Adoption	Develop/Institute New Curriculum & Materials	4	YES
	Professional Development	Provide PD for Teachers	4	YES
	Coaching of School-based Literacy Coach	Provide Coaching for Staff	4	YES
Boston Teacher Residency (BTR)				
The Boston Teacher Residency (BTR) aims to improve student engagement and academic learning in Boston Public Schools' turnaround and low-performing schools through teacher and school leadership training and support. The intervention is structured around four key components: (1) restructuring of the residency program; (2) concentrating BTR graduates in turnaround schools; (3) having Clinical Teacher Educators provide training and content-specific induction; and (4) providing data analysts and products. By implementing these key components, BTR expects to increase teachers' professional capacity, create a coherent instructional guidance system, and create strong instructional leadership, all of which will help develop (a) conditions necessary for ambitious teaching and instruction that meets the needs of diverse learners and (b) an increased number of effective teachers.	Restructured Residency	Provide PD for Teachers	3	YES
	Concentration of BTR Graduates in Schools	Select/Evaluate Staff	3	NO
	Induction Support	Provide Coaching for Teachers	3	YES
	Data Support	Plan for and Support Assessment & Data Use	3	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Building Assets, Reducing Risks (BARR) Model				
The goals of the Building Assets, Reducing Risks (BARR) model are to increase student coursework success and improve achievement. The BARR model reduces social/emotional barriers to learning through (1) professional development for teachers, staff, and administrators and (2) restructuring of the 9th grade blocks served by teacher teams. Fundamental elements of the approach are creating positive, intentional relationships (staff to staff, staff to students, and students to students); collaborative problem solving by teacher teams; and using existing technology platforms and real-time student data to guide instructional action.	Professional development	Provide PD for Teachers	3	YES
	Restructuring	Institute Structural Change	3	YES
	Parent involvement	Involve Parents/Community Members	3	YES
	Developmental Assets curriculum	Develop/Institute New Curriculum & Materials	3	YES
	Block meeting review	Provide Services Targeting Individualized Learning	3	YES
	Risk review	Provide Services Targeting Individualized Learning	3	YES
	Whole student emphasis	Support Staff Collaboration	3	YES
	Contextual support	Target Leadership Structures and Supports	3	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Collaborative Organizational Model to Promote Aligned Support Structures				
The Collaborative Organizational Model to Promote Aligned Support Structures (COMPASS) is a district-level reform initiative that works to align and better train the district's seven types of support staff: Instructional Facilitators, Instructional Technologists, Exceptional Children Specialists, RtI Coordinators, Limited English Proficiency Specialists, Intervention Specialists, and Differentiation Specialists. Increasing the expertise of support staff is expected to lead to teachers receiving higher quality support, which then will increase their knowledge, skills, and performance. Ultimately, COMPASS's long-term goal is to improve the academic achievement of all students, with a focus on students with high needs, disabilities, and limited English proficiency.	Educative Critical Component	Provide PD for Teachers and Other Staff	3	YES
	Procedural/Pedagogical Critical Component	Target Leadership Structures and Supports	2	YES
CollegeYES				
CollegeYES is an innovative way for schools to offer a technology certification program to middle school students. Students demonstrate technology literacy by creating projects that meet the International Society for Technology in Education (ISTE) National Educational Technology Standards for Students (NETS*S). A cohort of students in a structured peer-mentoring program assist other students as well as teachers in implementing the two project-based learning technology projects that are required for certification. CollegeYES allows schools to integrate opportunities for improved technology literacy into existing classes or programs. The program	Middle School: CollegeYES Summer Camp	Provide PD for Teachers and Other Staff	3	YES
	Middle School: CollegeYES Professional Development	Provide PD for Teachers	4	YES
	Middle School: Student Technology Leadership Club	Institute Structural Change	3	YES
	Middle School: CollegeYES Science and College/Career Project	Develop/Institute New Curriculum & Materials	3	YES
	High School: CollegeYES Summer Camp	Provide PD for Teachers and Other Staff	4	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
encourages all students to complete technology projects that are creative and personally interesting.	High School: CollegeYES Professional Development	Provide PD for Teachers	5	YES
	High School: Student Technology Leadership Club	Institute Structural Change	5	NO
	High School: CollegeYES Science and College/Career Project	Develop/Institute New Curriculum & Materials	4	YES
Convergence Academies				
Convergence Academies is a whole-school reform model that infuses digital media arts and technology throughout the school to engage students in challenging, interest-driven learning; improve their academic achievement; and develop critical 21st century skills, such as communication, collaboration, problem solving, and creativity.	Connected Learning Supports	Institute Structural Change	2	YES
	Instructional Framework in Digital Media	Develop/Institute New Curriculum & Materials	2	YES
	Professional Learning Supports	Provide PD for Teachers	2	NO
Curriculum 2.0				
The objectives of Curriculum 2.0, or the Elementary Integrated Curriculum (EIC), were to (1) develop digital K-5 curricula and assessments that integrate the four core subjects, along with the arts, around critical thinking and academic success skills; (2) create an online learning community that supports professional development focused on successful implementation of the integrated curriculum and assessments; and (3) increase percentages of traditionally under-represented students performing at advanced levels. Montgomery County Public Schools designed a staggered implementation of Curriculum 2.0, implementing the new curricula one grade at a time starting with kindergarten.	Grades K-5 digital integrated curriculum and assessments	Develop/Institute New Curriculum & Materials	2	YES
	Online learning community for all grade levels	Support Staff Collaboration	2	NO
	Professional development	Provide PD for Teachers and Other Staff	2	YES
	District and school-level structures and supports	Target Leadership Structures and Supports	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Data-Driven Decision Making and Information Technology Curricula in Schools				
The Data-Driven Decision Making and Information Technology Curricula in Schools project aims to explore the effectiveness of a data-driven decision making process that incorporates information technology and specific content driven interventions (in literacy, math, and science) in grades K-12. The project is implemented in low-performing schools and targets students with low academic performance. The intervention includes a sequential focus on content areas beginning with a focus on language arts in elementary school, continuing in middle school with a focus on mathematics, and then in high school with a focus on science, technology, engineering, and mathematics.	Professional Development	Provide PD for Teachers and Other Staff	2	YES
	Instruction	Develop/Institute New Curriculum & Materials	2	YES
	Assessment	Plan for and Support Assessment & Data Use	2	YES
E3TL Teacher Performance Evaluation System				
E3TL Teacher Performance Evaluation System is an initiative to support the implementation of rigorous and comprehensive performance-based teacher evaluation. The initiative encompasses the development of professional teaching standards for educators. Additionally, it includes performance rubrics and provides professional development to teachers around the newly-created standards.	Development of Training/Materials	Develop/Institute New Curriculum & Materials	4	YES
	Training/Ongoing Support	Provide PD for Other School or Non-School Staff	4	YES
	Teacher Evaluations	Select/Evaluate Staff	4	YES
EngageMe P.L.E.A.S.E.				
EngageMe P.L.E.A.S.E. provides educators in middle and high school math and English language arts (ELA) classes with an integrated data system of	Integrated Learning Platform and Data Management System (LMS)	Plan for and Support Assessment & Data Use	2	NO

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
student information that identifies student needs and furnishes relevant learning activities. Teachers have access to a data system that provides them with timely feedback about the current academic performance and learning needs of their students, and resources for learning activities aligned to standards and learning preferences.	Develop Professional Learning Support Teams: Train the Trainer	Provide PD for Other School or Non-School Staff	2	YES
	Provide Teacher Professional Development	Provide PD for Teachers	2	YES
	School-Level Support	Develop/Institute New Curriculum & Materials	2	YES
	District-Level Support	Develop/Institute New Curriculum & Materials	2	YES
Every Child Ready (ECR)				
Every Child Ready (ECR) is a full-day preschool program designed for at-risk children. ECR focuses on universal screening, regular progress monitoring, differentiated instruction based on children's progress, specialized support plans for children with IEPs, and professional development and individual coaching for teachers. ECR is theorized to enhance children's language, literacy, and numeracy development.	In-Class Coaching	Provide Coaching for Teachers	3	YES
	On-Site Professional Development	Provide PD for Teachers	3	YES
	Teacher Education/Degree	Select/Evaluate Staff	3	YES
	Classroom Instruction	Develop/Institute New Curriculum & Materials	3	YES
Everyday Arts for Special Education (EASE)				
Everyday Arts for Special Education (EASE) provides professional development to special education teachers and is designed to improve student achievement in the areas of communication, socialization, academic learning, and arts proficiency through integrated, arts-based approaches.	Professional Development Workshops	Provide PD for Teachers and Other Staff	2	YES
	Collaborative Classroom Modeling (CCM)	Provide Coaching for Teachers	2	YES
	On-Site Professional Development (OSPD)	Provide PD for Teachers	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
	Classroom Instruction (CI)	Develop/Institute New Curriculum & Materials	2	YES
Exceptional Coaching for Early Language and Literacy - Enhanced (ExCELL-E)				
The i3-funded intervention extended and refined the Exceptional Coaching for Early Language and Literacy (ExCELL) preschool professional development model so that it (a) integrated technology into the training protocol to allow more teachers and children to be served; (b) addressed the needs of English Language Learners (ELLs) as well as native speakers of English; and (c) served teachers and children in preschool, kindergarten, and first grade. By improving teachers' instructional strategies, particularly those related to language and literacy practices, the intervention aimed to improve students' language and literacy outcomes.	Coaching feedback: Coaches review videotapes of teachers' use of module strategies and provide feedback that is comprehensive, direct, and congruent w/ ExCELL-E strategies	Provide Coaching for Teachers	1	YES
	Progress monitoring: Teachers monitor child progress and adjust instruction	Plan for and Support Assessment & Data Use	1	YES
	Teacher PD: Teachers participate in and master online training modules	Provide PD for Teachers	1	YES
	Teacher practice: Teachers use key techniques with frequency and quality in classrooms	Develop/Institute New Curriculum & Materials	1	NO
The Expository Reading and Writing Course (ERWC)				
The Expository Reading and Writing Course (ERWC) is a 12th grade English course developed by a task force of California State University faculty and high school educators to improve the academic	Curriculum Materials	Develop/Institute New Curriculum & Materials	1	YES
	Professional Development	Provide PD for Teachers	1	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
literacy of high school seniors and thereby reduce the need for English remediation in college. The program is disseminated to schools and English teachers through professional development in which teachers further improve their pedagogical skills to teach the ERWC.	Teaching of Curriculum	Develop/Institute New Curriculum & Materials	1	NO
Facilitating Long-Term Improvements in Graduation and Higher Education for Tomorrow (FLIGHT)				
The goal of Facilitating Long-Term Improvements in Graduation and Higher Education for Tomorrow (FLIGHT) is to increase the extent to which low-income students with academic promise are prepared for, enrolled in, and successful in college. Prepaid student scholarships are a key input to the FLIGHT model; each participating student is guaranteed a four-year college scholarship provided he/she maintains good standing in the program. Pre-service activities include training new mentors; training FLIGHT staff in the creation and distribution of Student Detail Reports; and creating lesson plans for each of the six college access workshops and the three supplemental workshops (nine annual workshops in total). Student service activities are provided directly to students (e.g., wrap around case management; one-on-one mentoring; college access workshops; supplemental workshops) or to their parents and guidance counselors in the form of the Student Detail Reports.	Train mentors	Provide PD for Other School or Non-School Staff	2	YES
	Train FLIGHT staff on application of Student Detail Report (SDR)	Provide PD for Other School or Non-School Staff	2	YES
	Develop lesson plans for workshops	Develop/Institute New Curriculum & Materials	2	YES
	Case management meetings	Provide Services Targeting Individualized Learning	2	YES
	Student mentoring	Provide Services Targeting Individualized Learning	2	YES
	College Access and Success Workshops (CASWs)	Support College/Career Readiness	2	YES
	Supplemental Student Workshops (SWs)	Support College/Career Readiness	2	YES
	SDR	Plan for and Support Assessment & Data Use	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Florida Master Teacher Initiative (FMTI)				
The Florida Master Teacher Initiative (FMTI) supports teachers and administrators through a job-embedded master's degree program, a Teacher Fellows program, a Principal Fellows program, and a Summer Leadership Institute. The goal of FMTI is to foster professional learning communities, teacher professionalism, and improved classroom practices, and ultimately to improve students' reading and math achievement.	Early Childhood Education Graduate Program	Provide PD for Teachers	3	NO
	Teacher Fellows Program	Provide PD for Teachers	3	YES
	Principal Fellows Program	Provide PD for Other School or Non-School Staff	3	NO
	Summer Leadership Institutes	Provide PD for Teachers and Other Staff	3	NO
InnovateNYC Ecosystem				
InnovateNYC Ecosystem aims to improve student achievement, engagement, and self-efficacy by improving the quality of education technology tools that teachers use in the classroom, and thereby increasing teachers' use of technology. The intervention employs a bottom-up strategy through which teachers and other educators are directly engaged in the process of identifying student learning challenges and selecting the solutions to address those challenges.	School-App Matching	Develop/Institute New Curriculum & Materials	1	NO
	Teacher/Developer Engagement	Provide PD for Teachers	1	YES
	Teachers' Use of Apps as Prescribed	Develop/Institute New Curriculum & Materials	1	NO
Innovations in Early Mathematics				
The overarching goal of the Innovations in Early Mathematics program is to use a whole teacher approach to professional development to help students in Pre-K to 3rd grade reach or exceed state learning standards in mathematics. The intervention includes summer institutes, learning labs, and onsite	Learning labs	Provide PD for Teachers	2	YES
	Summer institutes	Provide PD for Teachers	1	NO
	Coaching	Provide Coaching for Teachers	2	YES
	Grade level meetings	Support Staff Collaboration	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
coaching for teachers, as well as leadership meetings to address three significant problems confronting mathematics education and learning: (1) the mathematics achievement gap between low-income, minority students and their more advantaged peers; (2) the low level of competence in math teaching of early childhood and early elementary teachers; and (3) the limited and ineffective in-service opportunities for early childhood and elementary teachers to develop mathematics competencies.	Leadership academy	Target Leadership Structures and Supports	2	YES
Integrating English Language Development (ELD) and Science				
The Integrating English Language Development (ELD) and Science intervention is designed to close the achievement gap between students with limited English proficiency and their peers at the K-5 level by supporting adoption of an integrated ELD and science instructional program. The intervention includes a professional development program comprised of workshops, study groups, and leadership development activities. The expectation is that strategically designed professional development focused on integrated ELD and science will increase the percentage of highly effective teachers, as measured by both teacher performance and student achievement in ELD and science.	Professional development (summer workshops and study groups)	Provide PD for Teachers	2	YES
	Curriculum units	Develop/Institute New Curriculum & Materials	2	YES
	Teacher and district leadership	Target Leadership Structures and Supports	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Internet-based Reading Apprenticeship Improving Science Education (iRAISE)				
The year-long Internet-based Reading Apprenticeship Improving Science Education (iRAISE) program seeks to provide completely online literacy professional development to approximately 100 teachers from 20 schools in Pennsylvania and Michigan. iRAISE starts with a 5-day foundational training, followed by monthly Ignite Sessions and Professional Learning Communities (PLCs) aimed at influencing teacher instruction and increasing teacher knowledge of literacy instruction. The goal is to improve students', especially high-needs students', general reading literacy skills and ability to engage in and understand a variety of scientific texts.	Delivery of PD Content Adheres to Standards	Provide PD for Teachers	1	YES
	Delivery of Professional Development by Program Developers	Provide PD for Teachers	1	YES
	Teacher Attendance at Professional Development and Program Activities	Provide PD for Teachers	1	NO
Making Time for What Matters Most				
Making Time for What Matters Most is a school-wide intervention consisting of three major elements: (1) provide structures and supports to facilitate student mastery of academic material and successful completion of all core courses in one year or less; (2) provide a range of personalized supports to students to increase engagement in school and promote college readiness; and (3) improve teachers' pedagogical and student support practices to maximize the effectiveness of increased learning time. The intervention aims to reduce dropout rates, increase academic achievement, and increase high school graduation rates while also strengthening students' college readiness skills and increasing the percentages of students who graduate high school and go on to college.	Flexible Student Scheduling	Institute Structural Change	2	YES
	College Access Time	Support College/Career Readiness	2	NO
	Professional Learning Communities	Support Staff Collaboration	2	NO

