
What Does the Most Good...and *for Whom?*

A Guide to Academic Return on Investment Analysis



**District
Management
Group**

What does the most good...and *for whom*?

The goal of academic return on investment (A-ROI) is to use data to analyze which programs and services are producing the best outcomes for students, while maximizing the use of resources.

Districts have increasingly turned to program evaluation to understand which of their programs are effective so that they can do more of what works for students and less of what does not. This frame of thinking is an excellent first step, but it is not sufficient to tackle fully the challenges in districts today. The A-ROI approach includes three key additions to the program evaluation already being conducted in some districts:

1. **A-ROI analysis considers program costs, including the cost of staff time.** A program that leads to a slight improvement in student reading scores might look like a successful investment if the analysis only includes outcomes; the fact that this program costs three times the alternative is information that is not captured in most program evaluations. Analysis of this program's A-ROI would show that this program may not be the most effective use of funds to help students.
2. **A-ROI analysis goes beyond looking at results by demographic groups, and looks more specifically at educational needs.** Since No Child Left Behind, student scores have been broken down by race, socioeconomic status, and special needs status. Grouping students by specific educational needs such as baseline reading ability or academic engagement enables more effective analysis of how well a program is in meeting specific student needs.
3. **The final recommendations from an A-ROI analysis allow districts to strategically target services to students with specific needs.** Oftentimes program evaluation is conducted to determine if a program should be maintained or ended. With an A-ROI analysis that incorporates thoughtful student segmentation, districts can take measures to more accurately target the services to the students who benefit from them most.

The A-ROI toolkit, which follows herein, serves as a guide for the district as it begins developing A-ROI capabilities. Included in this package is the "What Does the Most Good...*for Whom*?" guide and the A-ROI Workbook, which provides illustrative examples of and advice on the more technical steps in the A-ROI process. These tools are meant to support a highly-skilled team in the Office of Evaluation and Planning. Ultimately, it will be the work of this team, not the toolkit itself, that will bring cultural change across the district and embed the A-ROI mindset in decision-making at all levels.



Plan

1. Select Target
2. Define Success
3. Identify Comparison Group

Design Analysis

4. Collect Segment Data
5. Collect Outcomes Data
6. Map the Cost

Evaluate

7. Evaluate Program Effectiveness
8. Analyze Cost-Effectiveness
9. Draw Insight

Act

10. Take Action



Contents

What does the most good...and <i>for whom</i> ?	3
Contents.....	5
Step 1: Select Target.....	7
Step 2: Define Success	15
Step 3: Identify Comparison Group	24
Step 4: Collect Segment Data.....	36
Introduction to Segmentation.....	36
Step 5: Collect Outcomes Data	49
Step 6: Map the Cost	55
Introduction to A-ROI Cost Accounting.....	55
Step 7: Evaluate Program Effectiveness	66
Step 8: Analyze Cost-Effectiveness.....	75
Step 9: Draw Insight	82
Introduction to Drawing Insight with A-ROI.....	82
Step 10: Take Action	90
Appendix A: Toolkit for A-ROI Analysis.....	96
Tool 1.1: Roles of Key Stakeholders	96
Tool 1.2: Program, Strategy, and Effort Organizer	97
Tool 1.3: Program Selection Rubric	98
Tool 1.4: Program Information.....	Excel Workbook
Tool 2.1: Logic Model Template	99
Tool 2.2: Identifying Outcome Measures	100
Tool 2.3: Benchmark for Success Template	101
Tool 3.1: Comparing Forward- and Backward-Looking Analyses	102
Tool 4.1: Possible Key Variables for Segments.....	103
Tool 4.2: Student Segment Organizer.....	105
Tool 4.3: Student Segment and Outcomes Data Collection Template	Excel Workbook
Tool 6.1: Cost Map.....	106
Tool 6.2: Cost Category Outline	Excel Workbook
Tool 6.3: Cost Data Collection Templates	107
Tool 7.1: Segment Identification Process	108
Tool 8.1: Cost Analysis Summary.....	Excel Workbook



