November 2015

Improvement Science
in New York City Schools:

A White Paper on the Work of Eskolta School Research and Design and the New York City Department of Education's Office of Postsecondary Readiness

Written by:
Michael Rothman and Savanna Honerkamp-Smith
Eskolta School Research and Design

Edited by:
Deirdre Faughey
Eskolta School Research and Design

New York City Department of Education
Office of Postsecondary Readiness

eskolta
We help schools learn.

Table of Contents

Employing Principles of Improvement Science in New York City Schools

Successful education reform requires a process of focused learning. This simple fact is as true for educators as it is for students. The field of improvement science, which focuses on employing the scientific method to test and learn from changes within organizations, has helped inform the thinking of education reformers from John Dewey to Alfred North Whitehead to Richard DuFour, who have emphasized the value of disciplined, scientific investigation into education practice in order to improve education practice. Indeed, the scientific method is often employed in large-scale, multiyear, multimillion-dollar education evaluations; however, using these same methods in individual classrooms with individual students is far rarer.

This gap, between the opportunity to test out ideas with the rigor of the scientific method at the large scale without a similar opportunity on the small scale, has hurt our schools and our children by leaving practitioners without a clear way to identify, develop, and scale the practices that are best for their students. Tests, curricula, and activities standardized for distribution across entire districts or states too often fail to achieve intended impact at the local level. Education is too often treated as a field in which the next great idea will be the solution, when in fact what educators need is a process to test the many great ideas that are already available and implement the right ones in the right ways to work for their students. Today, important efforts are under way to make this happen.

Improvement science and the Carnegie Foundation for the Advancement of Teaching

Most recently, a new understanding of improvement science in education has been brought to the fore by education researcher Anthony Bryk, president of the Carnegie Foundation for the Advancement of Teaching (CFAT). Bryk advocates the use of scientific principles by education practitioners through small, measured experiments integrated into their everyday work. This white paper highlights concrete examples of how schools in New York City have done this, providing a new way forward for testing, refining, and scaling practices that offer the most promise for our students.

In their 2015 book, Learning to Improve, Bryk and his colleagues Louis Gomez, Alicia Grunow, and Paul LeMahiew set out a vision for education reform hewing to six core principles of improvement science:

1. **Make the work problem-specific and user-centered.** By clearly defining the users and the problem, we arrive at better solutions.
2. **Address variation in performance.** By looking for places where results differ, we can find opportunities to learn or improve.
3. **Adopt a "systems view."** When seeking to understand how to address one problem, we need to grasp the complexity of factors that are affecting that problem.
4. **Measure to improve at scale.** When engaging in change efforts, identifying manageable measures helps us to identify progress.
5. **Anchor practice improvement in disciplined inquiry.** By engaging in thoughtful but structured reflection on action, practitioners are able to learn from their experiences and data.
6. **Accelerate improvement through networked communities.** Bringing together practitioners who are working to solve the same problem enables sharing of resources, ideas, and learning.

Bryk and his colleagues propose that by following these principles, educators can more effectively shift from acting solely as practitioners, who are using practices that have been handed down to them by their colleagues, to acting as practitioner-researchers, who are simultaneously testing and improving upon these practices through their own insights, ingenuity, and scientific experimentation.

Eskolta, OPSR, and the work of school improvement

Bryk is among key luminaries whose thinking have informed the work of Eskolta School Research and Design, a nonprofit education reform organization based in New York City. Eskolta was a founding member of CFAT's Student Agency Improvement Community (SAIC), a national network of organizations and school districts using improvement science principles to help students improve persistence in learning. Since 2012, Eskolta has worked in partnership with the New York City Department of Education's Office of Postsecondary Readiness (OPSR) employing principles of improvement science in more than 50 education improvement projects in individual New York City schools.

These joint improvement projects have occurred primarily through two initiatives that began in the 2012–13 school year. Through the Transfer School Common Core Institute (TSCCI), twenty transfer schools—high schools serving overage, under-credited students who have previously attended another high school and fallen behind in their learning—have engaged in multiple change projects over three years. Through the Academic Behaviors Pilot (ABP), fourteen middle and high schools have engaged in change projects to build student persistence in learning.

In its assessment of these projects, Eskolta has found that more than 70 percent of the time, new practices brought into place through these projects remain in place in schools one, two, and three years after they were first introduced. This rate of sustainability is more than twice the estimated average of 30 percent for typical efforts to improve practice in organizations[1].

School principals who have been involved in these projects attribute dramatic progress in their schools, ranging from improvements in student test scores to changes in how educators think about differentiation and learning, to this work.

This white paper shares examples in which change efforts were particularly aligned to principles of improvement science at ten schools:

- **Bronx Leadership Academy II**, a public high school in the South Bronx that enrolls nearly 500 students, 90 percent of whom are eligible for free or reduced-price lunch. BLA II participated in ABP and SAIC for two years.
- **Brooklyn Frontiers High School**, a transfer school in downtown Brooklyn that enrolls 180 students, four-fifths of whom are eligible for free or reduced-price lunch. Brooklyn Frontiers participated in TSCCI for two years.
- **City-As-School**, a transfer school in Manhattan that enrolls 650 students, three-fifths of whom are eligible for free or reduced-price lunch. CAS participated in TSCCI for two years.
- **Coalition School for Social Change**, a public high school in Manhattan that enrolls 290 students, all of whom are eligible for free or reduced-price lunch. Coalition participated in ABP and SAIC for two years.
- **I.S. 229 Roland Patterson**, a middle school in the Bronx that enrolls approximately 250 students, about four-fifths of whom are eligible for free or reduced-price lunch. I.S. 229 participated in ABP for two years.
- **Jill Chaifetz Transfer High School**, a transfer school in the Bronx that enrolls 190 students, all of whom are eligible for free or reduced-price lunch. Jill Chaifetz participated in TSCCI for two years.
- **P.S. / I.S. 266**, an elementary and middle school in Queens that enrolls 690 students, half of whom are eligible for free or reduced-price lunch. I.S. 266 participated in ABP for two years.
- **North Queens Community High School**, a transfer school that enrolls approximately 200 students, about three-fifths of whom are eligible for free or reduced-price lunch. NQCHS participated in TSCCI for three years, ABP for one year, and SAIC for two years.
- **South Brooklyn Community High School**, a transfer school in South Brooklyn that enrolls 160 students, four-fifths of whom are eligible for free or reduced-price lunch. South Brooklyn participated in TSCCI for two years.
- **West Brooklyn Community High School**, a transfer school in Brooklyn that enrolls 200 students, nearly three-fourths of whom are eligible for free or reduced-price lunch. West Brooklyn participated in TSCCI for two years.