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
New England Network for Personalization and Performance (NETWORK)				
The New England Network for Personalization and Performance (NETWORK) aims to prepare high school students for postsecondary success by personalizing learning for all students in participating schools. The NETWORK accomplishes this through a collaborative, peer-mentored approach to professional development that supports the development of instructional experiences and related assessments and rubrics; a change leadership team in each network school that shepherds systemic change; a Performance Assessment Review Board (PAR), made up of nationally recognized experts who make visits to the network schools, that assesses student experiences and provides feedback for improvement; and a Project Steering Committee that oversees the project, including monitoring progress toward goals and adherence to the timeline and budget.	Development of inquiry-based curricular units and performance-based assessments, along with rubrics aligned with state content standards	Develop/Institute New Curriculum & Materials	5	YES
	Establishment of a broadly representative Project Steering Committee (PSC) responsible for project oversight, communication, dissemination, and evaluation	Target Leadership Structures and Supports	5	YES
	Establishment of a Performance Assessment Review Board (PAR) comprised of nationally recognized experts	Target Leadership Structures and Supports	5	YES
	Establishment of school-based change leadership (CL) team with the support of a school change coach	Target Leadership Structures and Supports	5	YES
	Objective evaluation process used to provide feedback and documentation for the Project Steering Committee (PSC)	Target Leadership Structures and Supports	5	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Oakland Accelerates				
Oakland Accelerates is a program to bring the College Board's EXCELerator district-wide process for college readiness to Oakland Unified School District. The intervention includes PD for teachers, leadership, and counselors to support college-going culture; support and coaching for College Readiness Specialists; technical assistance to develop policy and infrastructure to support college-going culture; and key resources to support teachers, students, and families to build an understanding of college preparation.	AP Teacher Professional Development	Provide PD for Teachers	2	YES
	Capacity and Policy Guidance	Support College/Career Readiness	2	NO
	College Readiness Specialist Coaching and Support	Provide Coaching for Staff	2	YES
	District Diagnostic	Support College/Career Readiness	1	YES
Ounce of Prevention Fund Professional Development Initiative (PDI)				
The primary goal of the Ounce of Prevention Fund's Professional Development Initiative is "to build birth-to-five teachers' capacity to design and deliver standards-aligned, data-driven instruction and to close developmental and learning gaps among high needs students to support their kindergarten readiness through simultaneous job-embedded PD for teachers, leaders, and their coaches.	Center Leader Professional Development	Provide PD for Other School or Non-School Staff	2	YES
	Coach Induction (Year 1) and Community of Practice Implemented by the Sponsor Organization (Year 2 and 3)	Provide PD for Other School or Non-School Staff	3	YES
	Coach Professional Development	Provide PD for Other School or Non-School Staff	2	YES
	Direct Supervisor Professional Development	Provide PD for Other School or Non-School Staff	2	YES
	Professional Development Initiative Implementation	Provide Coaching for Mixed Group	2	YES
	Teacher Professional Development	Provide PD for Teachers	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Pathways to STEM Initiative (PSI)				
The Pathways to STEM Initiative uses an integrative approach to introduce a STEM curriculum, and includes external mentoring, teacher professional development, and extra-curricular opportunities for students. The program begins in middle school (Level 1–Discover) when all students are introduced to STEM through Gateway to Technology (GTT) curriculum. Interested students can continue involvement by exploring STEM-related careers through extra-curricular activities (Level 2–Explore), and pursue STEM coursework and internships in high school and beyond (Level 3–Pursue).	Classroom Technology	Develop/Institute New Curriculum & Materials	2	YES
	Implementation of Clark County School District & Project Lead The Way Curricula	Develop/Institute New Curriculum & Materials	2	NO
	Implementation of STEM Club	Support College/Career Readiness	2	NO
	Math & Science Tutoring	Provide Services Targeting Individualized Learning	2	YES
	Ongoing Teacher Support	Provide PD for Teachers	2	NO
	STEM Summer Camp	Develop/Institute New Curriculum & Materials	2	YES
	Teacher Professional Development	Provide PD for Teachers	2	YES
	Weekly Sessions with STEM Professionals	Support College/Career Readiness	2	NO
Problem-Based Learning (PBL)				
Problem-Based Learning (PBL) is a whole-school problem-based learning intervention that aims to (1) redefine the relationship between the high school and the community by connecting students with local professionals in STEM fields to provide real-world validation for students' college and career questions and by having counselors help every student to complete at least one college application; (2) provide increased instructional time for limited English	Focusing on 1st Generation College Bound Students and Developing a PBL Laboratory: Starting Strong	Develop/Institute New Curriculum & Materials	4	YES
	Increasing Rigor and Focusing on 21st Century Skills in Curriculum: PBL Course Implementation	Develop/Institute New Curriculum & Materials	5	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
proficient students and students with disabilities, with a focus on mathematics and one-to-one mentoring; (3) design and implement a scalable, sustainable, 21st century skills-based program that includes PBL curriculum in both AP and non-AP courses, specific supports for struggling students, and professional development that trains teachers to design and implement effective problem-based curricula. The aim is to increase STEM learning in the classroom in order to improve students' college and career readiness and increase their performance on state math and science, as well as AP, exams.	Developing Distributed Expertise to Support a Rigorous Curriculum: Leadership Team	Target Leadership Structures and Supports	5	YES
	Development/Adaptation/Adoption of Research-Based Framework for PBL Design, Implementation and Evaluation	Develop/Institute New Curriculum & Materials	5	YES
	Monitoring and Supporting Career Readiness: EPIC Campus Ready Survey, Application Support	Provide Services Targeting Individualized Learning	5	YES
	Increasing Teacher Pedagogical Expertise (SILT)	Provide PD for Teachers	5	YES
	Designing a Rigorous Curriculum: PBL Courses through PBL Design Teams	Support Staff Collaboration	5	YES
PTA Comunitario				
PTA Comunitario seeks to establish successful partnerships between community-based organizations and schools with minority and low-income families in order to increase student achievement, particularly for students who are low-income, Hispanic, or limited English proficient (LEP). The project has three main objectives: (1) establish PTA Comunitarios in South Texas public school	Phase I: Community Organizing and Mobilization	Involve Parents/Community Members	2	YES
	Phase II: Initiation and Relationship Setting	Involve Parents/Community Members	2	YES
	Phase III: Taking Collaborative Action for Student Success, Educational Leadership Projects	Involve Parents/Community Members	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
districts with a minimum of 20 active members per PTA Comunitario; (2) establish a partnership between PTA Comunitarios and schools through training and technical assistance; and (3) facilitate the development of at least one educational leadership project informed by actionable data for each PTA Comunitario and partner school.	Phase IV: Instituting Sustainable Connections	Involve Parents/Community Members	2	YES
Public School Choice Initiative (PSCI)				
The Public School Choice Initiative (PSCI) supports the Los Angeles Unified School District's Bold Competition, which aims to create a diverse portfolio of innovative schools, supported and sustained, that can respond to the needs of the local community and turnaround chronically low performing schools. The intervention identifies "focus schools" (existing schools with low performance) and "relief schools" (newly constructed schools designed to ease overcrowding in low-performing schools), and includes elementary, middle, and high schools. Schools submit proposals for funding programs, which may include STEM academies; bilingual immersion programs; arts-focused schools; or the Linked Learning model, which provides students a college-preparatory academic curriculum, technical education, work-based learning opportunities, and various socio-economic and academic supports.	Identification of Public School Choice (PSC) schools	Select/Evaluate Staff	2	YES
	Facilitation of stakeholder involvement	Involve Parents/Community Members	2	YES
	Facilitation of support and oversight	Provide PD for Other School or Non-School Staff	2	YES
	Accountability and monitoring	Select/Evaluate Staff	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Responding Effectively to Assessments with Curriculum and Teaching (REACT)				
Responding Effectively to Assessments with Curriculum and Teaching (REACT) focuses on a district-wide implementation of support structures of data coaching, data informed instruction, and data professional learning community teams with the goal of increasing teacher effectiveness, improving student achievement, and closing the achievement gap as measured by scores on California state tests. As part of REACH, the Del Norte County Unified School District implemented data-driven instruction in K to 12th grade classrooms using data teams with site/district coordinators.	Administrators	Target Leadership Structures and Supports	2	YES
	Classroom Teachers	Plan for and Support Assessment & Data Use	2	YES
	Data Coaches	Plan for and Support Assessment & Data Use	2	YES
Rio Grande Valley Center for Teaching and Leading Excellence: New Teacher Training (NTT)				
The Rio Grande Valley Center for Teaching and Leading Excellence (RGV) New Teacher Training (NTT) programs aim to increase the supply of effective teachers by offering new teachers a five-day summer institute, ongoing professional development, and periodic coaching from an instructional coach during their first year. Through these services, NTT aims to improve teacher outcomes (e.g., job satisfaction, self-efficacy, leadership skills, teacher retention) as well as student outcomes (i.e., reading, math, social studies, and science achievement on the Texas state assessments).	New Teacher Recruitment and Selection	Select/Evaluate Staff	3	YES
	New Teacher Training	Provide PD for Teachers	3	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Rural Math Excellence Partnership (RMEP)				
The Rural Math Excellence Partnership (RMEP) Project aimed to develop a rural workforce qualified for science, technology, engineering, and math (STEM) jobs in their local communities. The goal of the RMEP Project was to develop and implement a model of shared responsibility among families, teachers, and communities in rural areas to prepare students to be successful in advanced high school and postsecondary STEM studies. Via a gap analysis and development of a Math Advanced Study guide; professional development and ongoing coaching for participating teachers; Family Math Nights conducted by teachers; a project website and social media presence; community-based STEM events; and access to technology for students, the project aimed to improve student achievement in math, attitudes toward math, and interest in STEM careers; to increase the percentage of 10th grade students enrolled in advanced STEM courses; and ultimately to ensure that students leave school ready, at a minimum, to enroll in a certificate program for a technician-level career in STEM-related fields.	Access to technology	Develop/Institute New Curriculum & Materials	2	NO
	Community-based STEM career event	Involve Parents/Community Members	3	YES
	Family math night	Involve Parents/Community Members	2	YES
	Identification of foundational math content gaps	Develop/Institute New Curriculum & Materials	2	YES
	Professional development and ongoing coaching	Provide PD for Teachers	2	YES
	Project website	Institute Structural Change	2	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
School of One				
School of One provides middle school students with personalized, effective, and dynamic math instruction that is tailored to their academic needs, interests, and learning preferences. Through an adaptive technology platform, the intervention incorporates multiple instructional modalities (including live, online, and collaborative instruction), which aim to improve math achievement, student academic behaviors and attitudes, teacher and student technological knowledge, and teacher fulfillment. The ultimate goal is to increase student learning and preparation for high school, college, work, and life.	Operational Aspects	Institute Structural Change	2	YES
	Required Resources	Develop/Institute New Curriculum & Materials	2	YES
	Sufficient Time	Institute Structural Change	2	YES
Schools to Watch (STW): School Transformation Network Project				
The Schools to Watch (STW): School Transformation Network Project is a whole school reform model for the middle grades consisting of a guiding framework for high performance, a multi-layered system of support, and school improvement strategies. The intervention provides a variety of resources including tools for assessment, goal setting, action planning and monitoring; a STW-trained coach to provide ongoing assistance; a principal leadership coach; a mentor school (i.e., a current STW school); a cross-school peer network; and professional development for teachers. The goals of the intervention are to build the capacity of persistently low-performing middle-grades schools	Mentor school	Provide Coaching for Staff	4	YES
	Implement early indicators program	Plan for and Support Assessment & Data Use	4	NO
	Participate in national and state STW network	Target Leadership Structures and Supports	4	YES
	Implement professional learning communities (PLCs)	Support Staff Collaboration	4	YES
	Focused professional development designed to build a learning community and address the needs of students who need additional support	Provide PD for Teachers and Other Staff	4	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
by: 1) developing organizational structures, norms, and processes for continuous improvement; 2) improving academic rigor; 3) promoting equity for all students; and 4) developing an array of supports designed to meet the developmental needs of students. The outcomes of the project are to improve the educational practices, experiences, and outcomes of low-performing middle-grades schools.	Create a powerful vision for high performance using the STW criteria	Target Leadership Structures and Supports	4	YES
	Engage in an in-depth assessment and planning process using the STW criteria	Provide Coaching for Mixed Group	4	YES
	STW coach	Provide Coaching for Mixed Group	4	YES
	Principal mentor	Provide Coaching for Administrators	4	NO
Science, Technology, Engineering and Math Education for the 21st Century (STEM21)				
Science, Technology, Engineering and Math Education for the 21st Century (STEM21) builds on the success of the Connecticut Career Choices program and aims to improve student achievement, interest, and self-efficacy in STEM, as well as college enrollment. The intervention provides professional development to teachers and delivers a STEM course sequence in an interactive environment in which students progress through online coursework that is guided by teachers and integrated with authentic learning experiences.	Teacher Professional Development	Provide PD for Teachers	3	YES
	Grade Level Implementation	Develop/Institute New Curriculum & Materials	3	YES
	Program Level Implementation	Develop/Institute New Curriculum & Materials	3	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Spheres of Proud Achievement in Reading for Kids (SPARK) Program				
Spheres of Proud Achievement in Reading for Kids (SPARK) program aims to improve reading achievement for students in grades K-3 by providing in-school tutoring and increasing parental engagement. SPARK students are provided one-on-one tutoring during the school day for 30 minutes, up to three times per week, for two years. At each SPARK site, a program manager, who is also a certified teacher, oversees the tutors. The tutoring component of SPARK is loosely based on the Reading Recovery program, which focuses on in-school tutoring with lesson plans written, and assessments analyzed, by the tutors themselves. To execute the family engagement component, each site has a parent partner who works with each participating student's family. Their work is designed to bridge the divide between school and home by translating literacy concepts, educating families about a variety of literacy activities, and validating the literacy practices already happening in the home. Parent partners stay connected with families through a monthly newsletter, monthly family events at each site, and phone calls or emails. These communications are designed to keep families aware of student progress in SPARK, help families promote literacy at home, and address any attendance issues that arise during the program. Parent partners also conduct home visits for all students twice during the summer between their first and second year of participation.	In-school tutoring	Provide Services Targeting Individualized Learning	1	YES
	Family engagement	Involve Parents/Community Members	1	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
STEM Learning Opportunities Providing Equity (SLOPE)				
STEM Learning Opportunities Providing Equity (SLOPE) aims to improve academic achievement in mathematics and increase interest in STEM courses through the use of real-world, project-based instruction with a focus on STEM topics. The intervention consists of professional development and coaching for teachers on the use of project-based Algebra and pre-Algebra curricula. Rising eighth graders who are low performing students are eligible to participate in a 4-week summer program (Summer College Awareness and Math Proficiency, or C.A.M.P.) focused on pre-Algebra concepts. The core curriculum is complemented by a college awareness curriculum that highlights pathways to STEM courses and careers.	C.A.M.P. Curricular Materials	Develop/Institute New Curriculum & Materials	2	YES
	C.A.M.P. Professional Development	Provide PD for Teachers	2	YES
	C.A.M.P. Coaching	Provide Coaching for Teachers	2	NO
	C.A.M.P. College Awareness Curriculum	Develop/Institute New Curriculum & Materials	2	NO
	Algebra I Curricular Materials	Develop/Institute New Curriculum & Materials	2	YES
	Algebra I Professional Development	Provide PD for Teachers	2	YES
	Algebra I Coaching	Provide Coaching for Teachers	2	YES
	Algebra I College Awareness Curriculum	Develop/Institute New Curriculum & Materials	2	NO
	C.A.M.P. Teacher Implementation of Classroom Intervention, Unit 1	Develop/Institute New Curriculum & Materials	2	NO
	C.A.M.P. Teacher Implementation of Classroom Intervention, Unit 2	Develop/Institute New Curriculum & Materials	2	NO
	C.A.M.P. Teacher Implementation of Classroom Intervention, Unit 3	Develop/Institute New Curriculum & Materials	2	NO

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
	Algebra I Teacher Implementation of Classroom Intervention, Unit 1	Develop/Institute New Curriculum & Materials	2	YES
	Algebra I Teacher Implementation of Classroom Intervention, Unit 2	Develop/Institute New Curriculum & Materials	2	YES
	Algebra I Teacher Implementation of Classroom Intervention, Unit 3	Develop/Institute New Curriculum & Materials	2	NO
STEM Summer Learning with VEX Robotics				
The Baltimore STEM Summer Learning Program with VEX Robotics aims to increase mathematics knowledge and interest in STEM careers and college majors by providing professional development in mathematics and robotics to teachers, as well as 5 weeks of summer instruction in these subjects to students. Specifically, the project provides professional development to participating teachers to equip them to lead students in construction of a robot and activities to prepare students for a robotics competition, as well as high quality delivery of math instruction. The robotics component is expected to increase student engagement (including attendance) and perception of the relevance of mathematics, leading to increased student effort and math achievement.	Instruction in mathematics	Develop/Institute New Curriculum & Materials	3	YES
	Professional development in robotics	Provide PD for Teachers	3	NO
	Professional development in math	Provide PD for Teachers	2	NO
	Instruction in robotics	Develop/Institute New Curriculum & Materials	3	YES

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
Texas Tech University "Tech Teach" Program				
The Texas Tech University "Tech Teach" Program is a competency-based teacher and school intervention. Part A focuses on competency-based strategies at the teacher candidate and program levels via three main components: (1) trainee screening and recruitment; (2) competency-based pre-certification coursework; and (3) student teaching using competency-based training. Part B focuses on implementing a model of technology-enabled competency-based teacher development in mathematics via three main components: (1) individual pre-and post-conferences; (2) a professional development cycle; and (3) Texas Instruments MathForward curriculum. Both parts include professional development with coaching, multiple assessments of teaching, and professional learning communities.	Competency-Based Pre-Certification Coursework	Provide PD for Teachers	2	YES
	Student Teaching Using Competency-Based Framework	Provide PD for Teachers	2	YES
	Trainee Recruitment and Screening	Select/Evaluate Staff	2	YES
Transforming Teacher Talent (t3) System				
Aspire Public Schools aims to increase student achievement in grades 3-11 through its Transforming Teacher Talent (t3) system. This intervention uses a train-the-trainer model in which experienced and highly effective Aspire teachers	Charter Management Organization provides adequate support	Provide PD for Teachers	1	YES
	PD content library	Provide PD for Teachers	1	NO

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
receive training on the use of three t3 tools and then provide professional development and coaching to other Aspire teachers. Tools include a professional development content library, peer observations cycles, and a virtual professional learning community. Under its i3 grant, Aspire implemented t3 in all 35 Aspire schools in California.	Peer observation/walkthrough in BloomBoard (an online personalized/customized observation data collection tool through which principals and peer observers conduct informal observations and provide targeted, frequent feedback and exposure to new teaching strategies).	Provide Coaching for Teachers	1	NO
	Virtual collaborations	Support Staff Collaboration	1	NO
Turnaround with Increased Learning Time (TILT)				
The Turnaround with Increased Learning Time (TILT) model was implemented in two low-performing Boston middle schools. Both of the TILT schools redesigned their school days to include 300 more hours per year of additional instructional time for all students. The increased time was allocated to meeting three goals: more time for core academic instruction that is data-based and targeted to meet the individual needs of students; more time for robust enrichment program supports by strong local partnerships; and more time for teachers to collaborate to improve their effectiveness and instructional quality.	More Time for Academics	Institute Structural Change	3	NO
	More Time for Enrichment	Institute Structural Change	3	NO
	Planning Component	Institute Structural Change	3	NO
	Sustainability Component	Institute Structural Change	1	NO
	Teacher Collaboration	Support Staff Collaboration	3	NO

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
We Are A Village				
We Are A Village focuses on improving the school readiness and academic achievement of young children, from preschool to third grade, by integrating parent and family engagement strategies into classrooms. Project activities are designed to support the successful transition of children from preschool to elementary school by providing opportunities for cross-system, cross-grade training for parents, families, and staff; parent and family school leadership opportunities; peer support school transition groups; and the Incredible Years (IY) Teacher Classroom Management and BASIC Parenting Programs. By conducting targeted outreach to parents and families, including parents and families of children with disabilities and/or limited English proficiency, the project encourages parents and families to become actively engaged in school activities.	"Parents are Power!" resource centers	Involve Parents/Community Members	2	YES
	Cross-system educational workshops for parents	Involve Parents/Community Members	2	YES
	Cross-system training for staff & Parent Peer Navigators	Provide PD for Teachers and Other Staff	2	NO
	Direct parent outreach	Involve Parents/Community Members	2	NO
	Incredible Year (IY) Program	Provide PD for Teachers and Other Staff	2	YES
	Kindergarten transition support for families	Involve Parents/Community Members	2	YES
	Parent leadership	Involve Parents/Community Members	2	YES
Write Up!				
Write Up! is intended to enhance the use of high-quality assessments and standards and improve teacher effectiveness through a variety of research-based program strategies including support of Step Up to Writing; launch of My Access!, an artificial intelligence scoring engine; and provision of an online writing program (i.e., United Streaming by Discovery Education). Write Up! also includes professional development to create and integrate online courses and staff development focused on improving student achievement through the use of	Teacher participation in trainings	Provide PD for Teachers	3	NO
	Teacher collaboration	Support Staff Collaboration	3	YES
	Classroom observations	Select/Evaluate Staff	3	YES
	Coach/Teacher on Special Assignment (TSA) participation in trainings	Provide PD for Teachers and Other Staff	3	YES
	Online lessons and courses in writing instruction	Develop/Institute New Curriculum & Materials	3	NO

Intervention Description	Key Component Name	Key Component Domain	Number of Times Fidelity Was Measured	Implemented with Adequate Fidelity?
technology. Write Up! is theorized to increase teacher effectiveness; upper elementary, middle, and high school student writing and English language arts achievement; and secondary students' college and career knowledge.	Online lessons and courses in career and college readiness	Support College/Career Readiness	3	NO
	Student use of writing software	Develop/Institute New Curriculum & Materials	3	NO

Appendix B: i3 Review Protocol, Version 1.0

The i3 Review Protocol Version 1.0 (April 7, 2016) was developed to apply the What Works Clearinghouse™ (WWC) Standards to reviews of the impact studies from i3-funded evaluations. The i3 Review Protocol defines the scope and purpose of the review of i3-funded evaluations relative to the WWC Standards and defines all topic-specific applications of the WWC Standards for the review of i3 evaluations. The reviews use a single protocol both to ensure consistency across reviews and to include evaluations whose topics fall beyond the scope of existing WWC review protocols. This protocol is designed to be as consistent with WWC review protocols as possible using a parsimonious set of rules that can be applied across topics covered by i3-funded evaluations.

Purpose Statement

The purpose of this Review Protocol is to describe how the Abt Team applies the WWC Standards in the review of the i3-funded evaluations. Reviews conducted by the Abt Team assign *unofficial* WWC ratings to i3 impact evaluations. The ratings are unofficial because they are not assigned by the WWC and will not be included in the WWC database of reviewed studies.[61] In addition, while the WWC reviews only publicly available reports, the Abt Team collects the information necessary for the reviews directly from the evaluators and includes in the i3 reviews only the prespecified impact analyses for each evaluation.