Tool 8.2: Cost Breakdown (Ranked) Excel Workbook

Tool 9.1: Draw Insight Summary..... Excel Workbook

Appendix B: Introductions to A-ROI Concepts 109

 Introduction to Segmentation 109

 Introduction to A-ROI Cost Accounting 110

 Introduction to Drawing Insight with A-ROI 112





Step 1: Select Target

Engaging key stakeholders when implementing a new strategy is critical to its success, particularly when it is a strategy to evaluate the effectiveness of initiatives. Imagine a governor who runs a successful campaign based on the premise that she will create jobs in the state. Two years after she takes office and her jobs program is implemented, her office publishes a report that claims the program has been wildly successful. Without any notice or information about the methods used, will key constituents believe the analysis was objective? Given her limited engagement and communication up to this point, the governor might experience significant resistance when trying to generate support for the findings. The same is true for A-ROI analysis. Communicating with stakeholders and inviting key constituents to participate in the process beginning with selecting which program to analyze can help ensure buy-in in the later steps. Start the A-ROI process by communicating effectively and engaging stakeholders to allow for a smooth transition to a new approach to evaluating the cost-effectiveness of programs.

.....

Key Outcomes

When the district has completed Step 1 it should be able to answer:

- Do key stakeholders understand the concept of A-ROI?
- Which program(s) will be the subject of A-ROI analysis?



1a. Introduce district leaders to the concept of A-ROI

Begin the conversation about A-ROI before taking even the first steps towards program analysis.

Before even identifying programs to analyze, district leaders should introduce key players to the concept of A-ROI analysis. In many districts, the idea of A-ROI is unfamiliar and may elicit negative responses. It is critical for key stakeholders to understand how A-ROI will be used to boost student achievement.

All of the key stakeholders who may be impacted by the outcomes of A-ROI or who will be asked to make decisions based on this work should be acclimated to A-ROI. Educating these groups early in the process will make the decision-making later on much easier.

A-ROI should be used and discussed as a tool for improving student outcomes.

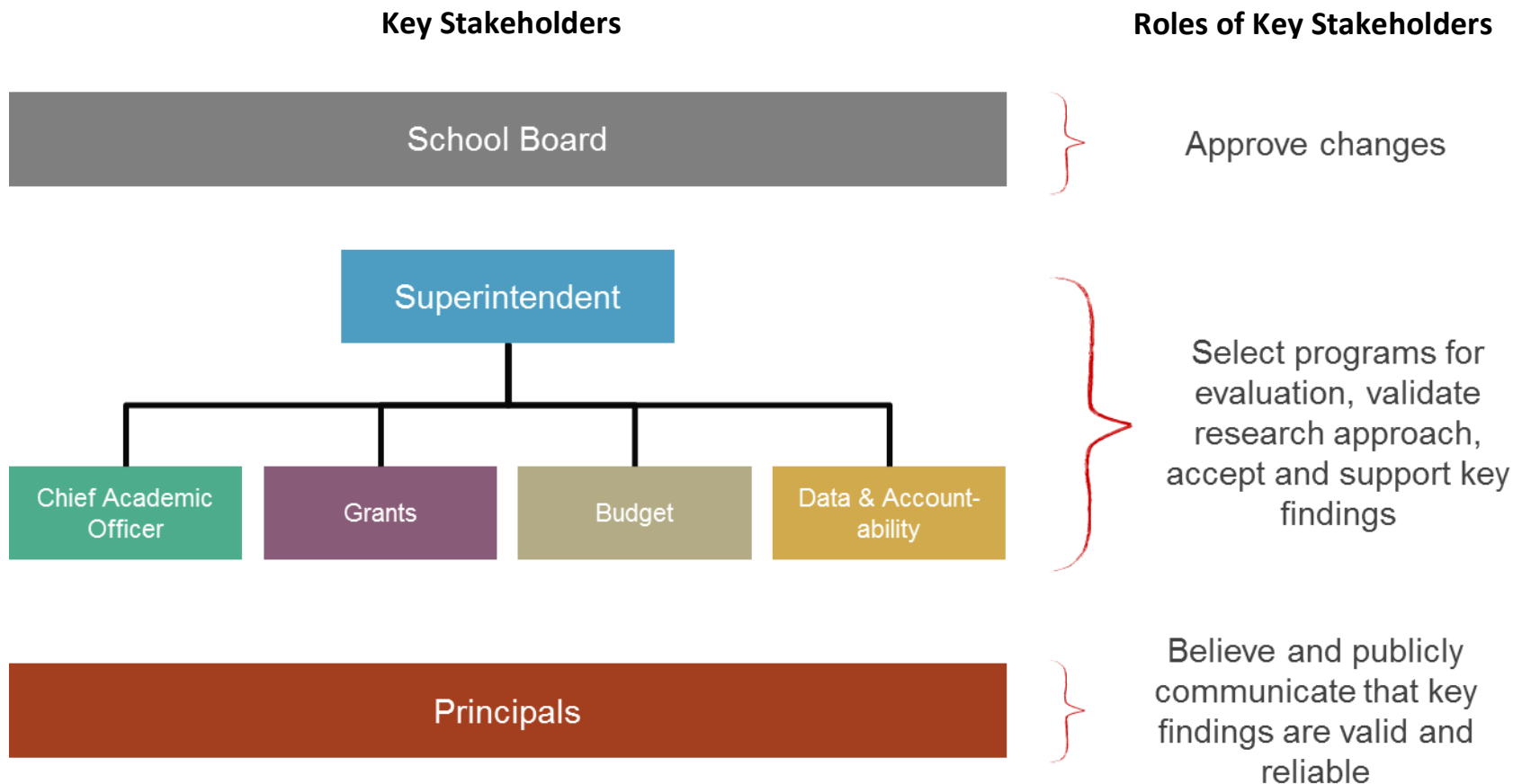
A-ROI is a tool to increase the quality and effectiveness of services to students, and it is important to explain it in this way. If a district emphasizes cost reduction rather than student success, the district runs the risk of alienating many key stakeholders and limiting the potential impact of A-ROI decision-making.