1. Kotter, J., "Leading Change: Why transformation efforts fail," Harvard Business Review, 1995

Prologue: Organizational Readiness before the Work Begins

The examples provided throughout this white paper highlight how schools have brought the principles of improvement science to life in practice. However, before such work can even begin, a few key structures need to be in place. While this white paper does not explore how to achieve this, it is important to note that significant effort goes into the early establishment of such structures in schools. These structures include:

Dedicated leadership. The principal of any participating school must actively identify the place of the change effort in his or her own instructional priorities for the school, developing goals that align with the effort and showing support for the effort by providing active support to those involved in the change effort. In addition, the principal must be involved in selecting a team and scheduling time and commit to changing school policy the following year based on the findings of the team.

Dedicated team. A group of two to six educators, ideally working with the same set of students, must be ready to actively engage in improvement science to rethink and redesign practices in their school. This design team should include some who are potential leaders on faculty, paying particular attention to who generally takes on new initiatives while still being respected by colleagues as someone from whom they can learn. In the projects described here, Eskolta staff worked with principals to identify such a team.

Dedicated time. The design team ideally has at least 60 minutes a week when they are regularly meeting with one another in order to manage and reflect on their shared work in implementing improvement science projects. In addition, the full school has at least two meeting times, ideally one in the winter and one near the end of the year, when the principal has empowered the design team to share learning from their experiments in order to inform school policy changes and scaling of learning.

Openness to professional learning. Finally, school leadership and team members need to enter this work with an understanding that they are not getting an off-the-shelf package for their students but are instead being treated as professionals who can test and improve upon practices. A belief that schools are places of adult learning and that school cultures must foster improvement is key to educators being able to engage in the work effectively.

1 | Make the Work Problem-Specific and User-Centered

"Being user-centered means respecting the people who actually do the work…" (p. 32)

To improve practice requires first knowing the purpose for the practice itself. Large-scale studies of practice can be hard to implement at the classroom level because of the many nuanced differences between schools, communities, and students. For this reason, educators who wish to use methods of improvement science do well to start with the question: "What specifically is the problem we are trying to solve at our school?"

Indeed, the first core principles of improvement science ask educators to be problem-specific and user-centered, paying close attention to exactly what is being solved and for whom. In so doing, one adopts a "co-development orientation" in which not only the educators but also their colleagues, their students, families, and other community members are engaged in understanding the problem. As Bryk et al. explain, this enables projects to focus on real, felt needs and thereby yield meaningful results while also helping everyone to "own the outcomes of their efforts" (p. 34).

While this principle—that change succeeds if people have taken the time to deeply understand the problem—may seem obvious in concept, in practice it can be difficult. Educators have often spent much time already contemplating ways to address problems they experience in classrooms and schools; school leaders often have their own visions of what they can accomplish. There may be an understandable desire to get to solutions without dwelling on investigating the problem. However, the very act of consulting students, teachers, or parents, and of seeking to better understand a problem, can help everyone to become part of the solution.

This section outlines two approaches that Eskolta has used effectively with school staff to engage in problem-specific and user-centered efforts: by using surveys of students to understand user perspectives, and by using rubric development to specify the problem being tackled.

Survey students to understand user perspectives

Once a team has identified an aim that they are trying to tackle, there is often strong impetus to move forward without deeply understanding the perspective of the students themselves—the "users." By conducting surveys or interviews of students, educators can address their own assumptions and deepen understanding of the users and the problems they face before they move from goal to idea. Surveys for this purpose do not need to meet every criterion for statistical validity, but designing a good survey does involve careful work to develop and test out questions that will yield thoughtful answers to a specific research question and hypothesis. Eskolta has drawn upon the work of the national SAIC network to use survey items that have been validated through national research.

One example of how educators used survey data to understand a problem comes from City-As-School. At this transfer school, students participate in four internship placements each year intended to serve as a central element to learning. The improvement science project began when the school identified a problem from their existing data: some students were not completing internships, while others were completing them

but not receiving credit. While they had many ideas of what lay at the root of this problem, they did not jump directly into addressing these ideas but instead sought to better understand their users first. Working with Eskolta, the team developed a survey of student perceptions of the internship experience. This work included:

- bringing together internship coordinators at the school to brainstorm possible causes of the problem;
- using this to generate a broad research question and sets of more detailed questions regarding student interest, attendance, personal relationships with site staff, and the school's internship seminar class;
- testing the questions with a few students before broader use, in order to ensure students understood the questions and that the responses yielded useful information;
- collecting surveys from dozens of students who were involved in internships, with clear framing that this would only be used to try to improve internships; and
- investigating patterns in responses and subgroups within the full sample.