The reviews assign one of four ratings to each estimated impact reported to the Abt Team:

- Unofficially Meets WWC Standards without Reservations.
- Unofficially Meets WWC Standards with Reservations.
- Unofficially Does Not Meet WWC Standards.
- Unofficially Ineligible for WWC Review

Eligibility Criteria

Eligible Interventions

All i3-funded interventions are eligible for review under this protocol.

Eligible Research

All evaluations supported by i3 funding are eligible for review under this protocol.

Eligible Designs

Randomized controlled trials (RCT) and quasi-experimental designs (QED) are eligible for review under this protocol. Regression discontinuity designs (RDD) and single-case designs (SCD) are not eligible for review under this protocol but will be reviewed by the WWC. All other designs are classified as ineligible for review under this protocol.

---

[61] The WWC will conduct official reviews of publicly available i3 evaluation reports as they are released. Once completed, the reviews can be found at https://ies.ed.gov/ncee/wwc/ReviewedStudies.

Eligible Analyses

Analyses that estimate the effect of an educational intervention relative to a comparison condition on an outcome are eligible for review under this protocol. The comparison could be to business-as-usual or to an alternative intervention.

Eligible Outcomes

All outcomes are eligible for review under this protocol. These outcomes may include student outcomes (including, but not restricted to, individual or aggregate measures of student achievement) as well as outcomes for teachers and other parties that could be affected by the i3 interventions.

Eligible Subgroups

Subgroup analyses are eligible for review; the reviews are not limited to particular subgroups.

Evidence Criteria

Estimated impacts from the impact study that meet the eligibility criteria listed above are assessed against the standards described in the *WWC Procedures and Standards Handbook, version 3.0,* Section III: Screening and Reviewing Studies (pp. 7–21). This section describes the application of these standards in the review of i3-funded evaluations.

Level of Causal Inference

For evaluations that assign *clusters* of individuals to groups (e.g., treatment and control), the evidence criteria for sample attrition and baseline equivalence described below account for the level at which inferences are made from the impact study results.

Causal inference at the cluster level: Inference from estimated impacts about an intervention's effect on clusters of students (e.g., clusters of students who are in the same classrooms, schools, or districts at the point the outcome is measured).

Causal inference at the individual level: Inference from estimated impacts about an intervention's effect on individuals (e.g., individual students or teachers).

Sample Attrition

The reviewers apply the WWC's liberal attrition standard unless there is reason to believe that attrition will yield substantial bias in the estimated effect—more specifically, when attrition is endogenous to the intervention and data are not missing at random. The reviewers assume that attrition will yield substantial bias in the estimated effect and therefore apply the conservative attrition standard if **all three** of the following conditions hold:

1. **The sample consists exclusively of high school students.** High school students typically have more choice about whether to remain in school, change schools, or change classes: these decisions are not likely to be random.
2. **Mobility decisions affect attrition—that is, whether a student is included in the analytic sample.** For example, this condition would be satisfied if the i3 local evaluator collects outcome data on outcomes only from the schools initially included in the impact study sample so that no outcome data are collected on students who moved out of these schools.

3. **Mobility decisions almost surely affect the value of the outcome variable.** The reviewers treat the following outcomes as almost surely affected by the mobility of high school students across schools or into and out of school (i.e., dropping in and dropping out): graduation, credit completion, promotion to the next grade level, dropout. The reviewer reserves the right to designate other related variables as being almost surely affected by student mobility.

Under these conditions, the reviewer applies the conservative attrition standards because an intervention's effect on student mobility could produce non-random differential attrition and biased impact estimates.[62]

To calculate attrition rates in RCTs, the reviewer divides the number of units in the analytic sample by the number of units that were in the sample at the point of random assignment.[63] For RCTs that randomize individuals, attrition is deemed acceptable if the combination of overall and differential attrition for the individuals that were randomized satisfies the selected standards (liberal or conservative). For RCTs that randomize clusters of individuals, attrition is deemed acceptable:

- *for causal inference at the individual level* if the analytic sample excludes joiners[64] and the combination of overall and differential attrition satisfies the selected standards *for both the individuals* and *the clusters that were randomized*.
- *for causal inference at the cluster level* if the analytic sample includes joiners and the combination of overall and differential attrition satisfies the selected standards *for the clusters that were randomized*.

For more information about the attrition standards, see the *WWC Procedures and Standards Handbook version 3.0*, Section III: Subsection B.2, "Sample Attrition: Is the combination of overall and differential attrition high?" (pp. 11–15).

Baseline Equivalence

If the impact study design is an RCT with high attrition or a QED, the reviewer assesses whether the treatment and comparison groups in the analytic sample are equivalent at baseline. For comparison group designs in which assignment occurs at the cluster level (e.g., QEDs that compare student outcomes in schools that received the intervention to student outcomes in schools that did not receive the intervention), baseline equivalence can be established:

- *for causal inference at the individual level* if baseline equivalence is established on the analytic sample of individuals prior to the intervention.

---

[62] For example, suppose an intervention reduces the dropout rate, and students who drop out are excluded from the analytic sample. The effects could lead to a lower attrition rate in the treatment group than in the control group, lower attrition bias in the treatment group mean than in the control group mean, and biased estimates of the treatment effect for outcomes like graduation, grade promotion, and credit accumulation.

[63] The reviewer defines the point of random assignment as the point at which the results of random assignment (treatment or control) are communicated to parties that could be influenced by the results (e.g., schools if the impact study randomized schools or teachers).

[64] Joiners are defined as individuals who are not in the sample before the point of random assignment but are in the sample at the point of outcome measurement—that is, they were included in the analytic sample.

- *for causal inference at the cluster level* if baseline equivalence is established on the analytic sample of clusters prior to the intervention (i.e., on the same cohort of individuals from an earlier point in time or an earlier, adjacent cohort in the same grade as the cohort used in the impact analysis[65]).

To meet the baseline equivalence criteria, the impact study must both affirmatively demonstrate baseline equivalence between the two groups and not provide evidence of nonequivalence. To affirmatively demonstrate baseline equivalence, the impact study must show that the two groups are equivalent at baseline on a measure that is highly correlated with the outcome measure.

Two types of pre-intervention measures are *always* considered highly correlated with a given outcome measure:

1. **A pre-intervention measure of the outcome measure.** To satisfy this condition, the baseline measure must be collected using the same testing instrument or protocol as the outcome measure (e.g., the Stanford 10 at both time points).
2. **A pre-intervention measure in the same evaluator-defined outcome domain or, where applicable, the same domain as defined for i3.** The outcome domains as defined for i3 are described in Tables 1 and 2. Briefly, there are seven student outcome domains consistent with the specific goals of i3. These "relevant outcomes" are: English language arts, mathematics, science, social studies, general achievement, academic performance in school, and educational progress and attainment. Additionally, there are also three "other" student outcome domains: student achievement in other subjects (beyond the four core academic subjects), labor market outcomes and social-emotional skills/development. For example, if the evaluator separates reading outcomes into the fluency and comprehension outcome domains, the reviewer will accept a fluency baseline measure for a comprehension outcome because both fall into the same outcome domain of English language arts defined for i3.

Additionally, for causal inference at the individual level, the reviewer considers additional pre-intervention measures to be highly correlated with the outcome measure in specific contexts and for particular outcome measures.[66]

- If the outcome does not have a natural pre-test (e.g., graduation) and the intervention is provided to students at the K-12 or postsecondary level, the following combination of measures are considered highly correlated with the outcome: a) race/ethnicity or a pre-intervention measure of disadvantage and b) a pre-intervention measure of academic performance or achievement. Acceptable measures of disadvantage include eligibility for free or reduced price lunch, immigrant status, parent's education,

---

65 The earlier adjacent cohort must be from the year that immediately precedes the implementation of the intervention—that is, the most recent cohort of students that did not receive the intervention in either of the two groups.

66 For causal inference at the individual level, the additional options for acceptable baseline measures account for the challenges in obtaining a pretest measure in the same domain as the outcome. However, for causal inference at the cluster level, these challenges are not present because it is usually possible to establish baseline equivalence at the cluster level for a pre-intervention measure in the same domain as the outcome (e.g., cluster level outcomes at the same grade level as the intervention but in the prior year). Therefore, additional options for establishing baseline equivalence for causal inference at the cluster level are not necessary.

single-parent household, special education, and teen parent status. Acceptable measures of academic performance or achievement include standardized test scores, credits earned, at or behind grade level, rate of school attendance, and grade point average.

- If the outcome is a measure of student achievement in an outcome domain that is not tested in every grade level (e.g., science or social studies), and the intervention is provided to students at the K-12 level, either of the following two types of measures are considered highly correlated with the outcome: (1) a pre-intervention measure in the mathematics outcome domain defined for i3 or (2) a pre-intervention measure in the English language arts outcome domain defined for i3.
- If the outcome is a measure of children's reading and literacy or mathematics and the intervention is provided to children below the age of 3 years 0 months, either of the following two types of measures are considered highly correlated with the outcome: (1) a pre-intervention measure in the general achievement outcome domain defined for i3 or (2) a pre-intervention measure in the English Language Arts outcome domain defined for i3.
- *If the intervention is provided to children younger than 18 months*, a pre-intervention measure of parental resources (e.g., maternal education, family income, and single-parent household status) or a standardized test of maternal intelligence (e.g., Kaufman Brief Intelligence test) is considered highly correlated with child outcomes.

Even if the two groups are equivalent on a pre-intervention measure that is highly correlated with the outcome measure, baseline equivalence will be rejected if the evaluator provides evidence of nonequivalence for any measure in the same evaluator-defined domain or in the same domain as defined for i3 as the outcome measure. Baseline equivalence will also be rejected under either of the following two scenarios:

1. The analysis plan indicates that the study is targeted toward a special population (e.g., English learners, students with behavioral problems, or students in special education), evidence from the baseline or analytic samples suggest that over 50 percent of the sample belongs to the special population, and the evaluator provides evidence of nonequivalence between the two groups in the analytic sample on either (1) the fraction of students that belong to the special population (e.g., has a behavioral problem) or (2) the mean of a continuous measure of the condition by which the special population is defined (e.g., scale scores from an assessment of behavioral problems).
2. The intervention is provided to children younger than 18 months, and the evaluator provides evidence of nonequivalence between the two groups on any measure of developmental delay or disability (e.g., a developmental screening instrument of cognitive development or social-emotional development).

Exhibit B.1: Relevant Student Outcome Domains Defined for i3

- **English language arts.** Outcomes related to English language arts measure the development of verbal, written, and other forms of communication skills or content knowledge. English language arts includes, but is not limited to, alphabetics, phonological processing, reading fluency, English language development, writing, print knowledge, oral language, comprehension (including reading comprehension, vocabulary, and listening comprehension), spelling, knowledge of features of the English language (e.g., grammar, syntax), general reading achievement and general literacy achievement.
- **Mathematics.** Outcomes related to mathematics measure mathematics content knowledge and skills, which demonstrate understanding of and application of mathematical concepts, procedures, and problem solving. Mathematics content includes, but is not limited to, knowledge of numbers, arithmetic, measurement, graphing, logical reasoning, pre-algebra, algebra, geometry, trigonometry, pre-calculus, calculus, and general mathematics achievement.
- **Science.** Outcomes related to science measure mastery of science concepts, inquiry, practices, procedures, and problem solving, such as reasoning and proof, making connections, and use of scientific representation. Science content includes, but is not limited to, life science (e.g., biology), Earth/space science (e.g., geology, astronomy), physical science (e.g., chemistry, physics), and general science achievement.
- **Social studies.** Outcomes related to social studies measure content knowledge of relationships and the functioning of society as well as content application. Social studies content includes, but is not limited to, civics, government, geography, history, economics, and general social studies achievement.
- **General achievement.** Either (A) a combination of two or more of the above domains related to academic achievement in a given subject area, (B) general academic achievement that is not content-specific (including standardized achievement tests like the ACT and SAT as well as state-mandated tests that are not content-specific), or (C) cognition, which includes, but is not limited to, memory, problem-solving, cognitive processing and flexibility, general knowledge, and IQ. Measures of general kindergarten readiness are considered cognition measures.
- **Academic performance in school.** Outcomes related to the extent to which students adequately complete expected coursework. As such, eligible measures of academic performance in schools are those that arise naturally from student educational experiences. Examples of ways that academic performance in school might be operationally defined in studies include: (a) final grade in a single course, (b) grade point average across courses, and (c) the ratio of courses passed vs. failed.
- **Educational progress and attainment.** Outcomes related to staying in school, progressing in schools, completing school, and postsecondary access and enrollment (from WWC's Transition to College protocol). (A) Staying in school refers to outcomes that measure whether the student has dropped out of school and the number of days the student was enrolled in school. (B) Progressing in school refers to outcomes that assess the number of high school course credits the student has earned, whether the student was promoted to the next grade, and the highest grade the student has completed. (C) Completing school refers to outcomes that measure whether the student has earned a high school diploma or GED or whether he or she has graduated from high school. (D) Postsecondary access and enrollment refers to the process of applying to, actually enrolling, and attending a postsecondary institution. Examples of ways that enrollment might be operationally defined in studies include: (a) actual enrollment in college, (b) number and/or selectivity of admitted and/or enrolling institutions, (c) enrollment by institution type (2 year vs. 4 year), (d) intensity of enrollment (full time vs. part time), (e) timing of enrollment (e.g., immediate vs. delayed enrollment after high school). (f) the completion of financial aid forms geared at college attendance, and (g) enrollment in college preparatory courses (e.g., Advanced Placement, dual enrollment, or International Baccalaureate).

Exhibit B.2: Other Student Outcome Domains Defined for i3

- **Achievement in Other (Non-Core Academic) Subjects.** Outcomes related to content knowledge and skills outside of the four core academic subjects (Mathematics, English Language Arts, Science, and Social Studies). Other subjects include, but are not limited to arts, drama, home economics, ancient or modern languages, music, physical development or education, and technology.
- **Labor Market Outcomes.** Outcomes related to measuring student employment status (e.g., employed vs. unemployed; employed full-time vs. employed part-time; employed in field vs. not employed in field); includes outcomes related to vocational skill development (e.g., training in a trade via collaborative education or career academies).
- **Social-emotional skills/development.** Outcomes related to social-emotional development measure behavioral, social, and emotional competencies. Social-emotional development includes, but is not limited to, measures of pro-social (or problem) behaviors, social interactions, cooperation, self-concept, engagement, attention, persistence, impulsivity, self-control, and initiative.

In general, to compute the baseline difference between the two groups, the reviewer uses the unweighted mean and standard deviation for the treatment group and the unweighted mean and standard deviation for the comparison group. However, when studies use weights or other analytic methods to account for the research design or to address nonresponse, the reviewer uses the same method to test for baseline equivalence. For example, if an evaluator of an RCT with high attrition uses weights for the impact analysis to account for differential rates of assignment to treatment within blocks, the reviewer uses weighted means and standard deviations in its assessment of baseline equivalence. If the study takes a modeling approach to account for design features of the study (e.g., the model includes fixed effects for the blocks within which schools were randomly assigned or selected and random effects for clusters of students selected into the study sample), the reviewer uses the following two pieces of information to assess baseline equivalence: (1) the unweighted standard deviations for both groups and (2) a "design-adjusted" difference in means between the two groups (i.e., the regression coefficient on the treatment indicator from a model that regresses the baseline pre-test variable on the treatment indicator and whatever fixed and/or random effects were included in the analysis model to account for study design features).

Lastly, the reviewer reserves the discretion to conclude that the study cannot establish baseline equivalence if the intervention and comparison groups were drawn from substantially different contexts, as reported by the evaluator (e.g., urban schools versus rural schools, schools in corrective action versus schools that consistently make Adequate Yearly Progress).

For more details on the WWC Standards for baseline equivalence, see the *WWC Procedures and Standards Handbook, version 3.0*, Section III: Subsection B.3, "Baseline equivalence: Is equivalence established at baseline for the groups in the analytic sample?" (pp. 15-16).

Outcomes

The reviewer assesses outcomes for face validity, reliability, consistency of data collection, and (lack of) overalignment. To satisfy the outcome standards for face validity, an outcome must be clearly defined, have a direct interpretation, and measure the outcome domain that it was intended to measure. To be clearly defined, the evaluator must provide as much detail about the outcome as might typically be expected in a journal article to provide readers with a basic understanding of how it was constructed. To

have a direct interpretation and measure the outcome domain it was designed to measure, the outcome must appear to measure the domain into which it was classified by the evaluator.

The reviewer treats standard educational measures (e.g., grades, credits earned, grade promotion or retention, graduation or dropout, and postsecondary attendance, persistence or completion) as face valid for the domains they most clearly measure. For example, grades are considered to be face valid measures of academic performance in school (but not of student achievement), and both credits earned and promotion to the next grade level are considered face valid measures of educational progress and attainment.

The minimum standards for the reliability of the outcome measure are:

- Internal consistency (such as Cronbach's alpha) of 0.60 or higher.
- Temporal stability/test-retest reliability of 0.40 or higher.
- Inter-rater reliability (such as percentage agreement, correlation, or kappa) of 0.50 or higher.

Each outcome measure must satisfy one of these three criteria to meet the outcome standards. The reviewer accepts both evaluator-reported reliability statistics and references cited in the literature, assumes that standardized tests (e.g., the Terra Nova or Stanford 10) and state-required tests satisfy the reliability criteria, and exempts standard educational measures from reliability criteria since reliability cannot be estimated for these measures.

In addition, to meet the outcome standards, outcomes must be collected in the same manner between treatment and comparison groups and not be overaligned with the intervention.

For more details on the outcome standards, see the *WWC Procedures and Standards Handbook, version 3.0*, Section III: Subsection B.4, "Outcome Eligibility and Reporting" (pp. 16-19).

Confounding Factors

The reviewer assesses whether the design used to produce a particular estimated impact suffers from a fundamental confound. For information on the criteria for identifying confounds, see the *WWC Procedures and Standards Handbook, Version 3.0*, Section III: Subsection B.5, "Confounding Factors" (pp. 19-20).

In principal, a design that estimates the effects of a bundle of interventions could be viewed as either free of confounds for the intervention bundle or confounded for each particular intervention in the bundle.

- *If the evaluation tests a bundle of i3-funded interventions*, the reviewer treats the estimated impacts as unconfounded for the effects of the bundle (i.e., the design does not suffer from a confound).
- If the evaluation tests a bundle that includes an i3-funded intervention and a non-i3-funded intervention, the reviewer treats the estimated impacts as:
  - unconfounded if the non-i3-funded intervention is offered by the developer of the i3-funded intervention (i.e., the design does not suffer from a confound).
  - confounded if the non-i3-funded intervention is offered by a different entity from the developer of the i3-funded intervention (i.e., the design suffers from a confound).

Statistical Adjustments

The reviewer makes clustering corrections when testing for statistical significance for analyses that have not adequately accounted for clustering (see the *WWC Procedures and Standards Handbook, version 3.0,* Section IV: Subsection B, "Statistical Significance of Findings," pp. 24–26). Like the WWC, the reviewer uses the Benjamini-Hochberg procedure to correct for multiple comparisons.

Correcting for Multiple Comparisons in Reporting Evidence of Effectiveness from i3 Evaluations

Most of the i3 local evaluations plan to conduct more than one hypothesis test to determine whether the intervention was effective. Conducting multiple hypothesis tests can lead to a high "false discovery rate," which is the proportion of estimated effects that are large enough to reject the null hypothesis when the true impact is zero. In addition, it can lead to a high "family-wise error rate," which refers to the fraction of random samples for which evaluators would falsely conclude, purely by chance, that the intervention has a positive effect on outcome variables in the same domain, which is defined as follows:

- **Outcome domain.** The outcome domain indicates, at a general level, the outcome that may be affected by the intervention; it can be thought of as a latent construct that can be measured with one or more outcome measures. For the purpose of correcting for multiple comparisons, the reviewer uses the relevant student outcome domains defined for i3 (see Table 1).