If the concept and rationale for A-ROI is only introduced after a program has been selected for evaluation, stakeholders may see A-ROI as a “cover story” for leadership to cut the chosen program. By bringing principals, district leaders, and board members into the conversation early, it is possible to build understanding for the value of A-ROI thinking.



Tool 1.1: Roles of Key Stakeholders

Communicate with these key stakeholders and clarify the roles they need to fulfill in order to have a successful implementation of an A-ROI approach.



1b. Identify important programs, strategies, or efforts

Create a thorough list of current and upcoming programs, strategies, or efforts as potential candidates for A-ROI analysis.

Districts generally do not have the resources to analyze multiple programs each year. Sifting through all of the programs to determine the best candidates for A-ROI analysis can be a resource-intensive and politically sensitive process if it is not organized in a thoughtful way.

First create a list of all of the district's major initiatives, either current or upcoming. Think broadly about what constitutes a program, strategy, or effort. Some types of programs are easy to identify:

- A purchased reading intervention
- A one-to-one laptop initiative
- A PD series focused on a specific teaching approach

It is also important to include things that are not typically called programs but are important elements of the district's strategy to serve students, for example:

- One-to-one paraprofessional support for special needs students
- Coaching for ELL teachers
- Co-teaching
- Maintaining small class sizes
- Extended day for struggling students
- Programs funded by Title I

DMGroup Tool

See the "Program, Strategy, and Effort Organizer" to guide the process for soliciting input from principals, department heads, and others.

Programs that are currently in place or efforts that are about to start are equally valuable for consideration.

Include multiple departments and schools in the process of creating the list to ensure completeness and accuracy.

Without soliciting input from across departments and schools, it might be difficult to generate a complete list of programs, strategies, and efforts. It may be easier for principals and department heads to create a complete and accurate list if they approach the task one category at a time.

- By type of program (e.g., curriculum initiative, instructional strategy, indirect services, or non-academic programs)
- By level (e.g., elementary, middle, high)



Tool 1.2: Program, Strategy, and Effort Organizer

Instructions: List all major programs currently in use in your school or department as well as any programs that you plan on launching in the near future. Major programs can comprise a variety of items, including, but not limited to, programs aligned with district priorities, programs funded by Title I, programs serving the district's neediest students, and programs that require significant investment of staff time or dollars.

Curriculum Initiatives	Instructional Strategies	Support For Teachers	Non-Academic Programs
<i>e.g., Read 180; elementary world language initiative</i>	<i>e.g., Co-teaching; additional reading block for struggling readers; class size reductions; alternative schools</i>	<i>e.g., Teacher mentorship program; focused PD initiative</i>	<i>e.g., Middle school social worker program; autism inclusion program; parent engagement initiative</i>

1c. Narrow the list using minimum threshold criteria

Setting minimum threshold criteria for size and scope will focus the list on high-potential candidate programs.

A minimum threshold requirement works as a filter to focus on programs of sufficient size and scope, ensuring that there will be significant value to applying an A-ROI analysis to the program.