The surveys helped the team assemble a set of characteristics of "excellent internships," which then informed further thinking and testing as they sought to convey these characteristics to internship mentors through introductory materials and practices for engaging with students in ongoing assessments of their internship experiences. As they did this, they moved into work that drew upon disciplined inquiry (core principle 5) and efforts to learn from stronger internships in order to address variation in the system (core principle 2).

Develop a rubric by drawing upon student artifacts

Many change efforts in schools begin with an understanding that there is a particular skill or mindset that educators hope to help students better develop. Such a need may be identified through educators' experience with their students: for example, they have seen students struggling to write persuasive essays and want to help them do this better. Often when such problems are identified, however, they have not reached a high level of specificity. Better writing, more persuasive writing, or greater persistence are all goals for improved student skills, but none have a high enough degree of specificity to be actionable for teachers. In fact, when focusing on skills and mindsets, Eskolta staff have often found it valuable to help educators "drill down" into a more specific definition of the problem by working together on developing rubrics to define the issue. In this way, rubrics serve not just as a tool for educators to later assess work but also as a tool for educators to refine their own definition and understanding of what they are assessing in the first place.

For example, at Jill Chaifetz Transfer High School a team focusing on development of students' reading skills sought to better define the specific problem they were addressing. To do this, they began work on a rubric drawing on the Common Core Learning Standards. The team:

- chose three skills based on the standards that clearly connected to the problem of reading comprehension;
- reviewed student work from prior assessments in order to break each apart into the subskills a student would need in order to be able to master the skill. This step was critical in development, as it helped educators to specify the problem in relation to students—their key "users";
- identified and discussed instances in which the team had been making assumptions about what students already knew or were able to do in relation to these subskills; and
- used this to revise how they were assessing skills and sequence steps in the rubric.

Thus, by developing a rubric together based on existing research and practice, the team was able to more deeply specify a problem they were seeking to solve. The team then used this draft language to begin looking at samples of student work and used scores on the rubric to anchor the work in small measures (core principle 4) as they engaged in disciplined inquiry (core principle 5) around the development of these skills.

2 | Address Variation in Performance in Order to Improve Implementation

"Understanding the sources of variation in outcomes, and responding effectively to them, lies at the heart of improvement…" (p. 35)

The second core principle of improvement science addresses the importance of being able to replicate results under diverse conditions. Quality improvement requires paying attention to instances in which the same process is yielding different results. In some cases, these results are a sign that something is working well but only in some places; in other cases, these are a sign of an unsolved problem. In either case, reducing variation in performance is at the heart of improvement: educators can get better at what they do by learning from the isolated success or by figuring out how to address the isolated failures.

Too often, however, such variation is missed because when looking at data across a whole school, staff and school leaders have only enough information to see broad trends. For instance, they might know that the passing rate for math exams was 57 percent and lay out a goal of increasing that passing rate to 62 percent. Identifying variation requires a second intermediate step, which comes after identifying an overall trend but before seeking a solution. Educators need to change their question from "What is not working?" to instead "For whom is it working or not?" and "In what instances is it working or not?" It is important to note that such variation may relate to variations in process—for example, in one school Eskolta worked with, all teachers were asking students to revise work, but the different ways that they requested revisions were yielding variation—as well as variations in outcomes.

In this section, we outline three approaches that Eskolta has used to help school staff exploring variation in order to work on reducing it: by exploring variation in results when two results that are expected to go together do not, by exploring variation when the same student is getting different results in different settings, and by exploring variation in process by looking for educators who are getting better results doing the same thing as their colleagues.

Use a three-by-three matrix to identify where items that should go together do not

Exploring variation can be a valuable step when a project is first getting under way and educators are trying to identify an area of focus. This was the case at North Queens Community High School, where leadership reviewed existing data at the outset of a change project to find variation. In particular, they sought to find examples where data that they would expect to go together in fact did not. To look for this variation in performance at North Queens Community High School (NQCHS), school leadership gathered data on each student's grades and attendance—two areas that one would expect to go together—and made a simple comparison, asking: For which students does high attendance not connect to strong grades? By looking at these students in particular, teachers could then ask what needed to change for them in order to improve results.

At other schools, this data has been turned into a visual representation. To do this:

- School leadership gather data for every student on two related data points, often either grades and attendance, or grades and test scores.
- Cutoffs are set for what counts as high, middle, and low in each of the data points. For example, at a transfer school like NQCHS, high, middle, and low attendance might be set at 85 to 100 percent, 65 to 84 percent, and below 65 percent. Similarly, cutoffs are set for high, middle, and low grade point averages or passing rates.
- Students are divided into categories: those with high attendance, middle attendance, and low attendance. The same students are then divided into categories of those with high passing rates, middle passing rates, and low passing rates.
- By plotting these in a three-by-three matrix, educators can then examine more closely the list of students for whom unexpected correlations arise: students with low grades despite high attendance, or with high grades despite low attendance.
- Educators review the lists of individual students in these "unexpected correlation" categories to investigate variations that can lead to deeper insights.

Looking at a list of students for whom attendance and grades were not connecting, educators at NQCHS asked themselves what they knew about these students and what the source of variation may be. This helped them to then move into steps to engage in disciplined inquiry (core principle 5) to test approaches to supporting these students. In other cases, such initial research can lead into a deeper exploration to understand the system around these students (core principle 3) to better understand what potential sources of variation are.

Use outlier analysis to identify how student performance varies from class to class

When a project is already under way, the search for variation continues to be of paramount importance. It is often the case that teachers are addressing reading, writing, or thinking skills that are interdisciplinary in nature but that appear differently in different classrooms. This variation can be a valuable source of learning during a change effort. For example, a school that seeks to improve students' ability to write a clear informational paragraph may find that some students get better marks on this skill in one class and others in another. By carefully examining such variation, educators can glean insights into ways that differences between classes can be reduced to help all students reach higher levels of achievement.