The reviewer makes adjustments for multiple comparisons that address outcomes in the same domain. To correct study estimated impacts from prespecified analyses for multiple comparisons, the reviewer uses the Benjamini-Hochberg procedure as described in the *WWC Procedures and Standards Handbook, version 3.0,* Section IV (pp. 25-26). The reviewer follows the recommendations of an IES methods report and does not adjust for multiple comparisons in analyses that were not prespecified.[67]

[67] Schochet, P. Z. (2008). *Technical methods report: Guidelines for multiple testing in impact evaluations* (NCEE 2008-4018). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Appendix C: Assessments of and Findings from 67 i3 Impact and Implementation Evaluations

Intervention	Evaluators prespecified and reported results of confirmatory impact analyses	Strength of the Impact Evaluation	Quality implementation evaluation	Impact evaluation sample adequately representative of population served	Impact evaluation is considered independent	Evidence that the intervention was implemented with adequate fidelity	Evidence of at least one positive, statistically significant impact on student outcomes
Scale-up							
KIPP	Yes[a]	Unofficially Meets WWC Standards with Reservations	Yes	No	Yes	Yes	Yes
Reading Recovery	Yes[b]	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	Yes
Success for All (SFA)	Yes	Unofficially Meets WWC Standards without Reservations	Yes	No	Yes	Yes	No (null)
Teach for America (TFA)	Yes	Unofficially Meets WWC Standards without Reservations	Yes	No	Yes	Yes	No (null)
Validation							
Advanced ASSET Professional Development	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	Yes	No (null)
The Baby Family and Child Education (FACE) Program	Yes	Unofficially Does Not Meet WWC Standards	Yes	Yes	Yes	Yes	No evidence (Unofficially DNM WWC)
Children's Literacy Initiative (CLI) Program	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	Yes
Collaborative Strategic Reading (CSR)	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	No (null)
College Readiness Program (CRP)	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	No	Yes
College-Ready Writers Program (CRWP)	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	Yes

Intervention	Evaluators prespecified and reported results of confirmatory impact analyses	Strength of the Impact Evaluation	Quality implementation evaluation	Impact evaluation sample adequately representative of population served	Impact evaluation is considered independent	Evidence that the intervention was implemented with adequate fidelity	Evidence of at least one positive, statistically significant impact on student outcomes
Diplomas Now	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	No (null)
enhancing Missouri's Instructional Networked Teaching Strategies (eMINTS)	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	Yes
GO College (An Enhanced Version of Talent Search)	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	No (null)
Leadership Assistance for Science Education Reform (LASER)	Yes	Unofficially Meets WWC Standards without Reservations	Yes	No	Yes	Yes	No (null)
Northeast Tennessee College and Career Ready Consortium	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	No	Yes
Reading Apprenticeship (RA)	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	No (null)
Reading Enhances Achievement During the Summer (READS)	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	No (null)
StartSmart K-3 Plus Program	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	Yes
Virginia Initiative for Science Teaching and Achievement (VISTA)	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	Yes	No (null)

Intervention	Evaluators prespecified and reported results of confirmatory impact analyses	Strength of the Impact Evaluation	Quality implementation evaluation	Impact evaluation sample adequately representative of population served	Impact evaluation is considered independent	Evidence that the intervention was implemented with adequate fidelity	Evidence of at least one positive, statistically significant impact on student outcomes
Development							
The Achievement Network (ANet)	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	No	No (null)
Advanced Placement (AP) Insight	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	Yes	No (null)
Advancement via Individual Determination (AVID)	Yes[c]	Unofficially Does Not Meet WWC Standards	Yes	Yes	Yes	Yes	No evidence (Unofficially DNM WWC)
Around the Corner (ATC)	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	No (null)
Arts Achieve: Impacting Student Success in the Arts	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	No	No (null)
Arts for Learning (A4L) Lessons	Yes[d]	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	No (null)
Bay State Reading Institute (BSRI)	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	Yes	No (null)
Boston Teacher Residency (BTR)	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	Yes	No (null)
Building Assets, Reducing Risks (BARR) Model	Yes	Unofficially Meets WWC Standards without Reservations	Yes	No	Yes	Yes	Yes
Collaborative Organizational Model to Promote Aligned Support Structures (COMPASS)	Yes	Unofficially Does Not Meet WWC Standards	Yes	No	Yes	Yes	No evidence (Unofficially DNM WWC)
CollegeYES	Yes	Unofficially Does Not Meet WWC Standards	Yes	No	Yes	Yes	No evidence (Unofficially DNM WWC)

Intervention	Evaluators prespecified and reported results of confirmatory impact analyses	Strength of the Impact Evaluation	Quality implementation evaluation	Impact evaluation sample adequately representative of population served	Impact evaluation is considered independent	Evidence that the intervention was implemented with adequate fidelity	Evidence of at least one positive, statistically significant impact on student outcomes
Convergence Academies	Yes	Unofficially Ineligible for WWC Review	Yes	Yes	Yes	Yes	No evidence (Unofficially Ineligible for WWC)
Curriculum 2.0	Yes	Unofficially Ineligible for WWC Review	No	Yes	Yes	No Evidence	No evidence (Unofficially Ineligible for WWC)
Data-Driven Decision Making and Information Technology Curricula in Schools	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	Yes	No (negative)
E3TL Teacher Performance Evaluation System	Yes	Unofficially Ineligible for WWC Review	Yes	Yes	Yes	Yes	No evidence (Unofficially Ineligible for WWC)
EngageMe P.L.E.A.S.E.	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	No (null)
Every Child Ready (ECR)	Yes	Unofficially Does Not Meet WWC Standards	Yes	Yes	No	Yes	No evidence (Unofficially DNM WWC)
Everyday Arts for Special Education (EASE)	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	Yes	No (null)
Exceptional Coaching for Early Language and Literacy - Enhanced (ExCELL-E)	Yes[e]	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	Yes	No (null)
The Expository Reading and Writing Course (ERWC)	Yes	Unofficially Meets WWC Standards with Reservations	Yes	No	Yes	Yes	Yes
Facilitating Long-Term Improvements in Graduation and Higher Education for Tomorrow (FLIGHT)	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	Yes	No (null)
Florida Master Teacher Initiative (FMTI)	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	No	No (null)

Intervention	Evaluators prespecified and reported results of confirmatory impact analyses	Strength of the Impact Evaluation	Quality implementation evaluation	Impact evaluation sample adequately representative of population served	Impact evaluation is considered independent	Evidence that the intervention was implemented with adequate fidelity	Evidence of at least one positive, statistically significant impact on student outcomes
InnovateNYC Ecosystem	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	No	No (null)
Innovations in Early Mathematics	Yes	Unofficially Meets WWC Standards with Reservations	Yes	No	Yes	Yes	No (null)
Integrating English Language Development (ELD) and Science	Yes	Unofficially Ineligible for WWC Review	Yes	Yes	Yes	Yes	No evidence (Unofficially Ineligible for WWC)
Internet-based Reading Apprenticeship Improving Science Education (iRAISE)	Yes	Unofficially Does Not Meet WWC Standards	Yes	Yes	Yes	Yes	No evidence (Unofficially DNM WWC)
Making Time for What Matters Most	Yes	Unofficially Ineligible for WWC Review	Yes	Yes	Yes	No	No evidence (Unofficially Ineligible for WWC)
New England Network for Personalization and Performance (NETWORK)	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	Yes	No (null)
Oakland Accelerates	Yes	Unofficially Ineligible for WWC Review	Yes	Yes	Yes	Yes	No evidence (Unofficially Ineligible for WWC)
Ounce of Prevention Fund Professional Development Initiative (PDI)	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	Yes	No (null)
Pathways to STEM Initiative (PSI)	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	No	No (null)
Problem-Based Learning (PBL)	No	No prespecified analyses[f]	Yes	No Evidence	No Evidence	Yes	No evidence (No prespecified analyses)
PTA Comunitario	Yes	Unofficially Meets WWC Standards with Reservations	Yes	No	Yes	Yes	No (null)
Public School Choice Initiative (PSCI)	Yes	Unofficially Does Not Meet WWC Standards	Yes	No	Yes	Yes	No evidence (Unofficially DNM WWC)

Intervention	Evaluators prespecified and reported results of confirmatory impact analyses	Strength of the Impact Evaluation	Quality implementation evaluation	Impact evaluation sample adequately representative of population served	Impact evaluation is considered independent	Evidence that the intervention was implemented with adequate fidelity	Evidence of at least one positive, statistically significant impact on student outcomes
Responding Effectively to Assessments with Curriculum and Teaching (REACT)	Yes	Unofficially Does Not Meet WWC Standards	Yes	No	Yes	Yes	No evidence (Unofficially DNM WWC)
Rio Grande Valley Center for Teaching and Leading Excellence: New Teacher Training (NTT)	Yes[g]	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	No (null)
Rural Math Excellence Partnership (RMEP)	Yes	Unofficially Meets WWC Standards with Reservations	No	Yes	Yes	No Evidence	No (null)
School of One	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes[h]	Yes	No (null)
Schools to Watch (STW): School Transformation Network Project	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	Yes	No (null)
Science, Technology, Engineering and Math Education for the 21st Century (STEM21)	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	Yes	No (null)
Spheres of Proud Achievement in Reading for Kids (SPARK) Program	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	Yes	Yes
STEM Learning Opportunities Providing Equity (SLOPE)	Yes	Unofficially Meets WWC Standards without Reservations	Yes	Yes	Yes	No	No (null)
STEM Summer Learning with VEX Robotics	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	No	No (null)
Texas Tech University "Tech Teach" Program	Yes	Unofficially Does Not Meet WWC Standards	Yes	Yes	Yes	Yes	No evidence (Unofficially DNM WWC)
Transforming Teacher Talent (t3) System	Yes	Unofficially Ineligible for WWC Review	Yes	Yes	Yes	No	No evidence (Unofficially Ineligible for WWC)
Turnaround with Increased Learning Time (TILT)	Yes	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	No	Yes

Intervention	Evaluators prespecified and reported results of confirmatory impact analyses	Strength of the Impact Evaluation	Quality implementation evaluation	Impact evaluation sample adequately representative of population served	Impact evaluation is considered independent	Evidence that the intervention was implemented with adequate fidelity	Evidence of at least one positive, statistically significant impact on student outcomes
We Are A Village	Yes[i]	Unofficially Ineligible for WWC Review	Yes	No	Yes	Yes	No evidence (Unofficially Ineligible for WWC)
Write Up!	Yes[j]	Unofficially Meets WWC Standards with Reservations	Yes	Yes	Yes	No	No (null)
Total number of evaluations meeting criteria (N=67 evaluations)	66 (99%)	49 (73%)	65 (97%)	53 (79%)	65 (97%)	52 (78%)	12 (18%)

Notes:
[a] KIPP reported 16 of 18 prespecified analyses.
[b] Evaluation findings included here are from the RCT design study only and do not include findings from an RD study that was also conducted by the evaluator.
[c] AVID reported 3 of 4 prespecified analyses.
[d] A4L reported 1 of 2 prespecified analyses.
[e] ExCELL-E reported 4 of 5 prespecified analyses.
[f] Our review of the findings reported by the evaluation of PBL concludes that they are considered independent, are Unofficially Ineligible for WWC Review, and included a sample that is adequately representative of those served by i3.
[g] Rio Grande Valley Center for Teaching and Leading Excellence: New Teacher Training (NTT) reported 5 of 6 prespecified analyses.
[h] School of One reported 2 analyses; 1 of these 2 was considered independent.
[i] We Are A Village reported 1 of 3 prespecified analyses.
[j] Write Up! reported 3 of 4 prespecified analyses.

Appendix D: Findings Reported by 67 i3 Impact Evaluations from Prespecified Analyses of Student Academic Outcomes

This appendix presents the findings from analyses that were found to Unofficially Meet WWC Standards with or without Reservations. Each row refers to an impact of an intervention versus the comparison condition for a particular outcome domain and educational level. The comparison condition is often denoted as *Business as Usual* unless more details were needed to distinguish between different impacts from the same i3 evaluation. *Business as Usual* reflects whatever educational inputs were in place in the control or counterfactual environment and therefore varies across locations and evaluations. The effect size column reports the magnitude of the impact in standard deviation units using *Hedges' g* to calculate the standardized effect sizes, as described in the *WWC Procedures and Standards Handbook, version 3.0,* Appendix F (pp. F.1–G.12). Findings that are statistically significant at the 5 percent level are indicated with an asterisk (*).

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Scale-up					
KIPP					
KIPP Admission	No KIPP Admission	English Language Arts	Elementary School	Unofficially Does Not Meet WWC Standards	–
KIPP Admission	No KIPP Admission	English Language Arts	Elementary School	Unofficially Does Not Meet WWC Standards	–
New KIPP Schools	No KIPP	English Language Arts	Middle School	Unofficially Meets WWC Standards with Reservations	+0.120*
KIPP Admission	No KIPP Admission	English Language Arts	Middle School	Unofficially Meets WWC Standards without Reservations	+0.143*
KIPP Charter Schools	Business as Usual	English Language Arts	Middle and High School	Unofficially Meets WWC Standards with Reservations	+0.295*
Non-KIPP Charter Led by a KIPP-Trained Principal	No KIPP and No Non-KIPP Charter Led by a KIPP-Trained Principal	English Language Arts	Middle and High School	Not Reported	–
KIPP Charter Schools	Business as Usual	English Language Arts	High School	Unofficially Meets WWC Standards with Reservations	+0.175*
Option of Continuing KIPP	No Option of Continuing KIPP	English Language Arts	High School	Unofficially Meets WWC Standards with Reservations	+0.160*
Option of Continuing KIPP	No Option of Continuing KIPP	English Language Arts	High School	Unofficially Does Not Meet WWC Standards	–
KIPP Admission	No KIPP Admission	Mathematics	Elementary School	Unofficially Does Not Meet WWC Standards	–
KIPP Admission	No KIPP Admission	Mathematics	Elementary School	Unofficially Does Not Meet WWC Standards	–
New KIPP Schools	No KIPP	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	+0.226*

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
KIPP Admission	No KIPP Admission	Mathematics	Middle School	Unofficially Meets WWC Standards without Reservations	+0.176*
KIPP Charter Schools	Business as Usual	Mathematics	Middle and High School	Unofficially Meets WWC Standards with Reservations	+0.341*
Non-KIPP Charter Led by a KIPP-Trained Principal	No KIPP and No Non-KIPP Charter Led by a KIPP-Trained Principal	Mathematics	Middle and High School	Not Reported	–
KIPP Charter Schools	Business as Usual	Mathematics	High School	Unofficially Meets WWC Standards with Reservations	+0.273*
Option of Continuing KIPP	No Option of Continuing KIPP	Mathematics	High School	Unofficially Meets WWC Standards with Reservations	+0.136
Option of Continuing KIPP	No Option of Continuing KIPP	Mathematics	High School	Unofficially Does Not Meet WWC Standards	–
Reading Recovery					
Reading Recovery[a]	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.503*
Success for All (SFA)					
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.159
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.075
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.075
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.033
Teach for America (TFA)					
TFA	Business as Usual	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.028
TFA	Business as Usual	Mathematics	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.051[68]

[68] The update to the report would not have changed our overall assessment of the findings.

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Validation					
Advanced ASSET Professional Development					
Advanced ASSET Professional Development after Prior Experience with Science: It's Elementary	No Current Advanced ASSET Professional Development after Prior Experience with Science: It's Elementary	Science	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.334
Advanced ASSET Professional Development after Experience with Science: It's Elementary	No Current or Prior State-wide Science Curriculum or Training Initiatives	Science	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.421
The Baby Family and Child Education (FACE) Program					
Baby FACE Program	Business as Usual	General Achievement	Early Childhood	Unofficially Does Not Meet WWC Standards	–
Baby FACE Program	Business as Usual	General Achievement	Early Childhood	Unofficially Does Not Meet WWC Standards	–
Children's Literacy Initiative (CLI) Program					
CLI	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.207*
CLI	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.138
CLI	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.047
CLI	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.008
CLI	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.014
Collaborative Strategic Reading (CSR)					
CSR	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards without Reservations	+0.073
CSR	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards without Reservations	+0.039

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
CSR	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards without Reservations	+0.013
CSR - School Wide	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards with Reservations	+0.025
CSR - School Wide	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards with Reservations	-0.056
CSR - School Wide	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards with Reservations	-0.090
College Readiness Program (CRP)					
CRP[b]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	+0.904*
CRP[b]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	+0.660*
College-Ready Writers Program (CRWP)					
CRWP	Business as Usual	English Language Arts	Middle and High School	Unofficially Meets WWC Standards without Reservations	+0.201*
Diplomas Now					
Diplomas Now	Business as Usual	Educational progress and attainment	Middle and High School	Unofficially Meets WWC Standards without Reservations	+0.054
Diplomas Now	Business as Usual	Educational progress and attainment	Middle and High School	Unofficially Meets WWC Standards without Reservations	+0.026
Diplomas Now	Business as Usual	Educational progress and attainment	Middle and High School	Unofficially Meets WWC Standards without Reservations	+0.003
enhancing Missouri's Instructional Networked Teaching Strategies (eMINTS)					
eMINTS	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards without Reservations	-0.035
eMINTS	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards without Reservations	-0.042
eMINTS with Intel Teach	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards without Reservations	-0.075
eMINTS	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards without Reservations	+0.130
eMINTS	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards without Reservations	+0.048

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
eMINTS with Intel Teach	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards without Reservations	+0.168*
GO College (An Enhanced Version of Talent Search)					
GO College Whole School Program with Learning Community	GO College Whole School Program	English Language Arts	High School	Unofficially Meets WWC Standards without Reservations	+0.027
GO College Whole School Program with Learning Community	GO College Whole School Program	Mathematics	High School	Unofficially Meets WWC Standards without Reservations	+0.040
GO College Whole School Program with Learning Community	GO College Whole School Program	Academic performance in school	High School	Unofficially Meets WWC Standards without Reservations	+0.045
GO College Whole School Program with Learning Community	GO College Whole School Program	Educational progress and attainment	High School	Unofficially Meets WWC Standards with Reservations	-0.006
Leadership Assistance for Science Education Reform (LASER)					
LASER	Business as Usual	Science	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.002
LASER	Business as Usual	Science	Middle School	Unofficially Meets WWC Standards without Reservations	-0.049
Northeast Tennessee College and Career Ready Consortium					
Consortium[c]	Business as Usual	Educational progress and attainment	High School	Unofficially Meets WWC Standards with Reservations	+0.084*
Consortium[c]	Business as Usual	Educational progress and attainment	High School	Unofficially Meets WWC Standards with Reservations	+0.079*
Consortium[c]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	+0.568*
Consortium[c]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	+0.049*
Reading Apprenticeship (RA)					
Reading Apprenticeship	Business as Usual	English Language Arts	High School	Unofficially Meets WWC Standards without Reservations	+0.147
Reading Apprenticeship	Business as Usual	English Language Arts	High School	Unofficially Meets WWC Standards without Reservations	+0.139
Reading Enhances Achievement During the Summer (READS)					
READS	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.008

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
READS	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.004
READS - Parental Involvement (READS - PI)	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.044
READS - Parental Involvement plus Followup (READS - PIF)	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.046
StartSmart K-3 Plus Program					
Extended School Year	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.150*
Extended School Year	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.113*
Extended School Year	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.030
Extended School Year	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.003
Extended School Year	Business as Usual	Mathematics	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.092*
Virginia Initiative for Science Teaching and Achievement (VISTA)					
VISTA Teacher Professional Development	Business as Usual	Science	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.068
VISTA Teacher Professional Development	Business as Usual	Science	Middle School	Unofficially Meets WWC Standards with Reservations	-0.103
VISTA Teacher Professional Development	Business as Usual	Science	Middle School	Unofficially Meets WWC Standards with Reservations	-0.169
VISTA Teacher Professional Development	Business as Usual	Science	High School	Unofficially Does Not Meet WWC Standards	–
VISTA Teacher Professional Development	Business as Usual	Science	High School	Unofficially Does Not Meet WWC Standards	–