The threshold criteria should eliminate programs that are not aligned to district strategy, are too politically sensitive, or are too small in scope.

There are a variety of reasons that a program may not be ripe for A-ROI analysis, including:

- It is not a key element of district strategy or theory of action.
- The scope (number of student served) or costs are too small for any changes resulting from the analysis to be impactful. The one exception to this guideline is a small effort that is being considered for expansion. Taking into consideration the future plans allows a district to identify programs that might be too small in scope today, but are exactly the type of programs that would benefit from an A-ROI analysis to help future guide decision-making.
- The program is a political hot topic. Such an effort may not be the best candidate for A-ROI analysis if the political situation will prevent any meaningful changes.

Example

For instance, if the school board recently strongly supported the introduction of a new elementary world language program, it may be very difficult to make changes to the program, regardless of what an A-ROI analysis reveals. As another example, if the district has been debating the value of a penmanship program to teach cursive, this debate centers more on alignment rather than effectiveness. The district should determine whether they still believe that cursive is a skill they believe students should have, rather than conducting an A-ROI analysis to determine if the penmanship program is effective at teaching cursive.



1d. Select target program(s)

Thoughtful program selection is one of the most important steps in the A-ROI process and can be managed effectively by using a structured program selection rubric.

High-potential candidate programs can be prioritized on a more granular level. This process seeks to identify which of the programs, strategies, or efforts has the greatest potential to generate actionable insights and have the most significant impact on students as a result of an A-ROI analysis. The program selection process should consider scope, investment of dollars and staff time, and the political context surrounding the program.

Selecting a program with a good chance of leading to actionable insights and tangible change is critical to gain momentum for the concept of A-ROI in the first few analyses. Ideally, the analysis should focus on a program whose owner wants to use the report to inform future implementation.

DMGroup Tool

See “Tool 1.3: Program Selection Rubric” in the A-ROI Workbook to help create a standardized selection rubric.

It is important to consider how internal and external stakeholders will perceive the program selection methodology.

Debate concerning why a given effort is “targeted” for A-ROI analysis is a distraction from the ultimate goal: to increase the quality of services to students. Selecting a program that has previously been a contentious topic of discussion, whether amongst district employees or in the community at large, creates an opportunity for critics to question the transparency and objectivity of the A-ROI process.

For instance, selecting a program about which the district leadership has been critical and using the A-ROI analysis as a basis for cutting the program will raise concerns about the objectivity of the process, regardless of its fidelity.

Mistakes to avoid

During Program Selection:

- Selecting a program for A-ROI analysis that is not aligned to strategic priorities.
- Overlooking small programs with plans for expansion.
- Creating a list of candidate programs without gathering input from across the district.
- Selecting a program of which the leadership has been critical; this creates the risk of being perceived as using the results to justify cutting the program.

During Communication:

- Communicating or thinking about A-ROI primarily as a cost-saving tool rather than a tool to raise student outcomes.
- Waiting until later in the process to communicate with the School Board and principals as to why A-ROI is an important approach and mindset.

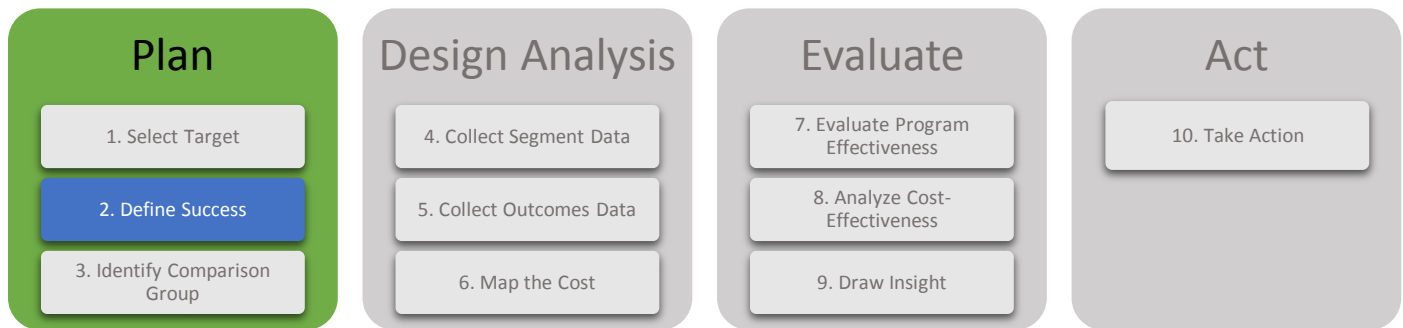


Tool 1.3: Program Selection Rubric

List the high-potential programs, strategies or efforts, then score each of them.