Such an analysis was part of the background work done at West Brooklyn Community High School. At this school, a team had been working on the question of how to improve student writing and had deployed a rubric to score student work across classes. The following process enabled deeper reflection about variation in teachers' practice, teachers' understanding of student skill level, and student results:

- Across the school, one assessment in every class was graded using a rubric focusing on how well students wrote using evidence from a text.
- Teachers were provided with baseline data on their own students to look for trends and patterns, but they were also, more importantly, provided with a set of outlier analyses, showing which students had a rubric score in one class that was noticeably different from their rubric score in all the rest of their classes. In some cases, for instance, a student who was receiving marks of a zero on a 4-point rubric in most classes received a 2 in one class. In another case, a student who was receiving a 2 in most classes received a zero in one class.
- The specific instances of outliers were provided with the name of the student and their scores in all their classes, highlighting the outlier class. Researchers were careful to find examples from different classes in order to avoid targeting one class too heavily.

- Each such example then became a source of discussion for teachers to learn from one another's experience about the possible reasons for the variation in results. This raised questions of whether the variation was caused by differences in teachers' expectations, in their instructional practices, or in the design of assessments. Each of the three possibilities then led to deeper reflection on how to reduce variation.

As educators compared results and discussed what was happening differently for the same students, they were able to arrive at new insights about their own teaching and about the support students were receiving. This informed the process of disciplined inquiry that participants were engaging in (as discussed in core principle 5) and helped them to raise new questions about how they were teaching writing and about how they were assessing students' writing skills.

Baseline Assessments

Student	Course	Score
Bryce	ELA 6	0
Bryce	Living Env. 3	2
Bryce	US Hist. 2	0
Garcia	ELA 4	2
Garcia	Living Env. 3	0
Garcia	Lab	0
Saboy	ELA 5	2
Saboy	Global 3	0
Saboy	Living Env. 3	2
Saboy	Lab	2

Look for positive outliers to emulate

Variation can also be sought, and reduced, when examining how different educators engage in the same practice across different contexts. In school reform today, this is often done with a focus on the negative: finding poor practices that need to be improved. Instead, collecting data from different teachers who are engaged in the same practice across a school or across different schools can yield insights about who is doing something especially well to help others learn from them. This can best be done after a change idea has been well under way and a process of using small measures (core principle 4) alongside a process of disciplined inquiry (core principle 5) have yielded insights about what is working, and can then be integrated into learning across a network to accelerate learning (core principle 6).

For example, across ten schools that had worked with Eskolta and OPSR on the problem of improving students' productive persistence for two years, results collected through measures of student growth stood out for one teacher at I.S. 229. The team actively sought to learn from this teacher and codify the variations in her practice for others. This was done by:

- Administering a brief survey to students in all classrooms that were working on the problem of developing persistence, in order to provide baseline data.
- Repeating the survey at the end of the year and looking for changes from individual students in the program. This review revealed the outlier classroom. That is, far more students in that classroom showed improvement in their mindsets from the beginning to end of the year than in any other classrooms.
- Engaging the teacher in reflection on her own process and practice and how these may have led to results. In this case, the teacher focused in particular on the fact that earlier in the year she had introduced a new six-week lesson to emphasize persistence through growth mindset.
- Comparing the teacher's reflection on practice to practices observed at other sites. Observers across multiple classrooms were able to identify this six-week lesson, both in its structure and implementation, as a key difference from other classrooms in schools with similar cultures that were not achieving the same impact.
- Observers proceeded to interview the teacher in greater depth to codify elements of the practice that could be shared with others.

In this way, variation in the system was found, reflected on, and codified in order to share with others. The example from I.S. 229 could then be integrated into processes to share learning in the networked improvement community that formed across participating schools, as described in core principle 6.

3 | Adopt a "Systems View" to Understand and Address Complexity

"It is hard to improve outcomes when we do not fully appreciate how our educational systems operate to produce the results we currently observe." (p. 84)

The third core principle of improvement science directs attention to understanding systems as well as organizing efforts to improve them. As Bryk and his colleagues explain, and as Eskolta's work with schools has shown, it can be hard to improve outcomes when we do not understand how the system operates. As multiple processes interact in classrooms, schools, and districts, taking a "systems view" can help to call attention to the complexity of the causes of individual students' successes or failures.

While educators often discuss and think about the "system" as a whole, taking the time and space to step back and consider the interactions occurring in the whole system is in fact somewhat rare. Often, there is a feeling of helplessness in the face of factors that feel beyond one's control, or there is a well-intentioned effort by educators to focus only on the individual students they see in front of them. On the one hand, it can become counterproductive for educators to spend excessive time considering the many institutions at play in a students' complex lives if this thinking does not lead them to actionable change to help their students. On the other hand, adding some complexity to the view that one gets in a single classroom can dramatically deepen insights into possibilities for improvement.

In this section, we describe two tools that Eskolta and OPSR have used specifically with practitioners in schools to understand the systems at play within their schools: causal analysis as a tool to understand the set of systems that are interacting with a problem, and deep case study as a tool to understand the set of systems that are interacting with a student.

Use causal analysis to understand why the problem is occurring

Causal analysis can be useful for considering the intersecting factors that contribute to a problem as educators are first delving into a project, or as they are discovering new problems that they hope to solve. It can be particularly useful for teachers who are trying to determine why students are not doing well. To develop a causal analysis, a group of educators begins with a problem presented by the data. For example, students might have a low passing rate on the math Regents exam, or students might not complete homework. The group then asks two questions: Why is this occurring? and What is within our control to influence? By carefully asking these questions in a way that is open to new insights and that avoids judgment, educators develop a deeper understanding of the problem.