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Development					
The Achievement Network (ANet)					
ANet[d]	Business as Usual	English Language Arts	Elementary and Middle School	Unofficially Meets WWC Standards with Reservations	-0.047
ANet[d]	Business as Usual	Mathematics	Elementary and Middle School	Unofficially Meets WWC Standards with Reservations	-0.040
Advanced Placement (AP) Insight					
AP Insight	Business as Usual	Science	High School	Unofficially Meets WWC Standards with Reservations	+0.054
AP Insight	Business as Usual	Science	High School	Unofficially Meets WWC Standards with Reservations	+0.047
AP Insight	Business as Usual	Science	High School	Unofficially Meets WWC Standards with Reservations	+0.009
AP Insight	Business as Usual	Science	High School	Unofficially Meets WWC Standards with Reservations	-0.004
Advancement via Individual Determination (AVID)					
AVID-core-plus schoolwide intervention	Business as Usual	English Language Arts	Middle and High School	Unofficially Does Not Meet WWC Standards	–
AVID-core-plus schoolwide intervention	Business as Usual	Mathematics	Middle and High School	Unofficially Does Not Meet WWC Standards	–
AVID-core-plus schoolwide intervention	Business as Usual	Educational progress and attainment	Middle and High School	Not Reported	–
AVID-core-plus schoolwide intervention	Business as Usual	Educational progress and attainment	High School	Unofficially Does Not Meet WWC Standards	–
Around the Corner (ATC)					
Around the Corner	Existing SFA Program	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.099
Arts Achieve: Impacting Student Success in the Arts					
Arts Achieve[e]	Business as Usual	English Language Arts	Elementary and Middle School	Unofficially Meets WWC Standards with Reservations	+0.039
Arts Achieve[e]	Business as Usual	English Language Arts	High School	Unofficially Meets WWC Standards with Reservations	+0.013
Arts for Learning (A4L) Lessons					
A4L	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.047
A4L	Business as Usual	English Language Arts	Elementary School	Not Reported	–

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Bay State Reading Institute (BSRI)					
Bay State Reading Institute	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.095
Boston Teacher Residency (BTR)					
Newly Hired BTR Graduates	Business as Usual	English Language Arts	Elementary and Middle School	Unofficially Meets WWC Standards with Reservations	-0.093
Newly Hired BTR Graduates	Business as Usual	Mathematics	Elementary and Middle School	Unofficially Meets WWC Standards with Reservations	+0.009
Building Assets, Reducing Risks (BARR) Model					
BARR	Business as Usual	Educational progress and attainment	High School	Unofficially Meets WWC Standards without Reservations	+0.227*
Collaborative Organizational Model to Promote Aligned Support Structures (COMPASS)					
COMPASS	Business as Usual	English Language Arts	Elementary and Middle School	Unofficially Does Not Meet WWC Standards	–
CollegeYES					
CollegeYES	Business as Usual	Science	High School	Unofficially Does Not Meet WWC Standards	–
Convergence Academies (CA)					
CA	Business as Usual	English Language Arts	Elementary and Middle School	Unofficially Ineligible for WWC Review	–
CA	Business as Usual	English Language Arts	High School	Unofficially Ineligible for WWC Review	–
CA	Business as Usual	English Language Arts	High School	Unofficially Ineligible for WWC Review	–
CA	Business as Usual	Mathematics	Elementary and Middle School	Unofficially Ineligible for WWC Review	–
CA	Business as Usual	Mathematics	High School	Unofficially Ineligible for WWC Review	–
CA	Business as Usual	Mathematics	High School	Unofficially Ineligible for WWC Review	–
CA	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
CA	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
CA	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Curriculum 2.0					
Curriculum 2.0	Business as Usual	English Language Arts	Elementary School	Unofficially Ineligible for WWC Review	–
Curriculum 2.0	Business as Usual	English Language Arts	Elementary School	Unofficially Ineligible for WWC Review	–
Curriculum 2.0	Business as Usual	English Language Arts	Elementary School	Unofficially Ineligible for WWC Review	–
Curriculum 2.0	Business as Usual	English Language Arts	Elementary School	Unofficially Ineligible for WWC Review	–
Curriculum 2.0	Business as Usual	Mathematics	Elementary School	Unofficially Ineligible for WWC Review	–
Data-Driven Decision Making and Information Technology Curricula in Schools					
Data Driven Decision Making	Business as Usual	English Language Arts	Elementary School	Unofficially Does Not Meet WWC Standards	–
Data Driven Decision Making	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	-0.096*
E3TL Teacher Performance Evaluation System					
E3TL Teacher Performance Evaluation System	Business as Usual	English Language Arts	Elementary and Middle School	Unofficially Ineligible for WWC Review	–
E3TL Teacher Performance Evaluation System	Business as Usual	English Language Arts	Elementary and Middle School	Unofficially Ineligible for WWC Review	–
E3TL Teacher Performance Evaluation System	Business as Usual	Mathematics	Elementary and Middle School	Unofficially Ineligible for WWC Review	–
E3TL Teacher Performance Evaluation System	Business as Usual	Mathematics	Elementary and Middle School	Unofficially Ineligible for WWC Review	–
EngageMe P.L.E.A.S.E.					
EngageMe P.L.E.A.S.E.	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards without Reservations	-0.053
EngageMe P.L.E.A.S.E.	Business as Usual	English Language Arts	High School	Unofficially Meets WWC Standards without Reservations	+0.125
EngageMe P.L.E.A.S.E.	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards without Reservations	-0.010
EngageMe P.L.E.A.S.E.	Business as Usual	Mathematics	High School	Unofficially Meets WWC Standards without Reservations	+0.239

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Every Child Ready (ECR)					
Every Child Ready[f]	Business as Usual	English Language Arts	Early Childhood	Unofficially Does Not Meet WWC Standards	–
Every Child Ready[f]	Business as Usual	English Language Arts	Early Childhood	Unofficially Does Not Meet WWC Standards	–
Every Child Ready[f]	Business as Usual	English Language Arts	Early Childhood	Unofficially Does Not Meet WWC Standards	–
Every Child Ready[f]	Business as Usual	English Language Arts	Early Childhood	Unofficially Does Not Meet WWC Standards	–
Every Child Ready[f]	Business as Usual	Mathematics	Early Childhood	Unofficially Does Not Meet WWC Standards	–
Everyday Arts for Special Education (EASE)					
EASE	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.419
EASE	Business as Usual	Mathematics	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.005
Exceptional Coaching for Early Language and Literacy – Enhanced (ExCELL-E)					
ExCELL-E	Business as Usual	English Language Arts	Early Childhood	Unofficially Meets WWC Standards with Reservations	-0.008
ExCELL-E	Business as Usual	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards with Reservations	+0.065
ExCELL-E	Business as Usual	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards with Reservations	+0.009
ExCELL-E	Business as Usual	English Language Arts	Early Childhood and Elementary School	Unofficially Does Not Meet WWC Standards	–
ExCELL-E	Business as Usual	English Language Arts	Early Childhood and Elementary School	Not Reported	–
The Expository Reading and Writing Course (ERWC)					
ERWC	Business as Usual	English Language Arts	High School	Unofficially Meets WWC Standards with Reservations	+0.133*
Facilitating Long-Term Improvements in Graduation and Higher Education for Tomorrow (FLIGHT)					
FLIGHT	Business as Usual	Academic performance in school	Middle and High school	Unofficially Meets WWC Standards with Reservations	-0.056
Florida Master Teacher Initiative (FMTI)					
FMTI[g]	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.038
FMTI with Early Childhood Master's Program[g]	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.170
FMTI[g]	Business as Usual	Mathematics	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.038

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
FMTI with Early Childhood Master's Program[g]	Business as Usual	Mathematics	Elementary School	Unofficially Meets WWC Standards with Reservations	-0.230
InnovateNYC Ecosystem					
Math Apps[h]	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	+0.075
Innovations in Early Mathematics					
Innovations in Early Mathematics	Business as Usual	Mathematics	Early Childhood and Elementary School	Unofficially Meets WWC Standards with Reservations	+0.013
Innovations in Early Mathematics	Business as Usual	Mathematics	Early Childhood and Elementary School	Unofficially Meets WWC Standards with Reservations	-0.073
Integrating English Language Development (ELD) and Science					
Exploratorium Professional Development	Business as Usual	English Language Arts	Elementary School	Unofficially Ineligible for WWC Review	–
Exploratorium Professional Development	Business as Usual	Science	Elementary School	Unofficially Ineligible for WWC Review	–
Internet-based Reading Apprenticeship Improving Science Education (iRAISE)					
iRAISE	Business as Usual	English Language Arts	High School	Unofficially Does Not Meet WWC Standards	–
Making Time for What Matters Most					
Making Time for What Matters Most[i]	Business as Usual	English Language Arts	High School	Unofficially Ineligible for WWC Review	–
Making Time for What Matters Most[i]	Business as Usual	English Language Arts	High School	Unofficially Ineligible for WWC Review	–
Making Time for What Matters Most[i]	Business as Usual	Mathematics	High School	Unofficially Ineligible for WWC Review	–
Making Time for What Matters Most[i]	Business as Usual	Science	High School	Unofficially Ineligible for WWC Review	–
Making Time for What Matters Most[i]	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
Making Time for What Matters Most[i]	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
Making Time for What Matters Most[i]	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
New England Network for Personalization and Performance (NETWORK)					
NETWORK	Business as Usual	English Language Arts	High School	Unofficially Meets WWC Standards with Reservations	-0.120
NETWORK	Business as Usual	Mathematics	High School	Unofficially Meets WWC Standards with Reservations	+0.062

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
NETWORK	Business as Usual	Educational progress and attainment	High School	Unofficially Meets WWC Standards with Reservations	+0.041
NETWORK	Business as Usual	Educational progress and attainment	High School	Unofficially Meets WWC Standards with Reservations	-0.055
NETWORK	Business as Usual	Educational progress and attainment	High School	Unofficially Meets WWC Standards with Reservations	-0.252
Oakland Accelerates					
Oakland Accelerates	Business as Usual	English Language Arts	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	Mathematics	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	General Achievement	High School	Unofficially Does Not Meet WWC Standards	–
Oakland Accelerates	Business as Usual	Educational progress and attainment	High School	Unofficially Does Not Meet WWC Standards	–
Oakland Accelerates	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
Ounce of Prevention Fund Professional Development Initiative (PDI)					
PDI	Business as Usual	English Language Arts	Early Childhood	Unofficially Meets WWC Standards with Reservations	+0.253
PDI	Business as Usual	English Language Arts	Early Childhood	Unofficially Meets WWC Standards with Reservations	+0.064
PDI	Business as Usual	English Language Arts	Early Childhood	Unofficially Does Not Meet WWC Standards	–
PDI	Business as Usual	Mathematics	Early Childhood	Unofficially Does Not Meet WWC Standards	–
PDI	Business as Usual	General Achievement	Early Childhood	Unofficially Meets WWC Standards with Reservations	+0.198
PDI	Business as Usual	General Achievement	Early Childhood	Unofficially Does Not Meet WWC Standards	–

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Pathways to STEM Initiative (PSI)					
PSI[j]	Business as Usual	Science	Middle School	Unofficially Meets WWC Standards with Reservations	+0.010
PSI[j]	Business as Usual	Science	Middle School	Unofficially Meets WWC Standards with Reservations	-0.056
Problem-Based Learning (PBL)					
				No prespecified analyses[k]	–
PTA Comunitario					
PTA Comunitario	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.025
PTA Comunitario	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards with Reservations	-0.003
PTA Comunitario	Business as Usual	English Language Arts	High School	Unofficially Does Not Meet WWC Standards	–
PTA Comunitario	Business as Usual	Mathematics	Elementary School	Unofficially Meets WWC Standards with Reservations	-0.021
PTA Comunitario	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	+0.008
PTA Comunitario	Business as Usual	Mathematics	High School	Unofficially Does Not Meet WWC Standards	–
Public School Choice Initiative (PSCI)					
PSCI: Focus Schools	Business as Usual	English Language Arts	Elementary, Middle and High School	Unofficially Does Not Meet WWC Standards	–
PSCI: Focus Schools	Business as Usual	English Language Arts	Elementary, Middle and High School	Unofficially Does Not Meet WWC Standards	–
PSCI: Relief Schools	Business as Usual	English Language Arts	Elementary, Middle and High School	Unofficially Does Not Meet WWC Standards	–
PSCI: Relief Schools	Business as Usual	English Language Arts	Elementary, Middle and High School	Unofficially Does Not Meet WWC Standards	–
PSCI: Focus Schools	Business as Usual	Mathematics	Elementary, Middle and High School	Unofficially Does Not Meet WWC Standards	–
PSCI: Focus Schools	Business as Usual	Mathematics	Elementary, Middle and High School	Unofficially Does Not Meet WWC Standards	–
PSCI: Relief Schools	Business as Usual	Mathematics	Elementary, Middle and High School	Unofficially Does Not Meet WWC Standards	–
PSCI: Relief Schools	Business as Usual	Mathematics	Elementary, Middle and High School	Unofficially Does Not Meet WWC Standards	–
Responding Effectively to Assessments with Curriculum and Teaching (REACT)					
REACT	Business as Usual	English Language Arts	Elementary, Middle and High School	Unofficially Does Not Meet WWC Standards	–
REACT	Business as Usual	Mathematics	Elementary, Middle and High School	Unofficially Does Not Meet WWC Standards	–

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Rio Grande Valley Center for Teaching and Leading Excellence: New Teacher Training (NTT)					
NTT	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.039
NTT	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards without Reservations	-0.128
NTT	Business as Usual	Mathematics	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.083
NTT	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards without Reservations	-0.298
NTT	Business as Usual	Science	Middle School	Not Reported	–
NTT	Business as Usual	Social Studies	Middle School	Unofficially Meets WWC Standards with Reservations	-0.120
Rural Math Excellence Partnership (RMEP)					
RMEP Project	Business as Usual	Mathematics	Middle and High School	Unofficially Meets WWC Standards with Reservations	-0.950
School of One					
School of One	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards without Reservations	+0.035
School of One[l]	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards without Reservations	-0.022
Schools to Watch (STW): School Transformation Network Project					
STW	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards with Reservations	-0.059
STW	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	-0.145
Science, Technology, Engineering and Math Education for the 21st Century (STEM21)					
STEM21 Academy	Business as Usual	Science	High School	Unofficially Meets WWC Standards with Reservations	+0.012
Spheres of Proud Achievement in Reading for Kids (SPARK) Program					
SPARK	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.355*
SPARK	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.226*

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
STEM Learning Opportunities Providing Equity (SLOPE)					
SLOPE: Algebra 1 Drop-in Units[l]	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	-0.101
SLOPE: Algebra 1 Drop-in Units with Summer College Awareness and Math Proficiency (C.A.M.P.)[m]	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards without Reservations	-0.147
STEM Summer Learning with VEX Robotics					
STEM Summer Learning Program with VEX Robotics[n]	No Participation in Any Summer Learning Program	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	0.000
Texas Tech University "Tech Teach" Program					
TechTeach	Business as Usual	English Language Arts	Elementary and Middle School	Unofficially Does Not Meet WWC Standards	–
Competency Based Training and Math Forward	Math Forward only	Mathematics	Middle and High School	Unofficially Does Not Meet WWC Standards	–
TechTeach	Business as Usual	Mathematics	Elementary, Middle and High School	Unofficially Meets WWC Standards with Reservations	-0.134[69]
Transforming Teacher Talent (t3) System					
Aspire t3 System[o]	Business as Usual	English Language Arts	Elementary, Middle and High School	Unofficially Does Not Meet WWC Standards	–
Aspire t3 System[o]	Business as Usual	Mathematics	Elementary and Middle School	Unofficially Does Not Meet WWC Standards	–
Turnaround with Increased Learning Time (TILT)					
TILT[p]	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards with Reservations	+0.384*
TILT[p]	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	+0.021
TILT[p]	Business as Usual	Science	Middle School	Unofficially Meets WWC Standards with Reservations	+0.089
We Are A Village					
We Are A Village	Business as Usual	English Language Arts	Early Childhood and Elementary School	Unofficially Ineligible for WWC Review	–
We Are A Village	Business as Usual	English Language Arts	Early Childhood and Elementary School	Not Reported	–

[69] Summary findings for TechTeach are not reported in Appendix C because the modal evidence rating across the findings reported is Unofficially Does Not Meet WWC Standards. One of the three findings is reported here because that finding received an evidence rating of Unofficially Meets WWC Standards with Reservations.

Intervention	Comparison	Outcome Domain (see Appendix B)	Educational Level	Unofficial WWC Evidence Rating	Effect Size
We Are A Village (Year 2)	We Are A Village (Year 1)	English Language Arts	Early Childhood and Elementary School	Not Reported	–
Write Up!					
Write Up![q]	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards with Reservations	-0.048
Write Up![q]	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards with Reservations	-0.198
Write Up![q]	Business as Usual	English Language Arts	Middle and High school	Unofficially Meets WWC Standards with Reservations	+0.022
Write Up![q]	Business as Usual	English Language Arts	Middle and High school	Not Reported	–

Notes:
[a] Evaluation findings included here are from the RCT design study only and do not include findings from an RD study that was also conducted by the evaluator.
[b] College Readiness Program was not implemented with adequate fidelity.
[c] Northeast Tennessee College and Career Ready Consortium was not implemented with adequate fidelity.
[d] ANet was not implemented with adequate fidelity.
[e] Arts Achieve was not implemented with adequate fidelity.
[f] These Every Child Ready analyses were not considered independent.
[g] FMTI, with and without Early Childhood Master's Program, was not implemented with adequate fidelity.
[h] Math Apps was not implemented with adequate fidelity.
[i] Making Time for What Matters Most was not implemented with adequate fidelity.
[j] Pathways to STEM Initiative was not implemented with adequate fidelity.
[k] See Appendix F for findings from analyses that were not prespecified.
[l] This School of One analysis was not considered independent.
[m] SLOPE: Algebra 1 Drop-in Units, with and without Summer C.A.M.P., was not implemented with adequate fidelity.
[n] STEM Summer Learning Program with VEX Robotics was not implemented with adequate fidelity
[o] Aspire t3 System was not implemented with adequate fidelity.
[p] Turnaround with Increased Learning Time was not implemented with adequate fidelity.
[q] Write Up! was not implemented with adequate fidelity.