Program or Strategy	Aligned to strategy?	Direct impact on learning?	Large reach or plans for expansion?	Significant investment of staff time?	Significant investment of cash?	Politically feasible to change?	Data available?	Large number of students?	Comparison group?	Uncertain effect?	Total
Example: <i>Technology based math program</i>	X	✓	✓	✓	✓	✓	X	X	✓	✓	7
1.											
2.											
3.											
4.											
5.											





Step 2: Define Success

Having a point of reference is essential for interpreting the outcomes of a program. Upon hearing that a basketball player makes 55% of his shots, one might conclude that he is a mediocre shooter. However, when compared to alternatives, in this case the league average of 45%, then he appears to be much better. Without having the league average as a point of reference, a scout might find it difficult to understand what making 55% of shots attempted actually means about the player's ability. When designing an A-ROI analysis, it is crucial to use benchmarks or comparisons to other students to set expectations about what results will be considered successful. Is a 10% increase in the number of students labeled proficient a success for the new reading program? A 20% increase? A 5% increase? Knowing that the old reading program typically helped 10% of students move to proficiency would be important in analyzing the new reading program's outcomes. Without having a concrete goal for the program's outcomes, it can be difficult to interpret the results, so it is beneficial for the district leadership to define a benchmark that would constitute success for the program at the outset of the A-ROI analysis.

.....

Key Outcomes

- When the district has completed Step 2 it should be able to answer:
 - What is the program designed to accomplish?
 - What metrics can measure these outcomes?
 - What level of change constitutes success?

2a. Clearly articulate the program's objectives and describe how the program functions to achieve the desired objectives

Explicitly describe the intended outcomes of the program—and how it achieves them—to ensure a clear link between the analysis and the intended outcomes.

The outcomes the program is intended to achieve and the methods used to achieve those outcomes must be clearly understood. One way to achieve clarity around this is to outline a logic model for the program. A logic model is a tool to map out the direct link between the inputs, the actions, and the intended outputs of a program. A clear and comprehensive logic model helps identify which aspects of a program are successful and which might need improvement.

The stakeholders most invested in the program should be consulted to understand how, in their mind, the program functions and how it creates the intended benefits.