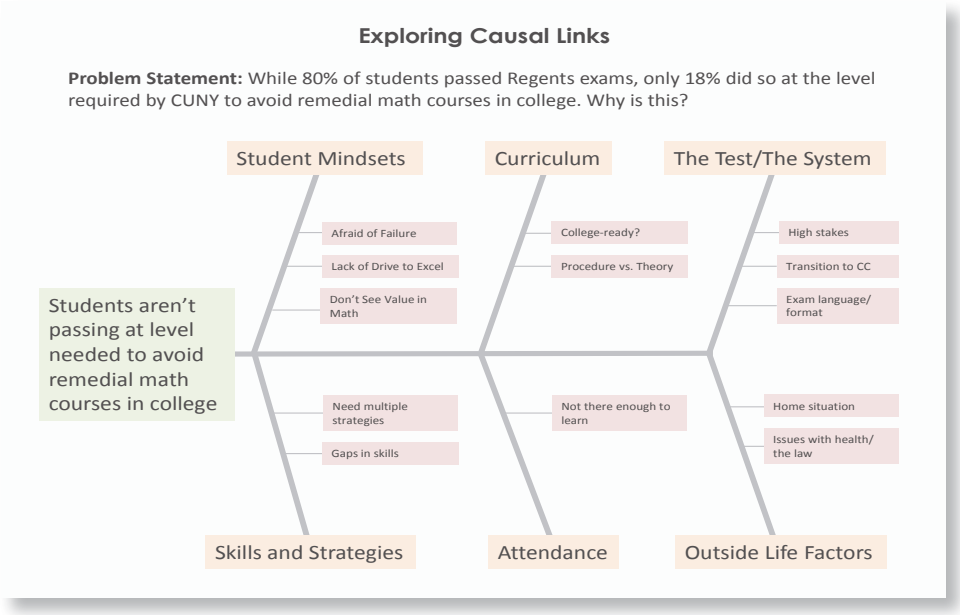
As a part of identifying an approach for their work, three teachers at Bronx Leadership Academy II engaged in a series of activities to learn more about the problem they were trying to solve. The team's causal analysis followed these steps:

- First, the team articulated a problem based on data they had available. Specifically, they saw that students were not passing state exams at a level needed to avoid remedial math courses in college.
- They then considered what students were experiencing inside and outside their school and used this to

consider potential causes of the problem. To do this, they repeatedly asked "why" a problem existed, not being content with the initial explanation and thereby forcing deeper understanding by repeating the question.

- As they did this, they made an effort to actively see each problem from the student's point of view in order to avoid simply blaming the student. In doing this, they arrived at factors like: students are afraid to fail, the curriculum is not sufficiently rigorous, the exam format is unfamiliar to students, students do not see the value in math, and so on. They also sought to identify those that were most actionable for the team.
- The group reviewed the areas they had identified and categorized them into similar topics: several items related to attendance, others related to the curriculum, others to student mindsets.

In this way, the group arrived at a map of the system of factors influencing student success on the math exam. This visual helped them organize their thinking and provided a document that they could refer to throughout their project. From there, they selected specific aspects of student mindsets to tackle in their first efforts, engaging in a process of disciplined inquiry (core principle 5) as they collected small measures (core principle 4) to understand whether and how their efforts were having an effect on students.



Use a deep case study to witness students interacting with the system

A deep case study is a useful tool for thinking about the system as comprising all the places and interactions in a school building. It can be particularly useful when teachers have reviewed a list of students based on variations in data (see core principle 2) or on user interviews (see core principle 1), and out of this realize that they need to investigate their own assumptions about the actual experiences of their students and the systems affecting them.

One of the best ways for educators to do this is by closely studying a handful of students over the course of an entire school day. In order to do this, a small group of staff and faculty pick three students who are grappling with a particular problem that the school is trying to address. The group then creates a schedule for the participating staff and faculty to observe each of the three students for one entire school day. Group members take notes throughout the observation process that focus exclusively on student behavior and the surrounding environment. These notes are then shared, as the educators reconstruct the system of influences that are affecting the student throughout the day. In some cases, a study such as this can then be turned into a visual that shows these relationships through a map drawing interconnecting processes or a spatial map placing elements of the system in different locations within and outside the school building.

Brooklyn Frontiers Low-Inference Observation Sheet

Student: _____ Class: _____ Pd.: ____ Date: _____

Time	Student is saying or doing…	Teacher is saying or doing…	The rest of the environment…
10:03-10:05	Looks at problem set. Twirls pencil. Raises hand. Turns to Carlos and says, "This is stupid." Raises hand, then puts it down as T. looks up.	T. circulates room talking to students individually on problem set, looking up from time to time.	Most students are focused on problem set. 3 or 4 on other end of room are talking and looking at Carlos.

Artifacts to Collect:	Questions/Inferences that Come to Mind:

Observed by: _____ Page #___

One example of such an effort comes from Eskolta's work with Brooklyn Frontiers High School. At this school, the full staff engaged in a study to better understand the system affecting struggling students:

- A group of teachers identified five students in the school who had consistently low passing rates across multiple classes.
- The principal then arranged coverage for teachers to visit one another's classes and for counselors to interview students at different times of day.
- Using this time, the staff arranged a schedule for each of these five students to be observed throughout every period of the day, both inside and outside classrooms. The observers took notes in which they wrote exactly what they observed the student doing and what they observed the teacher and other students doing in the room at the same time.
- The group convened weekly and shared observations of individual students as they were being completed. This venue provided an opportunity to paint a picture of the different places in which the student interacted: in different classrooms, in hallways, outside the school building, and so on.