Appendix E: Findings Reported by 21 i3 Impact Evaluations from Prespecified Analyses on Non-Academic Outcomes

Intervention	Comparison	Outcome Domain[a]	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Development					
Arts Achieve: Impacting Student Success in the Arts					
Arts Achieve[b]	Business as Usual	Art Achievement	Elementary, Middle, and High School	Unofficially Meets WWC Standards with Reservations	+0.095
Everyday Arts for Special Education (EASE)					
EASE	Business as Usual	Social/Emotional Learning	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.179*
Exceptional Coaching for Early Language and Literacy – Enhanced (ExCELL-E)					
ExCELL-E	Business as Usual	Overall Quality of Instructional Environment	Early Childhood and Elementary School	Unofficially Does Not Meet WWC Standards	–
Facilitating Long-Term Improvements in Graduation and Higher Education for Tomorrow (FLIGHT)					
FLIGHT	Business as Usual	Behavior	Middle and High School	Unofficially Meets WWC Standards with Reservations	+0.007
FLIGHT	Business as Usual	Motivation	Middle and High School	Unofficially Meets WWC Standards with Reservations	-0.085
Florida Master Teacher Initiative (FMTI)					
FMTI[c]	Business as Usual	Assessment Informed Instruction	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.082
FMTI with Early Childhood Master's Program[c]	Business as Usual	Assessment Informed Instruction	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.076
FMTI with Early Childhood Master's Program[c]	Business as Usual	Classroom Instruction: Classroom Organization	Elementary School	Unofficially Meets WWC Standards with Reservations	-0.475
FMTI with Early Childhood Master's Program[c]	Business as Usual	Classroom Instruction: Emotional Support	Elementary School	Unofficially Meets WWC Standards with Reservations	-0.046
FMTI with Early Childhood Master's Program[c]	Business as Usual	Classroom Instruction: Instructional Support	Elementary School	Unofficially Does Not Meet WWC Standards	–
FMTI[c]	Business as Usual	Collaboration Around Instruction	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.018
FMTI with Early Childhood Master's Program[c]	Business as Usual	Collaboration around Instruction	Elementary School	Unofficially Does Not Meet WWC Standards	–

Intervention	Comparison	Outcome Domain[a]	Educational Level	Unofficial WWC Evidence Rating	Effect Size
FMTI[c]	Business as Usual	Culturally Responsive Instruction	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.046
FMTI with Early Childhood Master's Program[c]	Business as Usual	Culturally Responsive Instruction	Elementary School	Unofficially Does Not Meet WWC Standards	–
FMTI[c]	Business as Usual	Developmentally Appropriate Instruction	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.041
FMTI with Early Childhood Master's Program[c]	Business as Usual	Developmentally Appropriate Instruction	Elementary School	Unofficially Meets WWC Standards with Reservations	-0.087
FMTI[c]	Business as Usual	Differentiated Instruction	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.147*
FMTI with Early Childhood Master's Program[c]	Business as Usual	Differentiated Instruction	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.100
FMTI[c]	Business as Usual	Early Childhood Teaching Knowledge	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.055
FMTI with Early Childhood Master's Program[c]	Business as Usual	Early Childhood Teaching Knowledge	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.629*
FMTI[c]	Business as Usual	Effective Principal P-3 Leadership	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.102
FMTI with Early Childhood Master's Program[c]	Business as Usual	Effective Principal P-3 Leadership	Elementary School	Unofficially Does Not Meet WWC Standards	–
FMTI[c]	Business as Usual	Emphasis on Higher-order Thinking Skills	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.004
FMTI with Early Childhood Master's Program[c]	Business as Usual	Emphasis on Higher-order Thinking Skills	Elementary School	Unofficially Does Not Meet WWC Standards	–
FMTI[c]	Business as Usual	Family Partnerships	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.066
FMTI with Early Childhood Master's Program[c]	Business as Usual	Family Partnerships	Elementary School	Unofficially Does Not Meet WWC Standards	–
FMTI[c]	Business as Usual	Frequent Use of a Variety of Assessments	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.059
FMTI with Early Childhood Master's Program[c]	Business as Usual	Frequent Use of a Variety of Assessments	Elementary School	Unofficially Meets WWC Standards with Reservations	-0.212

Intervention	Comparison	Outcome Domain[a]	Educational Level	Unofficial WWC Evidence Rating	Effect Size
FMTI[c]	Business as Usual	Governance Activities	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.931*
FMTI with Early Childhood Master's Program[c]	Business as Usual	Governance Activities	Elementary School	Unofficially Meets WWC Standards with Reservations	+6.537*
FMTI[c]	Business as Usual	Learner-centered Instruction	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.094
FMTI with Early Childhood Master's Program[c]	Business as Usual	Learner-centered Instruction	Elementary School	Unofficially Does Not Meet WWC Standards	–
FMTI[c]	Business as Usual	Outreach Activities	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.811*
FMTI with Early Childhood Master's Program[c]	Business as Usual	Outreach Activities	Elementary School	Unofficially Meets WWC Standards with Reservations	+2.941*
FMTI[c]	Business as Usual	Teacher Leadership	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.049
FMTI with Early Childhood Master's Program[c]	Business as Usual	Teacher Leadership	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.460*
FMTI[c]	Business as Usual	Teaching Knowledge	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.047
FMTI with Early Childhood Master's Program[c]	Business as Usual	Teaching Knowledge	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.564*
FMTI[c]	Business as Usual	Trusting Relationship between Teachers	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.024
FMTI with Early Childhood Master's Program[c]	Business as Usual	Trusting Relationship between Teachers	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.184
InnovateNYC Ecosystem					
Gap Apps[d]	Business as Usual	Opinions about Technology	Middle and High School	Unofficially Meets WWC Standards with Reservations	+0.098
Math Apps[d]	Business as Usual	Perceptions about Math	Middle and High School	Unofficially Meets WWC Standards with Reservations	+0.149
Math Apps[d]	Business as Usual	Perceptions about Math	Middle and High School	Unofficially Does Not Meet WWC Standards	–
Math Apps[d]	Business as Usual	Perceptions about Math	Middle and High School	Unofficially Does Not Meet WWC Standards	–
Gap Apps[d]	Business as Usual	Teacher's Use of Technology in the Classroom	Middle and High School	Unofficially Meets WWC Standards with Reservations	+0.187

Intervention	Comparison	Outcome Domain[a]	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Gap Apps[d]	Business as Usual	Teacher's Use of Technology in the Classroom	Middle and High School	Unofficially Meets WWC Standards with Reservations	-0.260
Gap Apps[d]	Business as Usual	Teacher's Use of Technology in the Classroom	Middle and High School	Unofficially Does Not Meet WWC Standards	–
Making Time for What Matters Most					
Making Time for What Matters Most[e]	Business as Usual	Attendance	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates					
Oakland Accelerates	Business as Usual	Participation in College Entrance Placement and Performance Exams	High School	Unofficially Does Not Meet WWC Standards	–
Oakland Accelerates	Business as Usual	Participation in College Entrance Placement and Performance Exams	High School	Unofficially Does Not Meet WWC Standards	–
Ounce of Prevention Fund Professional Development Initiative (PDI)					
PDI	Business as Usual	Child's Social & Emotional Development	Early Childhood	Unofficially Meets WWC Standards with Reservations	+0.224
PDI	Business as Usual	Classroom Organization	Early Childhood	Unofficially Meets WWC Standards with Reservations	+0.187
PDI	Business as Usual	Diverse Learning Opportunities	Early Childhood	Not Reported	–
PDI	Business as Usual	Diverse Learning Opportunities	Early Childhood	Not Reported	–
PDI	Business as Usual	Diverse Learning Opportunities	Early Childhood	Not Reported	–
PDI	Business as Usual	Emotional Supports	Early Childhood	Unofficially Meets WWC Standards with Reservations	+1.146
PDI	Business as Usual	Emotional Supports	Early Childhood	Not Reported	–
PDI	Business as Usual	Facilitation of Learning	Early Childhood	Unofficially Meets WWC Standards with Reservations	+0.825
PDI	Business as Usual	Facilitation of Learning	Early Childhood	Not Reported	–
PDI	Business as Usual	Oral Language and Linguistic Environment	Early Childhood	Not Reported	–

Intervention	Comparison	Outcome Domain[a]	Educational Level	Unofficial WWC Evidence Rating	Effect Size
PDI	Business as Usual	Oral Language and Linguistic Environment	Early Childhood	Not Reported	–
PDI	Business as Usual	Oral Language and Linguistic Environment	Early Childhood	Not Reported	–
PDI	Business as Usual	Program Structure	Early Childhood	Not Reported	–
PDI	Business as Usual	Program Structure	Early Childhood	Not Reported	–
PDI	Business as Usual	Program Structure	Early Childhood	Not Reported	–
PDI	Business as Usual	Quality of Interactions	Early Childhood	Not Reported	–
PDI	Business as Usual	Quality of Interactions	Early Childhood	Not Reported	–
PDI	Business as Usual	Quality of Interactions	Early Childhood	Not Reported	–
Rio Grande Valley Center for Teaching and Leading Excellence: New Teacher Training (NTT)					
Leadership Skills Training	Business as Usual	Teacher Leader Attrition	Elementary, Middle, and High School	Unofficially Meets WWC Standards without Reservations	-0.280
Skillful Teacher Institute	Business as Usual	Teacher Leader Attrition	Elementary, Middle, and High School	Unofficially Meets WWC Standards without Reservations	-2.393
Leadership Skills Training	Business as Usual	Teacher Leader Job Satisfaction	Elementary, Middle, and High School	Unofficially Meets WWC Standards without Reservations	+0.237
Skillful Teacher Institute	Business as Usual	Teacher Leader Job Satisfaction	Elementary, Middle, and High School	Unofficially Meets WWC Standards without Reservations	-0.034
Leadership Skills Training	Business as Usual	Teacher Leader Self-efficacy	Elementary, Middle, and High School	Unofficially Meets WWC Standards without Reservations	+0.071
Skillful Teacher Institute	Business as Usual	Teacher Leader Self-efficacy	Elementary, Middle, and High School	Unofficially Meets WWC Standards without Reservations	-0.395*
Rural Math Excellence Partnership (RMEP)					
RMEP Project	Same students prior to exposure	Attitudes toward Math	Middle and High School	Unofficially Does Not Meet WWC Standards	–
RMEP Project	Same students prior to exposure	Attitudes toward Math	Middle and High School	Unofficially Does Not Meet WWC Standards	–
RMEP Project	Same students prior to exposure	Attitudes toward Math	Middle and High School	Unofficially Does Not Meet WWC Standards	–
RMEP Project	Same students prior to exposure	Interest in STEM Careers	Middle and High School	Unofficially Does Not Meet WWC Standards	–

Intervention	Comparison	Outcome Domain[a]	Educational Level	Unofficial WWC Evidence Rating	Effect Size
RMEP Project	Same students prior to exposure	Interest in STEM Careers	Middle and High School	Unofficially Does Not Meet WWC Standards	–
RMEP Project	Same students prior to exposure	Interest in STEM Careers	Middle and High School	Unofficially Does Not Meet WWC Standards	–
RMEP Project	Same students prior to exposure	Interest in STEM Careers	Middle and High School	Unofficially Does Not Meet WWC Standards	–
RMEP Project	Same students prior to exposure	Interest in STEM Careers	Middle and High School	Unofficially Does Not Meet WWC Standards	–
RMEP Project	Same students prior to exposure	Interest in STEM Careers	Middle and High School	Unofficially Does Not Meet WWC Standards	–
RMEP Project	Same students prior to exposure	Interest in STEM Careers	Middle and High School	Unofficially Does Not Meet WWC Standards	–
School of One					
School of One	Business as Usual	Academic Behaviors and Attitudes	Middle School	Unofficially Meets WWC Standards without Reservations	-0.186
School of One	Business as Usual	Teacher Fulfillment	Middle School	Unofficially Does Not Meet WWC Standards	–
School of One	Business as Usual	Technological Knowledge	Middle School	Not Reported	–
Science, Technology, Engineering and Math Education for the 21st Century (STEM21)					
STEM21 Academy	Business as Usual	Career and College-Going Interest in STEM	High School	Unofficially Meets WWC Standards with Reservations	+0.187
STEM 21 Academy Afterschool Program[f]	Business as Usual	Interest in STEM Subjects	Middle School	Unofficially Meets WWC Standards with Reservations	-0.186
STEM Summer Learning with VEX Robotics					
STEM Summer Learning Program with VEX Robotics[g]	STEM Summer Learning Program without VEX Robotics (with Arts or Sports)	Aspirations to Study STEM in College	Middle School	Unofficially Meets WWC Standards with Reservations	+0.233
STEM Summer Learning Program with VEX Robotics[g]	STEM Summer Learning Program without VEX Robotics (with Arts or Sports)	Aspirations to Study STEM in College	Middle School	Unofficially Meets WWC Standards with Reservations	+0.068
STEM Summer Learning Program with VEX Robotics[g]	No Participation in Any Summer Learning Program	Attendance	Middle School	Unofficially Meets WWC Standards with Reservations	+0.094
STEM Summer Learning Program with VEX Robotics[g]	STEM Summer Learning Program without VEX Robotics (with Arts or Sports)	College Aspirations	Middle School	Unofficially Meets WWC Standards with Reservations	+0.146

Intervention	Comparison	Outcome Domain[a]	Educational Level	Unofficial WWC Evidence Rating	Effect Size
STEM Summer Learning Program with VEX Robotics[g]	STEM Summer Learning Program without VEX Robotics (with Arts or Sports)	College Aspirations	Middle School	Unofficially Meets WWC Standards with Reservations	-0.173
STEM Summer Learning Program with VEX Robotics[g]	STEM Summer Learning Program without VEX Robotics (with Arts or Sports)	STEM Career Aspirations	Middle School	Unofficially Meets WWC Standards with Reservations	+0.055
STEM Summer Learning Program with VEX Robotics[g]	STEM Summer Learning Program without VEX Robotics (with Arts or Sports)	STEM Career Aspirations	Middle School	Unofficially Meets WWC Standards with Reservations	-0.014
Turnaround with Increased Learning Time (TILT)					
TILT[h]	Business as Usual	Attendance	Middle School	Unofficially Meets WWC Standards with Reservations	+0.032
TILT[h]	Business as Usual	Student Engagement	Middle School	Unofficially Does Not Meet WWC Standards	–
We Are A Village					
We Are A Village	Business as Usual	Attendance	Early Childhood and Elementary School	Unofficially Ineligible for WWC Review	–
We Are A Village (Year 2)	We Are A Village (Year 1)	Attendance	Early Childhood and Elementary School	Not Reported	–
We Are A Village (Year 2)	We Are A Village (Year 1)	Attendance	Early Childhood and Elementary School	Not Reported	–
We Are A Village[i]	Business as Usual	Behavior	Early Childhood and Elementary School	Unofficially Ineligible for WWC Review	–
Validation					
The Baby Family and Child Education (FACE) Program					
Baby FACE Program	Business as Usual	Parent Perception of Socio-emotional Development	Early Childhood	Unofficially Does Not Meet WWC Standards	–
Baby FACE Program	Business as Usual	Parent Perception of Socio-emotional Development	Early Childhood	Unofficially Does Not Meet WWC Standards	–
Children's Literacy Initiative (CLI) Program					
CLI	Business as Usual	Classroom Literacy Environment	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.576*

Intervention	Comparison	Outcome Domain[a]	Educational Level	Unofficial WWC Evidence Rating	Effect Size
CLI	Business as Usual	Teacher Instruction	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.785*
enhancing Missouri's Instructional Networked Teaching Strategies (eMINTS)					
eMINTS	Business as Usual	21st Century Skills	Middle School	Unofficially Meets WWC Standards without Reservations	+0.025
eMINTS	Business as Usual	21st Century Skills	Middle School	Unofficially Meets WWC Standards with Reservations	+0.107
eMINTS with Intel Teach	Business as Usual	21st Century Skills	Middle School	Unofficially Meets WWC Standards without Reservations	+0.077
GO College (An Enhanced Version of Talent Search)					
GO College Whole School Program with Learning Community	GO College Whole School Program	Attendance	High School	Unofficially Meets WWC Standards without Reservations	-0.078
GO College Whole School Program with Learning Community	GO College Whole School Program	Behavior	High School	Unofficially Meets WWC Standards with Reservations	-0.016
GO College Whole School Program with Learning Community	GO College Whole School Program	Course Taking	High School	Unofficially Meets WWC Standards with Reservations	+0.031
StartSmart K-3 Plus Program					
Extended School Year	Business as Usual	Social Skills	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.066

Notes:
[a] These outcome domains were established by the evaluators, as these outcomes are outside of the domains defined in Appendix B.
[b] Arts Achieve was not implemented with adequate fidelity.
[c] FMTI, with and without Early Childhood Master's Program, was not implemented with adequate fidelity.
[d] Gap Apps and Math Apps were not implemented with adequate fidelity.
[e] Making Time for What Matters Most was not implemented with adequate fidelity.
[f] This STEM Education for the 21st Century analysis was not considered independent.
[g] STEM Summer Learning Program with VEX Robotics was not implemented with adequate fidelity.
[h] Turnaround with Increased Learning Time was not implemented with adequate fidelity.
[i] This We Are A Village analysis was not considered independent.

Appendix F: Findings Reported by 40 i3 Impact Evaluations from Analyses that Were Not Prespecified

The findings included in this appendix were reported by evaluators but were not from analyses that were prespecified. And while they are often from analyses of the same intervention condition, comparison condition, outcome domain, and educational level, they differ from those reported in Appendix D and Appendix E along one or more dimensions. These dimensions include limiting the analysis to a subgroup of those who were served, to those served in a particular year, to those served for a particular amount of time, or estimating the impact on a different outcome measure within the same outcome domain. Details on the differences between each of the findings reported here and those reported from prespecified analyses will be included in the data set to be released after this report.

Intervention	Comparison	Outcome Domain	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Scale-up					
KIPP					
Option of Continuing KIPP	No Option of Continuing KIPP	English Language Arts	High School	Unofficially Meets WWC Standards with Reservations	+0.123
Option of Continuing KIPP	No Option of Continuing KIPP	English Language Arts	High School	Unofficially Does Not Meet WWC Standards	–
Reading Recovery					
Reading Recovery	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+1.167*
Reading Recovery	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+1.065*
Reading Recovery	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.949*
Reading Recovery	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.641*
Reading Recovery	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.589*
Reading Recovery	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.546*
Reading Recovery	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.521*
Reading Recovery	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.498*
Reading Recovery	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.474*
Reading Recovery	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.467*
Reading Recovery	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.444*

Intervention	Comparison	Outcome Domain	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Success for All (SFA)					
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.177*
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.091
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.050
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.050
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.029
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.028
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.013
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.004
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.019
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.020
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.045
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.047
SFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.116
Teach for America (TFA)					
TFA	Business as Usual	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.138
TFA	Business as Usual	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.131*
TFA	Business as Usual	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.085
TFA	Business as Usual	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.030
TFA	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.049

Intervention	Comparison	Outcome Domain	Educational Level	Unofficial WWC Evidence Rating	Effect Size
TFA	Business as Usual	Mathematics	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.142
TFA	Business as Usual	Mathematics	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.095
TFA	Business as Usual	Mathematics	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.060
TFA	Business as Usual	Mathematics	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.041
TFA	Business as Usual	Mathematics	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.009
Validation					
Advanced ASSET Professional Development					
Advanced ASSET Professional Development after Science: It's Elementary	No Current Advanced ASSET Professional Development after Prior Experience with Science: It's Elementary	Mathematics	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.304
Advanced ASSET Professional Development after Science: It's Elementary	No Current Advanced ASSET Professional Development after Prior Experience with Science: It's Elementary	Mathematics	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.204
Advanced ASSET Professional Development after Science: It's Elementary	No Current Advanced ASSET Professional Development after Prior Experience with Science: It's Elementary	Mathematics	Elementary School	Unofficially Does Not Meet WWC Standards	–
Advanced ASSET Professional Development after Science: It's Elementary	No Current Advanced ASSET Professional Development after Prior Experience with Science: It's Elementary	Mathematics	Elementary School	Unofficially Does Not Meet WWC Standards	–

Intervention	Comparison	Outcome Domain	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Advanced ASSET Professional Development after Science: It's Elementary	No Current or Prior State-wide Science Curriculum or Training Initiatives	Mathematics	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.658*
Advanced ASSET Professional Development after Science: It's Elementary	No Current or Prior State-wide Science Curriculum or Training Initiatives	Mathematics	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.489
Advanced ASSET Professional Development after Science: It's Elementary	No Current or Prior State-wide Science Curriculum or Training Initiatives	Mathematics	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.290
Advanced ASSET Professional Development after Science: It's Elementary	No Current or Prior State-wide Science Curriculum or Training Initiatives	Mathematics	Elementary School	Unofficially Meets WWC Standards with Reservations	-0.066
Children's Literacy Initiative (CLI) Program					
CLI	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.198*
CLI	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.163*
CLI	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.114
College Readiness Program (CRP)					
CRP[a]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	+1.092*
CRP[a]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	+0.895*
CRP[a]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	+0.885*
CRP[a]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	+0.796*
CRP[a]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	+0.779*
CRP[a]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	+0.760*
CRP[a]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	+0.750*