Example

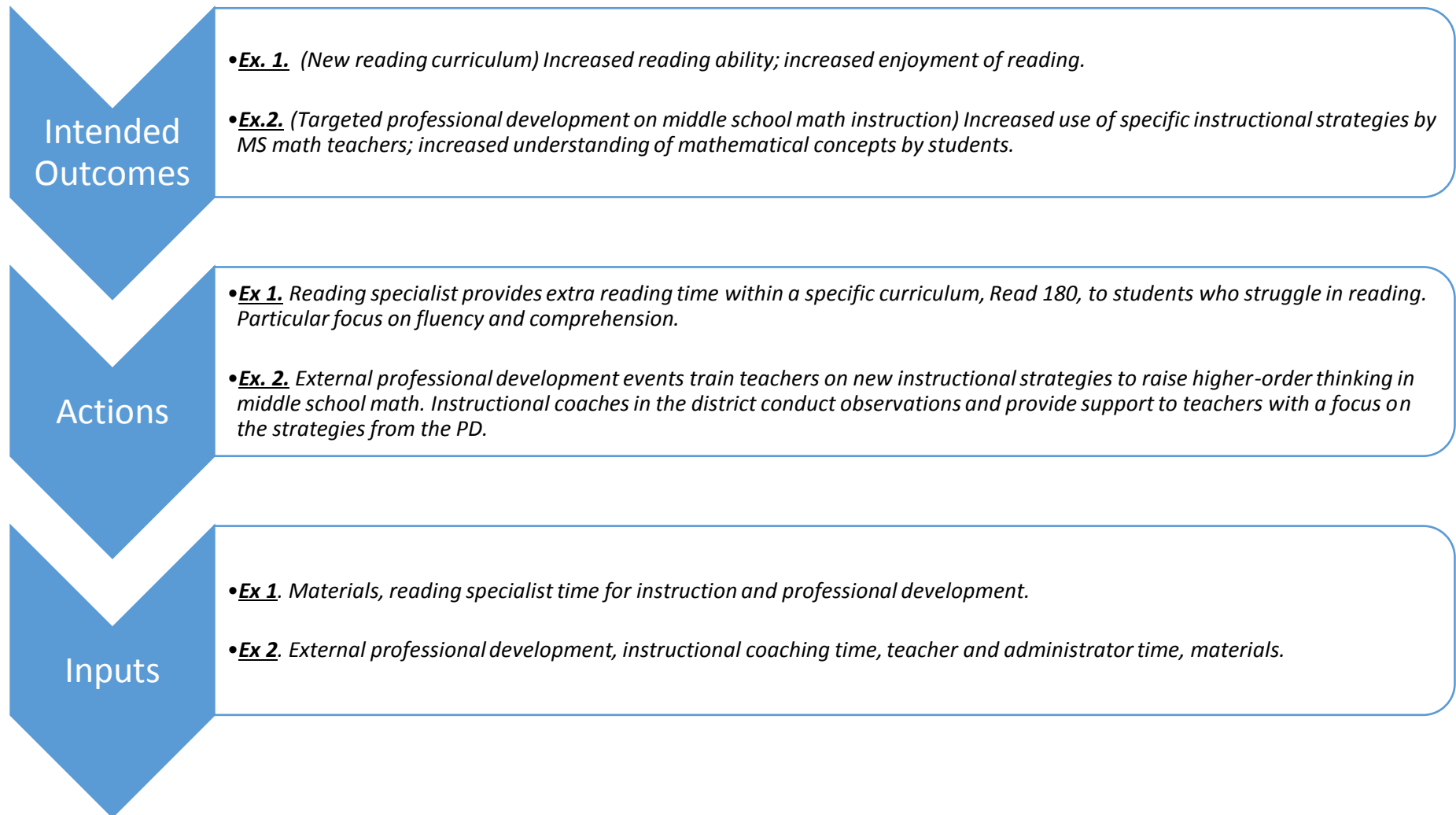
For instance, a district that analyzes a reading program may determine that both reading ability and the student's enjoyment of reading are intended outcomes. In this case, different types of data will need to be collected and different anticipated gains articulated to measure each intended outcome of the program.

Once the outcomes are identified, the next step is to describe a clear connection with how the program is intended to produce those outcomes (the link between actions and outcomes). After this, continue to work backwards to describe the connection between the inputs to the program and how those inputs support the important action steps that lead to outcomes. This chain of cause-and-effect is the logic model for the program.



Tool 2.1: Logic Model Template (Example)

Begin by writing the intended outcomes of the program, then describe the important actions that take place in the program to create those outcomes for students. Finally, identify the necessary inputs that will be required for those actions to take place.



Tool 2.1: Logic Model Template

Begin by writing the intended outcomes of the program, then describe the important actions that take place in the program to create those outcomes for students. Finally, identify the necessary inputs that will be required for those actions to take place.

Intended Outcomes

Actions

Inputs



2b. Establish concrete, specific measures of program's effects

Identify specific metrics to accurately measure each intended outcome.

It is important to gain agreement on the metrics that will be used to measure the impact for each intended outcome. Some outcomes have much clearer metrics of success than others, while other outcomes will have a variety of possible metrics. Articulating the metrics that will be used to measure the success of the program before conducting the analysis can help to ensure that stakeholders will later agree on the evaluation of the program.

DMGroup Tool

See "Identifying Outcome Measures" to map out which types of data will measure the effectiveness of the program.

Choose a metric that provides the best measurement, rather than settling for the one that is easiest to collect.

It may be that no current system tracks the specific data needed to measure the outcomes of the program, but settling for available data rather than effective data could compromise the analysis and reduce the chances of being able to take action after the analysis is complete.

Example

Outcomes from professional development can be difficult to capture. It might be tempting to measure the attendance rate or self-reported impact on teachers from the professional development sessions to determine the effects on teacher behavior. While these data, such as attendance data and satisfaction surveys, are readily available and easy to quantify, they do not indicate if teacher practice in the classroom actually *changed*.

Rather than measuring whether a teacher attended a training session on higher-order thinking (HOT) questions, or if teachers believe they have changed their practice, the district might measure the percentage increase of HOT questions asked in middle school math classes in the six months after the program began. Data could be collected by conducting random observations with specific "look-fors" around HOT questioning after establishing a baseline. This higher-quality data will take more time and effort to collect, and still be dependent on the capacity of the person doing the observations, but it will increase the quality and impact of the analysis significantly.