At Brooklyn Frontiers, this series of observations yielded a deeper understanding of the system around the student and led the group to identify the idea of introducing checklists for struggling students listing the detailed tasks to prepare for and engage in class at each period. Staff members proceeded to engage in a series of tests, using the checklist with individual students and reflecting on the work. This process is discussed further in sections on disciplined inquiry (core principle 5) and the small measures that help guide the work (core principle 4).

4 | Measure to Improve at Scale

"Since the intent of the data collection is to advance continuous improvement, data need to be collected frequently to identify opportunities for change and to assess whether positive changes are in fact occurring…" (p. 100)

The fourth core principle of improvement science addresses the importance of identifying manageable measures to track changes. Bryk and his colleagues argue that if something is not measured, it is impossible to improve it at scale. As Bryk and his colleagues write, "common data are the key to sensing overall system performance, identifying sources of variation in outcomes, and continuing the processes of learning to improve" (p. 111).

While the last two decades of standards-based reform have left no paucity of measures in schools, one of the most common mistakes in collecting measures is that of focusing exclusively on data that already exist. Grades, attendance, and test scores are regularly collected in New York City schools, but they are poor measures for an improvement effort. Improvement efforts succeed when teachers are able to take thoughtful risks, trying something different or new in order to learn from this experiment, and then have the opportunity to quickly learn from the risk they have taken. But data that are gathered by most school districts are items of high-stakes accountability that bring with them so much pressure to demonstrate good results that teachers are often unwilling to take the risks necessary to improve their own practice. In addition, most of these available measures are collected only a few times a year—hardly frequently enough for educators to make quick changes and adjustments that allow them to improve in a reasonable amount of time. Smaller, lower-stakes measures can enable educators to be more comfortable taking risks and learn more frequently from those risks so they can get better faster. However, finding smaller measures than these can also be taken too far. If educators are asked to start tracking multiple measures for every student every day, the very act of making check marks on a clipboard can distract from quality education, and educators may find their practice worsens because of the effort instead put into collecting data.

In this section, we describe two data-collection strategies that Eskolta has engaged in with schools to avoid such pitfalls: by identifying small measures that are carefully integrated into educators' existing practice, and by establishing small samples of students so that additional work is required for this group only.

Collect small measures embedded in the improvement effort

One challenge in improvement work is that of selecting a metric to measure impact without adding yet another task for educators to do. One way to skirt this problem is to carefully identify measures that are embedded in student work or in activities that already would be occurring frequently even if teachers were not asked to collect the data. Typically, these measures can be identified after a team has first identified the problem they seek to solve and spent enough time studying that problem and the systems around it that they have a clear idea of the changes they will first try and what they expect those changes to achieve.

An example comes from North Queens Community High School (NQCHS), where the math department had set a goal of establishing classroom structures to build student persistence in the face of failure. The team was then forced to consider how they would measure whether their change ideas were having an impact:

- Because the math department gave quizzes regularly throughout the week, they chose to use students' work on quizzes as a source of frequent measures.
- Students were given the option of submitting revisions after their quizzes were graded. To assess students' persistence, teachers counted the number of students who opted to revise their quizzes after receiving feedback.
- Questions on quizzes were arranged in increasing difficulty. To assess whether students persisted even as work became challenging, teachers counted how many questions students attempted the first time they took a quiz.

In this way, the team at NQCHS established simple measures that required only counting items that were already evident through their practice. These measures could then anchor discussion that occurred through the process of disciplined inquiry (core principle 5) and enabled the team to see small variations in results (core principle 2) across different efforts.

Establish small samples to reduce strain on data collection

Once educators have identified a problem they seek to address for their students, they often find it difficult to arrive at measures that are easy to collect. This can be paralyzing, and educators do not move forward with anything because they cannot arrive at data to collect. In such cases, it is more important to find a way to move forward than to wait until systems are in place. It can be most beneficial to identify a small sample of carefully selected students and collect data only with these students. At times, Eskolta has asked educators to focus on ten, five, or even just two or three students in order to learn from an initial experiment.

One example comes from Eskolta's work with Bronx Leadership Academy II, where a math teacher was focusing on how to help students improve persistence. He was able to collect measures on a series of improvement efforts by:

- Analyzing the results of a student survey on academic mindsets to identify a subset of students who had notable potential for growth.
- Drawing on his knowledge of the students, their grades, and their behaviors to narrow the group down to four students who appeared to have strong math skills yet did not appear to want to apply them. Refining the research question to match this small group, the teacher focused on how he could help students for whom skills were strong but persistence was low.
- Collecting data on how many questions this small group of students completed on every assignment, how many revisions on every assignment, and the students' own reflection on their growth in writing and conferencing.

This data helped the math teacher to see that by explicitly teaching students that their brains could grow from effort and then providing individualized feedback to students focusing on their effort to take on new problems, he was able to achieve dramatic changes for specific struggling students. This yielded new ideas for practices and lessons he could share with colleagues in the school.

5 | Anchor Improvement in Disciplined Inquiry

"Introducing change is structured so that participants can learn their way into what it will actually take to enact some new practice reliably with quality at scale." (p. 114)

The fifth core principle of improvement science highlights that participants need to engage in a disciplined inquiry process that can assist in carrying out this improvement work over the long haul. Of all the principles of improvement science, this principle is relevant across the largest portion of time that educators spend in the process, as they must return again and again to the question of how they will accomplish what they set out to do.

While educators across New York City are familiar with the term "inquiry" as a practice that has been promoted in City schools for more than a decade, the practice of inquiry takes on many forms in different schools, to different degrees of success in changing practice constructively. One common pitfall in the inquiry process is that educators quickly move from questions to answers: they offer suggestions and ideas without having had opportunity or structure to first question and listen. The opposite pitfall, however, is also common: educators reflect on the possibilities and the outcomes they want to achieve, but there is no disciplined effort to move from these ideas quickly and agilely into action in order to test them.