Intervention	Comparison	Outcome Domain	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Diplomas Now					
Diplomas Now	Business as Usual	Educational Progress And Attainment	Middle and High School	Unofficially Meets WWC Standards without Reservations	+0.062
Diplomas Now	Business as Usual	Educational Progress And Attainment	Middle and High School	Unofficially Meets WWC Standards without Reservations	+0.050
Diplomas Now	Business as Usual	Educational Progress And Attainment	Middle and High School	Unofficially Meets WWC Standards without Reservations	+0.018
Diplomas Now	Business as Usual	Educational Progress And Attainment	Middle and High School	Unofficially Meets WWC Standards without Reservations	+0.016
GO College (An Enhanced Version of Talent Search)					
GO College Whole School Program	Business as Usual	English Language Arts	High School	Unofficially Meets WWC Standards with Reservations	+0.270
GO College Whole School Program with Learning Community	GO College Whole School Program	English Language Arts	High School	Unofficially Meets WWC Standards without Reservations	+0.054
GO College Whole School Program	Business as Usual	Mathematics	High School	Unofficially Meets WWC Standards with Reservations	+0.070
GO College Whole School Program with Learning Community	GO College Whole School Program	Mathematics	High School	Unofficially Meets WWC Standards without Reservations	-0.041
GO College Whole School Program	Business as Usual	Educational progress and attainment	High School	Unofficially Meets WWC Standards with Reservations	+0.533
GO College Whole School Program with Learning Community	GO College Whole School Program	Educational progress and attainment	High School	Unofficially Meets WWC Standards without Reservations	-0.062
GO College Whole School Program with Learning Community	GO College Whole School Program	Educational progress and attainment	High School	Unofficially Meets WWC Standards with Reservations	-0.004
GO College Whole School Program with Learning Community	GO College Whole School Program	Educational progress and attainment	High School	Unofficially Meets WWC Standards with Reservations	-0.013

Intervention	Comparison	Outcome Domain	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Northeast Tennessee College and Career Ready Consortium					
Consortium[b]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	-0.007
Consortium[b]	Business as Usual	Educational progress and attainment	High School	Unofficially Meets WWC Standards with Reservations	+0.519*
Consortium[b]	Business as Usual	Educational progress and attainment	High School	Unofficially Meets WWC Standards with Reservations	+0.081*
Consortium[b]	Business as Usual	Educational progress and attainment	High School	Unofficially Meets WWC Standards with Reservations	+0.054
Consortium[b]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	+1.375*
Consortium[b]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	+0.341
Consortium[b]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	+0.004
Consortium[b]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	-0.018
Consortium[b]	Business as Usual	General Achievement	High School	Unofficially Meets WWC Standards with Reservations	-0.620
Consortium[b]	Business as Usual	Educational progress and attainment	High School	Unofficially Meets WWC Standards with Reservations	+0.529*
Consortium[b]	Business as Usual	Educational progress and attainment	High School	Unofficially Meets WWC Standards with Reservations	+0.010
Reading Apprenticeship (RA)					
Reading Apprenticeship	Business as Usual	English Language Arts	High School	Unofficially Meets WWC Standards without Reservations	+0.327*
Reading Enhances Achievement During the Summer (READS)					
READS	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.056
READS	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.045*
READS	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.034*
READS	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.020
READS	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.007

Intervention	Comparison	Outcome Domain	Educational Level	Unofficial WWC Evidence Rating	Effect Size
READS	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.053
READS - PIF	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+1.157*
READS - PIF	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.873*
StartSmart K-3 Plus Program					
Extended School Year	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.310*
Extended School Year	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.278*
Extended School Year	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.212*
Extended School Year	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.194*
Extended School Year	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.099*
Extended School Year	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.059
Extended School Year	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.046
Extended School Year	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.041
Extended School Year	Business as Usual	Mathematics	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.202*
Extended School Year	Business as Usual	Mathematics	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.156*
Virginia Initiative for Science Teaching and Achievement (VISTA)					
VISTA Teacher Professional Development	Business as Usual	Science	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.068
VISTA Teacher Professional Development	Business as Usual	Science	Middle School	Unofficially Does Not Meet WWC Standards	–
VISTA Teacher Professional Development	Business as Usual	Science	High School	Unofficially Does Not Meet WWC Standards	–
Development					
Advanced Placement (AP) Insight					
AP Insight	Business as Usual	Science	High School	Unofficially Meets WWC Standards with Reservations	+0.199
AP Insight	Business as Usual	Science	High School	Unofficially Meets WWC Standards with Reservations	+0.128

Intervention	Comparison	Outcome Domain	Educational Level	Unofficial WWC Evidence Rating	Effect Size
AP Insight	Business as Usual	Science	High School	Unofficially Meets WWC Standards with Reservations	+0.071
AP Insight	Business as Usual	Science	High School	Unofficially Meets WWC Standards with Reservations	+0.027
Advancement via Individual Determination (AVID)					
AVID schoolwide intervention	Business as Usual	English Language Arts	Middle and High School	Unofficially Does Not Meet WWC Standards	–
AVID schoolwide intervention	Business as Usual	Mathematics	Middle and High School	Unofficially Does Not Meet WWC Standards	–
AVID schoolwide intervention	Business as Usual	Educational progress and attainment	High School	Unofficially Does Not Meet WWC Standards	–
Around the Corner (ATC)					
Around the Corner	Existing SFA Program	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	-0.569*
Around the Corner	Existing SFA Program	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.234
Around the Corner	Existing SFA Program	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.226
Around the Corner	Existing SFA Program	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.021
Around the Corner	Existing SFA Program	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	+0.011
Around the Corner	Existing SFA Program	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	-0.088
Around the Corner	Existing SFA Program	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards without Reservations	-0.144
Around the Corner	Existing SFA Program	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards with Reservations	+0.240
Around the Corner	Existing SFA Program	English Language Arts	Early Childhood and Elementary School	Unofficially Meets WWC Standards with Reservations	+0.007
Arts Achieve: Impacting Student Success in the Arts					
Arts Achieve[c]	Business as Usual	English Language Arts	Elementary, Middle, and High School	Unofficially Does Not Meet WWC Standards	–

Intervention	Comparison	Outcome Domain	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Arts for Learning (A4L) Lessons					
A4L	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.009
A4L	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.029
Bay State Reading Institute (BSRI)					
Bay State Reading Institute	Business as Usual	English Language Arts	Elementary School	Unofficially Does Not Meet WWC Standards	–
Bay State Reading Institute	Business as Usual	English Language Arts	Elementary School	Unofficially Does Not Meet WWC Standards	–
Bay State Reading Institute	Business as Usual	English Language Arts	Elementary School	Unofficially Does Not Meet WWC Standards	–
Bay State Reading Institute	Business as Usual	English Language Arts	Elementary School	Unofficially Does Not Meet WWC Standards	–
Bay State Reading Institute	Business as Usual	English Language Arts	Elementary School	Unofficially Does Not Meet WWC Standards	–
Boston Teacher Residency (BTR)					
Newly Hired BTR Graduates	Business as Usual	English Language Arts	Elementary and Middle School	Unofficially Meets WWC Standards with Reservations	-0.162*
Newly Hired BTR Graduates	Business as Usual	Mathematics	Elementary and Middle School	Unofficially Meets WWC Standards with Reservations	-0.054
Building Assets, Reducing Risks (BARR) Model					
BARR	Business as Usual	English Language Arts	High School	Unofficially Meets WWC Standards without Reservations	+0.238*
BARR	Business as Usual	English Language Arts	High School	Unofficially Meets WWC Standards without Reservations	+0.150*
BARR	Business as Usual	English Language Arts	High School	Unofficially Meets WWC Standards without Reservations	+0.139
BARR	Business as Usual	Mathematics	High School	Unofficially Meets WWC Standards without Reservations	+0.518*
BARR	Business as Usual	Mathematics	High School	Unofficially Meets WWC Standards without Reservations	+0.381*
BARR	Business as Usual	Mathematics	High School	Unofficially Meets WWC Standards without Reservations	+0.333*
BARR	Business as Usual	Educational Progress And Attainment	High School	Unofficially Meets WWC Standards without Reservations	+0.251*
BARR	Business as Usual	Educational Progress And Attainment	High School	Unofficially Meets WWC Standards without Reservations	+0.194

Intervention	Comparison	Outcome Domain	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Collaborative Organizational Model to Promote Aligned Support Structures (COMPASS)					
COMPASS	Business as Usual	English Language Arts	Elementary School	Unofficially Does Not Meet WWC Standards	–
COMPASS	Business as Usual	English Language Arts	Elementary School	Unofficially Does Not Meet WWC Standards	–
COMPASS	Business as Usual	English Language Arts	Elementary School	Unofficially Does Not Meet WWC Standards	–
COMPASS	Business as Usual	English Language Arts	Middle School	Unofficially Does Not Meet WWC Standards	–
COMPASS	Business as Usual	English Language Arts	Middle School	Unofficially Does Not Meet WWC Standards	–
COMPASS	Business as Usual	English Language Arts	Middle School	Unofficially Does Not Meet WWC Standards	–
CollegeYES					
CollegeYES	Business as Usual	Science	High School	Unofficially Does Not Meet WWC Standards	–
CollegeYES	Business as Usual	Science	High School	Unofficially Does Not Meet WWC Standards	–
CollegeYES	Business as Usual	Science	High School	Unofficially Does Not Meet WWC Standards	–
CollegeYES	Business as Usual	Science	High School	Unofficially Does Not Meet WWC Standards	–
CollegeYES	Business as Usual	Science	High School	Unofficially Does Not Meet WWC Standards	–
CollegeYES	Business as Usual	Academic Performance In School	High School	Unofficially Does Not Meet WWC Standards	–
CollegeYES	Business as Usual	Academic Performance In School	High School	Unofficially Does Not Meet WWC Standards	–
CollegeYES	Business as Usual	Academic Performance In School	High School	Unofficially Does Not Meet WWC Standards	–
CollegeYES	Business as Usual	Educational Progress And Attainment	High School	Unofficially Does Not Meet WWC Standards	–
CollegeYES	Business as Usual	Educational Progress And Attainment	High School	Unofficially Does Not Meet WWC Standards	–

Intervention	Comparison	Outcome Domain	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Convergence Academies (CA)					
CA	Business as Usual	English Language Arts	Elementary and Middle School	Unofficially Ineligible for WWC Review	–
CA	Business as Usual	Mathematics	Elementary and Middle School	Unofficially Ineligible for WWC Review	–
Curriculum 2.0					
Curriculum 2.0	Business as Usual	English Language Arts	Elementary School	Unofficially Ineligible for WWC Review	–
Curriculum 2.0	Business as Usual	English Language Arts	Elementary School	Unofficially Ineligible for WWC Review	–
Curriculum 2.0	Business as Usual	English Language Arts	Elementary School	Unofficially Ineligible for WWC Review	–
Exceptional Coaching for Early Language and Literacy – Enhanced (ExCELL-E)					
ExCELL-E	Business as Usual	English Language Arts	Early Childhood	Unofficially Does Not Meet WWC Standards	–
ExCELL-E	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.051
Facilitating Long-Term Improvements In Graduation and Higher Education for Tomorrow (FLIGHT)					
FLIGHT	Business as Usual	Educational Progress And Attainment	High School	Unofficially Does Not Meet WWC Standards	–
Florida Master Teacher Initiative (FMTI)					
FMTI[d]	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.005
FMTI[d]	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.022
FMTI[d]	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards with Reservations	-0.081
FMTI[d]	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards with Reservations	-0.114
FMTI[d]	Business as Usual	Mathematics	Elementary School	Unofficially Meets WWC Standards without Reservations	+0.012
FMTI[d]	Business as Usual	Mathematics	Elementary School	Unofficially Meets WWC Standards without Reservations	-0.024
FMTI[d]	Business as Usual	Mathematics	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.021
FMTI[d]	Business as Usual	Mathematics	Elementary School	Unofficially Meets WWC Standards with Reservations	-0.080
InnovateNYC Ecosystem					
Math Apps[e]	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	+0.246
Math Apps[e]	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	+0.099

Intervention	Comparison	Outcome Domain	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Math Apps[e]	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	-0.154
Math Apps[e]	Business as Usual	Mathematics	Middle School	Unofficially Does Not Meet WWC Standards	–
Oakland Accelerates					
Oakland Accelerates	Business as Usual	English Language Arts	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	English Language Arts	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	English Language Arts	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	Mathematics	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	Mathematics	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	Mathematics	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–
Oakland Accelerates	Business as Usual	Educational progress and attainment	High School	Unofficially Ineligible for WWC Review	–

Intervention	Comparison	Outcome Domain	Educational Level	Unofficial WWC Evidence Rating	Effect Size
Pathways to STEM Initiative (PSI)					
PSI[f]	Business as Usual	Science	Middle School	Unofficially Meets WWC Standards with Reservations	+0.031
Problem-Based Learning (PBL)					
PBL	Business as Usual	English Language Arts	High School	Unofficially Ineligible for WWC Review	–
PBL	Business as Usual	Mathematics	High School	Unofficially Ineligible for WWC Review	–
PBL	Business as Usual	Science	High School	Unofficially Ineligible for WWC Review	–
PBL	Business as Usual	Social Studies	High School	Unofficially Ineligible for WWC Review	–
PBL	Business as Usual	General Achievement	High School	Unofficially Ineligible for WWC Review	–
Rural Math Excellence Partnership (RMEP)					
RMEP Project	Business as Usual	Mathematics	Middle and High School	Unofficially Does Not Meet WWC Standards	–
RMEP Project	Business as Usual	Mathematics	Middle and High School	Unofficially Does Not Meet WWC Standards	–
RMEP Project	Business as Usual	Mathematics	Middle and High School	Unofficially Meets WWC Standards with Reservations	-0.190
STEM Learning Opportunities Providing Equity (SLOPE)					
SLOPE[g]	Business as Usual	Mathematics	Middle School	Unofficially Does Not Meet WWC Standards	–
STEM Summer Learning with VEX Robotics					
STEM Summer Learning Program with VEX Robotics[h]	No Summer Learning Program	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	+0.181
STEM Summer Learning Program with VEX Robotics[h]	No Summer Learning Program	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	+0.159
STEM Summer Learning Program with VEX Robotics[h]	No Summer Learning Program	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	+0.100
STEM Summer Learning Program with VEX Robotics[h]	No Summer Learning Program	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	+0.071

Intervention	Comparison	Outcome Domain	Educational Level	Unofficial WWC Evidence Rating	Effect Size
STEM Summer Learning Program with VEX Robotics[h]	No Summer Learning Program	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	-0.090
Texas Tech University "Tech Teach" Program					
Competency Based Training and Math Forward	Math Forward only	Mathematics	Middle and High School	Unofficially Meets WWC Standards with Reservations	+0.192*
Competency Based Training and Math Forward	Math Forward only	Mathematics	Middle and High School	Unofficially Does Not Meet WWC Standards	–
Turnaround with Increased Learning Time (TILT)					
TILT[i]	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards with Reservations	+0.233*
TILT[i]	Business as Usual	English Language Arts	Middle School	Unofficially Meets WWC Standards with Reservations	+0.055
TILT[i]	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	+0.067*
TILT[i]	Business as Usual	Mathematics	Middle School	Unofficially Meets WWC Standards with Reservations	-0.077*
TILT[i]	Business as Usual	Science	Middle School	Unofficially Meets WWC Standards with Reservations	-0.093
TILT[i]	Business as Usual	Science	Middle School	Unofficially Meets WWC Standards with Reservations	-0.105
We Are A Village					
We Are A Village	Business as Usual	English Language Arts	Elementary School	Unofficially Ineligible for WWC Review	–
Write Up!					
Write Up![j]	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards with Reservations	+0.111
Write Up![j]	Business as Usual	English Language Arts	Elementary School	Unofficially Meets WWC Standards with Reservations	-0.211
Write Up![j]	Business as Usual	English Language Arts	Middle and High School	Unofficially Meets WWC Standards with Reservations	-0.075

Notes:
[a] College Readiness Program was not implemented with adequate fidelity.
[b] Northeast Tennessee College and Career Ready Consortium was not implemented with adequate fidelity.
[c] Arts Achieve was not implemented with adequate fidelity.
[d] FMTI was not implemented with adequate fidelity.
[e] Math Apps was not implemented with adequate fidelity.
[f] Pathways to STEM Initiative was not implemented with adequate fidelity.
[g] SLOPE was not implemented with adequate fidelity.
[h] STEM Summer Learning Program with VEX Robotics was not implemented with adequate fidelity
[i] Turnaround with Increased Learning Time was not implemented with adequate fidelity.
[j] Write Up! was not implemented with adequate fidelity.

Appendix G: Alternative Approaches to Assessing Representativeness of the 67 i3 Impact Evaluations

Chapter 4 reports the percentage of i3 impact evaluations that were based on a sample that was representative of the population served by the i3 intervention. An evaluation sample was not considered representative of this population if it excluded more than 25 percent of the schools that received the i3 intervention (see Section 4.1.4 for more details). However, because thresholds of this type are inherently arbitrary, and the 25 percent threshold could be considered generous, the Abt Team explored whether the share of evaluations that satisfy the representativeness criteria would be much smaller under a more stringent threshold of 10 percent.

Exhibit G.1 compares the results from the two approaches. It shows that changing this threshold from 25 percent to 10 percent would reduce the percentage of i3 evaluations that are based on a representative sample from 79 percent to 63 percent.

Exhibit G.1: Representativeness Results under Different Exclusion Thresholds for Schools that Received the i3 Intervention

Representativeness of i3 Impact Evaluations	Primary Approach – Cannot Exclude More than 25% of Schools	Alternative Approach – Cannot Exclude More than 10% of Schools
No prespecified analyses	1 (1%)	1 (1%)
Yes	53 (79%)	42 (63%)
No	13 (19%)	24 (36%)

Note: Columns do not sum to 100% due to rounding.