In this section, we describe disciplined inquiry strategies that Eskolta has engaged in with schools to avoid such pitfalls. These tools help educators to engage in systematic reflection on practice that in turn helps them to improve practice and better serve their students. These include: using a PDSA cycle to help participants move quickly between trying out an idea and reflecting on its results, using a LUCE matrix to consider how participants are getting the opportunity to question and develop their own thinking, and using active note-taking to maintain a record of learning over time.

Use a PDSA cycle with run charts to move from action to reflection

To test strategies in a disciplined way, it can help to actively make a plan and predictions about how a new strategy will help meet a goal and then compare results against those plans and predictions: this pairing of planning and predictions enables educators to articulate their own beliefs and test those against their practices. A helpful tool to do this is a plan-do-study-act (PDSA) protocol, which prompts educators to think through a series of steps whereby they review the aim of their work and arrive at one idea of a small change that they want to test out, make a prediction about what will happen, identify small measures to assess success, compare what they tried to the prediction, then repeat.

With such plans, it is key that the team is using measures that can be gauged within days after they are made. This often requires that educators move away from longer-term aspirations for students to improve grades in classes and on to much shorter-term aspirations—for example, that specific students will participate in class tomorrow or that specific students will complete assignments that they had not completed last week (also see core principle 4). It is also important that predictions address what educators hope will happen as well as what they fear will happen—for example, a new strategy is being tested with the hope that it will yield

greater participation but the fear is that it will not. By articulating both, educators have greater latitude to reflect afterward on what actually did occur.

An example of effective use of the PDSA tool comes from Coalition School for Social Change, where a group of math and social studies teachers wanted students to take on more challenging work on their own, without the promise of an external reward. They landed on an initial idea to test: offering students a "challenge question" once a week as an opportunity to push themselves beyond completing only assigned work. In a series of PDSA cycles, the team:

- Came together once a week to articulate predictions about what would happen when they used different methods of encouraging students to take on "challenge questions."
- Over the next few days each week, they tried one new way of encouraging students to try challenge questions in the classroom, collecting data about how many students did.
- Each week, the team counted what percentage of students had chosen to try challenge questions, and charted this on a graph so that they could see from variation in the data from week to week. This run chart helped them to see a simple visual of the impact of their efforts.
- The team also put the names of each student in the class on a grid where they highlighted in a new column each week whether or not the individual student had tried a challenge question. This allowed them to structure their inquiry to focus more deeply the question of how individual students were or were not responding to each intervention.
- At the next week's meeting, the team reviewed their experiences and the data collected, looked for variations in the data (see core principle 2) and used this information to decide whether to continue the idea, adapt it in some way to be tested again, or discard it completely and try something different.

From this series of PDSAs, the teacher team at Coalition discovered that with specific words of encouragement and framing of challenge as a worthy part of the education process, an increasing number of students took on challenge questions.

Draw out insight by drawing upon the LUCE matrix

Active listening is key to disciplined inquiry. As educators are engaging in the process of testing one idea after another, keeping an open mind to the possible successes and failures of those efforts, and being able to turn failures into new learning, depends on listening effectively. Too often, those charged with facilitating improvement efforts are not adept at helping educators to gain their own insights through scientific study and instead offer their own solutions or create an environment in which new ideas are not tested.

One approach to help educators engage in active listening is the listen-understand-convince-explain (LUCE) matrix. This matrix helps facilitators to be aware of how they balance the time they spend listening and speaking, as well as the time they spend understanding and convincing. In effective projects, participants spend the most time listening in order to understand participants' thinking and learning rather than trying to convince them based on their own ideas.

One example of the use of these active listening skills arises from Eskolta's work with I.S. 266. At that school, a team of educators engaged in an effort to help students improve persistence in learning by developing students' belief in their own ability to learn. One teacher on the team was seeking to identify the appropriate strategy to test. The teacher was intent on using an inventory of student learning styles as her next step but was also experiencing frustration at the sense that she was not certain this was the right solution. The facilitator:

- Began in a mode of listening to understand. She asked the teacher questions that helped her to clarify her thinking about the activity she was planning with questions such as "Would you tell me about the activity you're doing?" and "Talk it out a little bit more." These questions provided the teacher with room to develop her thinking without feeling that she had to resolve all her questions or that she was being assessed.

	Listen	Explain
Understand	Ask questions to help participants explain their thinking	Make statements to help participants explain their thinking
Convince	Ask questions to help participants learn something new	Make statements to help participants learn something new

- Shifted to a mode of speaking to understand. The facilitator paraphrased the teacher's original statements, using this to help solidify a few succinct points with phrases such as "So what I'm hearing you say is that you tried having students identify their own learning styles and you have found this helps them be aware of the areas where they can persist more." In this way, the facilitator was able to help the educator to connect her own insights to the research at hand.
- Resolved to a mode of listening to convince. Realizing that there were some areas where the teacher's ideas did not appear to connect with the research, the facilitator sought to help the teacher realize this herself. In particular, while cognitive science research highlights the importance of students taking on new challenges, the teacher's proposed use of a learning styles inventory was going to ask students to focus solely on what they perceived as their existing areas of strength. The facilitator asked: "How would I feel if I were a student and I identified as just a visual learner. How would I feel about tasks where I had to use auditory learning?" Such questions, asking the teacher to put herself in the shoes of the learner and to question her own assumptions, helped the teacher to make new connections.

This approach yielded an insight from the teacher who changed her effort to ask students to use learning styles to identify areas where they perceived the possibility for improvement and growth. The result was an effort that built on the teacher's understanding and strength while aligning to the research. The team continued forward with testing this and other ideas in an ongoing inquiry process.

Manage disciplined inquiry through disciplined note-taking

Along with active listening, active note-taking is key to disciplined inquiry. Inquiry often suffers from well-intentioned but unsuccessful efforts at note-taking. Verbatim minutes replete in detail are never read; notes that are too cursory or not taken at all leave participants repeating the same discussion they had at their previous meeting without building upon learning, so that by the time the third or fourth meeting rolls around everyone has tired of the effort. Notes that keep inquiry on track include a few features: clear enumeration of what was predicted and proven (or not) from prior attempts, next steps to push people to the next inquiry discussion, and a connection to original objectives so that participants remain rooted in the overarching goal of their work.

For example, at South Brooklyn Community High School, a transfer school with a large counseling staff to help students reengage in high school, Eskolta worked with a group of counselors engaged in a series of meetings to develop new structures for one-on-one conversations with students about persistence and organization skills. To keep the insights that surfaced over multiple meetings at the forefront:

- Notes were taken at each group discussion, synthesized to fit on one page, and sent back to participants within seven days after each discussion.
- A graphic organizer, brought to each meeting, placed each idea from the prior meeting alongside space for participants to brainstorm new ideas, enabling the team to reference highlights from earlier conversations.
- As participants moved from discussing ideas to making plans for implementation, they wrote notes highlighting specific next steps, tying these to plans that had been outlined in the graphic organizer.

The team was able to use these public notes to develop a deeper understanding of the practices they were trying out and discussing from one meeting to another, and to then see themes that cut across multiple efforts. The team eventually designed materials synthesizing their learning into recommended steps for meetings between students, teachers, and counselors.

6 | Accelerate Improvement through Networked Communities

"Networks organized around common conceptual frameworks, informed by common measures, and embracing the rudiments of disciplined inquiry open extraordinary possibilities for accelerating our learning to improve." (p. 168)

The sixth core principle of improvement science involves leveraging the creativity and commitment of the professionals who are engaged in this work to learn from one another, constituting what Bryk and his colleagues call a networked improvement community (NIC). NICs have four essential characteristics: they are focused on a well-specified common aim, guided by deep understanding of the problem, disciplined by methods of improvement research, and organized to accelerate the spread of good ideas into the field. These networks, Bryk and his colleagues argue, make NICs a "powerful system for accelerating improvement" (p. 144).

When NICs work well, they yield a whole that is greater than the sum of their parts. By bringing educators together to consider common problems and share common strategies for solving those problems, participants gain new insights from one another and increase the chance of achieving success. However, teacher collaboration does not always bear all the characteristics of networked improvement communities. Pitfalls include collaboration taking on the characteristics of a presentation or lecture format that does not value the perspective of multiple participants, unstructured collaboration leaving participants jumping into problem-solving mode without sharing ideas, or group meetings focusing only on administrative matters. Teachers are all too familiar with such pitfalls, recognizing professional development meetings that waste their time rather than develop their practice.

Elements of a NIC as viewed in the Academic Behaviors Pilot

The Academic Behaviors Pilot that Eskolta engaged in with OPSR was designed to comprise many features of a networked improvement community. In the design used in the 2013–14 year of the program, these features could be seen in the following aspects of the program:

- In order to participate, educators and schools had to apply based on their interest in the common aim of developing students' growth mindset by using feedback to help them persist in difficult work. This well-specified aim made it clear to all participating educators that they were seeking to tackle the same challenging problem.
- Educators in the effort met together across schools in seven meetings, one a month throughout the year. In these meetings, educators were exposed to the work of various researchers who were studying the impact of student mindsets on learning. This enabled them to gain a deep understanding of the problem from extant research at the same time as they were studying it in their own schools.
- Participants received coaching support at their own schools to engage in cycles of testing out approaches to introducing topics to students, giving feedback to students, engaging students in self-assessment, and so on. In this way, NIC members were using the tools of improvement science to take ideas from the research and test them in their own school contexts.

- As this work continued, Eskolta gathered survey data from students and kept track of small measures gathered through each project at ten different schools. From this, four case studies and an analytic study highlighting themes and lessons learned across one year of improvement efforts were written. These were published and shared among participants in the Academic Behaviors Pilot as a way to spread learning.

Future Directions: Deepening Improvement Science Work in NYC Schools

The joint work of Eskolta and OPSR to deepen a culture of innovation and learning in New York City schools continues into the 2015–16 school year. In 2015–16, the Academic Behaviors Institute and the Advanced Institute were designed to provide multiple entry points to educators while highlighting lessons learned from three years of improvement work on the aim of improving student persistence in the face of challenging work. The same year, the Transfer School Common Core Institute was designed around four NICs, each of which consisted of one host school that had achieved success through three years of improvement work, with three to four partner schools that had received support to learn from the host school and engage in improvement efforts to tackle the problems addressed by the host school.

The work described here at the city level is also supported through participation of Eskolta and OPSR staff in the national Student Agency Improvement Community convened by the Carnegie Foundation for the Advancement of Teaching. This network brings together practitioners engaging in change at the district level with researchers studying change across various districts nationally. In national convenings, researchers share their latest findings, while local teams translate those findings to the problems of practice they have identified at their own districts. In between, teams across the country working on the challenge of improving student agency share updates on efforts with schools and provide coaching to one another.

By continuing to harness the promise of improvement science, Eskolta and OPSR strive to continue building models for sustainable change in schools. As they do, the two organizations continue to engage in new experiments to determine how best to support educators as agents of change while ensuring that the changes these educators move forward are responsive to the needs of students, sustainable within the unique context of each school, and scalable to be shared with more educators in more contexts. Improvement science brings great promise as a method of recognizing that the greatest challenge in improving our schools is not coming up with the next great idea, but rather helping educators use the scientific method to pick the best out of those ideas, test them out, and make them stick.