Appendix H: Short-term and Long-term Goals Met by 67 i3 Impact and Implementation Evaluations

Intervention	Quality implementation evaluation	Impact evaluation considered independent	Unofficially Meets WWC Standards	Impact evaluation sample adequately representative of population served	MEETS SHORT-TERM GOALS OF I3	Evidence that the intervention was implemented with adequate fidelity	Evidence of at least one positive, statistically significant impact on student outcomes	MEETS LONG-TERM GOALS OF I3
Scale-up								
KIPP	✓	✓	✓	×	No	✓	✓	Yes
Reading Recovery	✓	✓	✓	✓	Yes	✓	✓	Yes
Success for All (SFA)	✓	✓	✓	×	No	✓	×	No
Teach for America (TFA)	✓	✓	✓	×	No	✓	×	No
Validation								
Advanced ASSET Professional Development	✓	✓	✓	✓	Yes	✓	×	No
The Baby Family and Child Education (FACE) Program	✓	✓	×	✓	No	✓	×	No
Children's Literacy Initiative (CLI) Program	✓	✓	✓	✓	Yes	✓	✓	Yes
Collaborative Strategic Reading (CSR)	✓	✓	✓	✓	Yes	✓	×	No
College Readiness Program (CRP)	✓	✓	✓	✓	Yes	×	✓	No
College-Ready Writers Program (CRWP)	✓	✓	✓	✓	Yes	✓	✓	Yes
Diplomas Now	✓	✓	✓	✓	Yes	✓	×	No
enhancing Missouri's Instructional Networked Teaching Strategies (eMINTS)	✓	✓	✓	✓	Yes	✓	✓	Yes
GO College (An Enhanced Version of Talent Search)	✓	✓	✓	✓	Yes	✓	×	No
Leadership Assistance for Science Education Reform (LASER)	✓	✓	✓	×	No	✓	×	No
Northeast Tennessee College and Career Ready Consortium	✓	✓	✓	✓	Yes	×	✓	No
Reading Apprenticeship (RA)	✓	✓	✓	✓	Yes	✓	×	No
Reading Enhances Achievement During the Summer (READS)	✓	✓	✓	✓	Yes	✓	×	No

Intervention	Quality implementation evaluation	Impact evaluation considered independent	Unofficially Meets WWC Standards	Impact evaluation sample adequately representative of population served	MEETS SHORT-TERM GOALS OF I3	Evidence that the intervention was implemented with adequate fidelity	Evidence of at least one positive, statistically significant impact on student outcomes	MEETS LONG-TERM GOALS OF I3
StartSmart K-3 Plus Program	✓	✓	✓	✓	Yes	✓	✓	Yes
Virginia Initiative for Science Teaching and Achievement (VISTA)	✓	✓	✓	✓	Yes	✓	×	No
Development								
The Achievement Network (ANet)	✓	✓	✓	✓	Yes	×	×	No
Advanced Placement (AP) Insight	✓	✓	✓	✓	Yes	✓	×	No
Advancement via Individual Determination (AVID)	✓	✓	×	✓	No	✓	×	No
Around the Corner (ATC)	✓	✓	✓	✓	Yes	✓	×	No
Arts Achieve: Impacting Student Success in the Arts	✓	✓	✓	✓	Yes	×	×	No
Arts for Learning (A4L) Lessons	✓	✓	✓	✓	Yes	✓	×	No
Bay State Reading Institute (BSRI)	✓	✓	✓	✓	Yes	✓	×	No
Boston Teacher Residency (BTR)	✓	✓	✓	✓	Yes	✓	×	No
Building Assets, Reducing Risks (BARR) Model	✓	✓	✓	×	No	✓	✓	Yes
Collaborative Organizational Model to Promote Aligned Support Structures (COMPASS)	✓	✓	×	×	No	✓	×	No
CollegeYES	✓	✓	×	×	No	✓	×	No
Convergence Academies	✓	✓	×	✓	No	✓	×	No
Curriculum 2.0	×	✓	×	✓	No	×	×	No
Data-Driven Decision Making and Information Technology Curricula in Schools	✓	✓	✓	✓	Yes	✓	×	No
E3TL Teacher Performance Evaluation System	✓	✓	×	✓	No	✓	×	No
EngageMe P.L.E.A.S.E.	✓	✓	✓	✓	Yes	✓	×	No
Every Child Ready (ECR)	✓	×	×	✓	No	✓	×	No
Everyday Arts for Special Education (EASE)	✓	✓	✓	✓	Yes	✓	×	No

Intervention	Quality implementation evaluation	Impact evaluation considered independent	Unofficially Meets WWC Standards	Impact evaluation sample adequately representative of population served	MEETS SHORT-TERM GOALS OF I3	Evidence that the intervention was implemented with adequate fidelity	Evidence of at least one positive, statistically significant impact on student outcomes	MEETS LONG-TERM GOALS OF I3
Exceptional Coaching for Early Language and Literacy - Enhanced (ExCELL-E)	✓	✓	✓	✓	Yes	✓	×	No
The Expository Reading and Writing Course (ERWC)	✓	✓	✓	×	No	✓	✓	Yes
Facilitating Long-Term Improvements in Graduation and Higher Education for Tomorrow (FLIGHT)	✓	✓	✓	✓	Yes	✓	×	No
Florida Master Teacher Initiative (FMTI)	✓	✓	✓	✓	Yes	×	×	No
InnovateNYC Ecosystem	✓	✓	✓	✓	Yes	×	×	No
Innovations in Early Mathematics	✓	✓	✓	×	No	✓	×	No
Integrating English Language Development (ELD) and Science	✓	✓	×	✓	No	✓	×	No
Internet-based Reading Apprenticeship Improving Science Education (iRAISE)	✓	✓	×	✓	No	✓	×	No
Making Time for What Matters Most	✓	✓	×	✓	No	×	×	No
New England Network for Personalization and Performance (NETWORK)	✓	✓	✓	✓	Yes	✓	×	No
Oakland Accelerates	✓	✓	×	✓	No	✓	×	No
Ounce of Prevention Fund Professional Development Initiative (PDI)	✓	✓	✓	✓	Yes	✓	×	No
Pathways to STEM Initiative (PSI)	✓	✓	✓	✓	Yes	×	×	No
Problem-Based Learning (PBL)[a]	✓	✓	×	✓	No	✓	×	No
PTA Comunitario	✓	✓	✓	×	No	✓	×	No
Public School Choice Initiative (PSCI)	✓	✓	×	×	No	✓	×	No
Responding Effectively to Assessments with Curriculum and Teaching (REACT)	✓	✓	×	×	No	✓	×	No

Intervention	Quality implementation evaluation	Impact evaluation considered independent	Unofficially Meets WWC Standards	Impact evaluation sample adequately representative of population served	MEETS SHORT-TERM GOALS OF I3	Evidence that the intervention was implemented with adequate fidelity	Evidence of at least one positive, statistically significant impact on student outcomes	MEETS LONG-TERM GOALS OF I3
Rio Grande Valley Center for Teaching and Leading Excellence: New Teacher Training (NTT)	✓	✓	✓	✓	Yes	✓	×	No
Rural Math Excellence Partnership (RMEP)	×	✓	✓	✓	No	×	×	No
School of One	✓	✓	✓	✓	Yes	✓	×	No
Schools to Watch (STW): School Transformation Network Project	✓	✓	✓	✓	Yes	✓	×	No
Science, Technology, Education and Math Education for the 21st Century (STEM21)	✓	✓	✓	✓	Yes	✓	×	No
Spheres of Proud Achievement in Reading for Kids (SPARK) Program	✓	✓	✓	✓	Yes	✓	✓	Yes
STEM Learning Opportunities Providing Equity (SLOPE)	✓	✓	✓	✓	Yes	×	×	No
STEM Summer Learning with VEX Robotics	✓	✓	✓	✓	Yes	×	×	No
Texas Tech University "Tech Teach" Program	✓	✓	×	✓	No	✓	×	No
Transforming Teacher Talent (t3) System	✓	✓	×	✓	No	×	×	No
Turnaround with Increased Learning Time (TILT)	✓	✓	✓	✓	Yes	×	✓	No
We Are A Village	✓	✓	×	×	No	✓	×	No
Write Up!	✓	✓	✓	✓	Yes	×	×	No
Total number of evaluations that meet the short- and long-term goals of i3 (N=67 evaluations)					40 (60%)			9 (13%)

Notes:
✓ indicates that the evaluation met the expectation;
× indicates that the evaluation did not meet the expectation;
[a] The impact evaluation of PBL did not prespecify analyses.

Appendix I: Alternative Approaches to Summarizing the Findings from the 67 i3 Impact Evaluations

Chapter 4 reports the percentage of i3 interventions that improved student academic outcomes, by summarizing findings from prespecified analyses that Unofficially Meet WWC Standards with or without Reservations. An intervention was considered to have improved student academic outcomes if it had a positive and statistically significant impact on at least one student academic outcome and no negative and statistically significant impacts on these outcomes, after correcting for multiple comparisons.[70] This appendix explores whether the percentage of interventions that improved student academic outcomes is sensitive to (1) the method of summarizing multiple findings or (2) the inclusion of findings based on analyses that were not prespecified.[71]

Alternative Methods for Summarizing Multiple Findings

We explored two alternative approaches to summarizing multiple findings. The first alternative used the most common, or modal, finding reported by the evaluation. Each finding was classified into one of the following categories:

- **Positive** if the estimate was positive and statistically significant, without any multiple comparisons adjustment;
- **Negative** if the estimate was negative and statistically significant, without any multiple comparisons adjustment; or
- **Null** if the estimate was not statistically significant.

Under this alternative, the summary finding for the intervention is based on the modal category across findings. An intervention was considered to have improved student academic outcomes if the modal category of findings was positive.

The second alternative used the average effect size across the impact estimates. We computed effect sizes from the estimates reported for prespecified analyses using procedures from the WWC Procedures & Standards Handbook, Version 3.0, calculated the simple average of those effect sizes, and tested whether the average effect size was statistically significant. It is important to note that this hypothesis test is affected by the correlation among the impact estimates caused by the fact that they were based on the same or similar samples. For each evaluation, we either determined that significance of the average effect size was not dependent on the size of the correlation—that is, the conclusion would be the same for all plausible values of the correlation—or we were able to estimate the correlation using information reported by the evaluator.

Under the second alternative, an intervention was considered to have improved student academic outcomes if the average effect size is positive and statistically significant—and *not* to have improved student academic outcomes if (1) the average effect size was negative and statistically significant or (2) the average effect size was not statistically significant (null).

---

[70] For a description of how multiple comparisons adjustments were made, see Appendix B.

[71] Findings reported by evaluators from analyses that were not prespecified can be found in Appendix F.

Exhibit I.1 compares the results from the two alternative approaches to the results from our primary approach reported in Chapter 4. Exhibits I.2 and I.3 show the pie charts that correspond to the alternative approaches, and that can be compared to Exhibit 4.6 in the report. Both alternatives give very similar results to our primary approach.[72]

Exhibit I.1: Summary of the Impacts on Student Academic Outcomes Using Our Primary Approach and Two Alternative Approaches (N=67 Impact Evaluations)

Impact	Primary Approach	Alternative #1 – Modal Finding	Alternative #2 – Average Effect Size
Positive	12	10	11
Null	36	38	35
Negative	1	1	3
No evidence	18	18	18

[72] The findings that are summarized in these three analyses can be found in Appendix D.

Exhibit I.2: Summary of Whether the i3 Impact Evaluations Found Statistically Significant Positive Impacts on Student Academic Outcomes – *Based on Modal Findings*

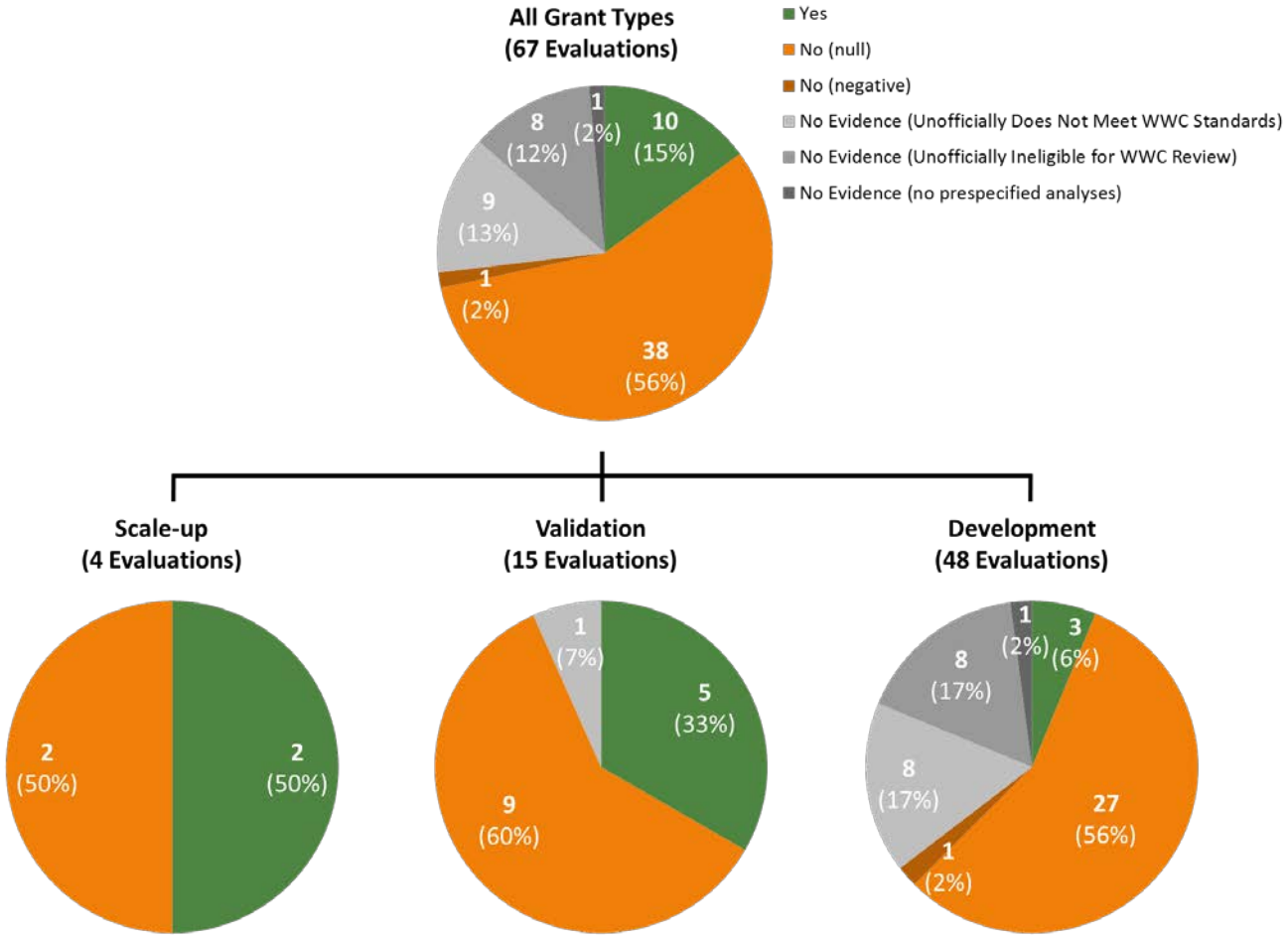
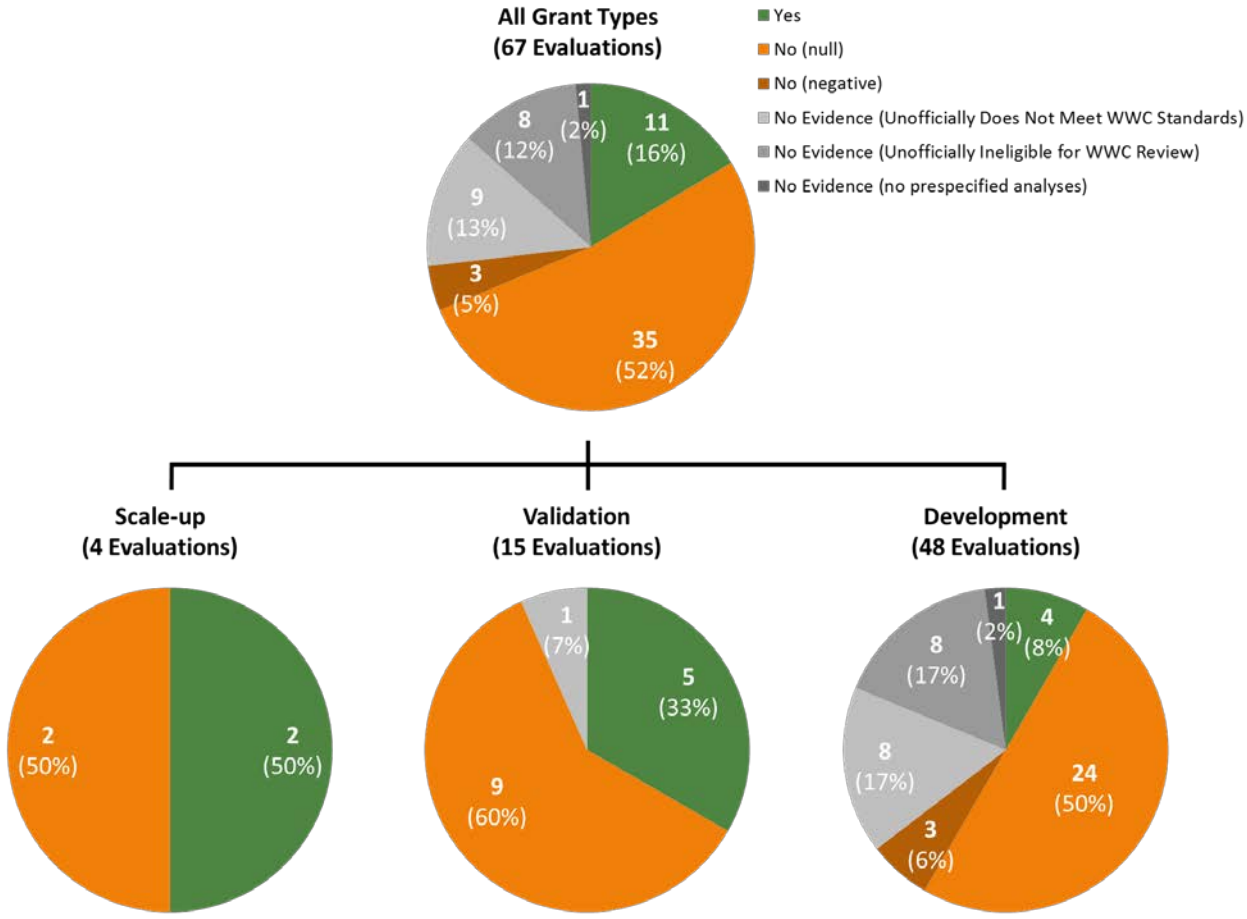


Exhibit I.3: Summary of Whether the i3 Impact Evaluations Found Statistically Significant Positive Impacts on Student Academic Outcomes – *Based on Average Effect Size*



Inclusion of Non-Prespecified Analyses

This report focuses primarily on findings from analyses that were prespecified by the evaluators (see Chapter 3). However, we collected all of the findings that the evaluators were willing to share, whether or not they were from prespecified analyses.[73] We conducted a sensitivity analysis to assess whether the conclusions about the impacts of the i3 interventions on student academic outcomes presented in Chapter 4 would be substantially different if the assessment had included findings from both prespecified and non-prespecified analyses.

In principle, including findings from non-prespecified analyses could either increase or decrease the percentage of i3 interventions that found at least one statistically significant positive impact on student academic outcomes. If evaluators selectively reported more favorable findings from non-prespecified analysis, we would expect the inclusion of these findings to increase the percentage of interventions that

[73] Findings reported by evaluators from analyses that were not prespecified can be found in Appendix F.

had a significant positive impact on at least one student academic outcome. However, if the analyses that were not prespecified are ones that would increase the number of comparisons among which we will apply multiple comparisons corrections, we might expect the inclusion of these findings to decrease the percentage of interventions that had a significant positive impact on at least one student academic outcome.

Exhibit I.4 shows the distribution of summary findings, based on findings from both prespecified and non-prespecified analyses, using our primary approach to summarizing multiple findings. Including findings from non-prespecified analyses changes the summary impact finding for only six evaluations. The evaluation that did not prespecify analyses receives a summary impact finding of "no evidence" because all of the findings from non-prespecified analyses were Unofficially Ineligible for WWC Review. Including findings from non-prespecified analyses reduces the number of evaluations with null findings by five; two of these evaluations would be summarized as having negative impact findings, one as having mixed findings, and one as having positive findings. The remaining evaluation receives a summary impact finding of "no evidence' because including findings from non-prespecified analyses increases the number of findings that Unofficially DNM WWC Standards such that it is now the modal rating for that evaluation. Including these additional findings in our summary does not change the conclusion that almost one-fifth of the i3 interventions improved student academic outcomes.

Exhibit I.4: Summary of Whether the i3 Impact Evaluations Found Statistically Significant Positive Impacts on Student Academic Outcomes – *Including Prespecified and Non-Prespecified Analyses*