Example: Professional Development on Higher-Order Thinking

Easily Available Data	Teacher attendance at PD sessions or self-reported impact
High-Quality Data	Percentage change in HOT thinking questions asked in classrooms in the 6 months after the PD began based on 100 randomly selected 30 minute observations



Tool 2.2: Identifying Outcome Measures

List each intended outcome and a high quality metric to measure the specific impact of the program.

Intended Outcome	Metric
<i>Ex 1. Increased reading ability</i>	<i>DIBELS scores, ELA state test scores</i>
<i>Ex 1. Increased enjoyment of reading</i>	<i>Number of minutes read outside of school or number of pages read outside of school</i>
<i>Ex 2. Increased use of higher-order thinking (HOT) questions for MS math teachers</i>	<i>Count of HOT questions asked by teachers from random observations by instructional coaches</i>
<i>Ex 2. Increased understanding of mathematical concepts</i>	<i>MS math state test scores in particular strands</i>



2c. Set the benchmark for success

It is important to determine the benchmark for success for each intended outcome before collecting data to minimize bias when interpreting the results of the analysis.

There are many ways to analyze a program and even more ways to interpret the results. It is important to “put a stake in the ground” and to state expectations for the program before collecting any data. It is helpful to seek agreement of key stakeholders on the specific metrics and benchmarks for success of the program before starting the analysis. Articulating and agreeing upon these items before the analysis will help ensure objectivity during the “Draw Insight” and “Take Action” steps later in the process.

DMGroup Tool

See the “Benchmark for Success Template” to develop a specific definition of success.

It can be helpful to define benchmarks for success for multiple years of the program, if there is concern about implementation dip in the first year or belief that it will take multiple years to see the full benefit.

Use baseline data as a guide when creating benchmarks for success to ensure the goals are meaningful.

- Historical data
 - Academic history of students in the program
 - Historical data from students in the same grade, at the same school, or with the same teachers
- Benchmark data
 - District performance standards
 - State averages for a particular outcome

Example

Using baseline data will provide context when the district is defining the benchmarks for success for a program. For instance, imagine that a district anticipates a 25% increase in the number of 3rd graders reading on grade level due to its new reading program. If sixty out of one hundred 3rd graders read on grade level, then the goal is meaningful and significant (an increase of fifteen students on grade level). However, this gain may not be meaningful if only twelve out of one hundred 3rd graders currently read on grade level (adding only three students on grade level would not be a sufficient improvement if it leaves eighty-five students below grade level). In the second scenario, a district might be better off setting an anticipated gain based on a different metric, such as years of growth per year or a higher bar, such as tripling the number of students reading on grade level.



Mistakes to avoid

When Defining Success:

- Settling for data that is available rather than data that will be the best measure of the program's effects.
- Waiting until the analysis is complete to set benchmark for success.

During Communication:

- Waiting until after the analysis is completed to gain buy-in for the intended outcomes, metrics, and benchmark for success.
- Using metrics or a benchmark for success with which key stakeholders are either unfamiliar or feel are not relevant to evaluate the program.



Tool 2.3: Benchmark for Success Template

Using your information from “Tool 2.2: Identifying Outcome Measures” for the first two columns, set a specific benchmark for anticipated gains for each outcome to help measure the success of the program.

Intended Outcome	Data Needed	Anticipated Gains
<u>Ex 1.</u> Increased reading ability	DIBELS scores ELA state test scores	>1 year of growth in 1 year on DIBELS 15% growth in # of students proficient/advanced
<u>Ex 1.</u> Increased enjoyment of reading	Records of outside of school reading logs	>50% increase in minutes read outside of school for struggling readers
<u>Ex 2.</u> Increased use of higher-order thinking (HOT) questions for MS math teachers	Count of HOT questions asked by teachers from random observations by instructional coaches	>3x of questions from teacher are HOT
<u>Ex 2.</u> Increased student understanding of mathematical concepts	MS math state test scores	>15% growth in # of students proficient/advanced in specific strands





Step 3: Identify Comparison Group

Constructing a comparison between different efforts can provide deeper understanding into their effectiveness. A group of doctors might be testing a new treatment for pediatric asthma. Children who receive the treatment show an average increase in their lung function by 18%. Initially, the doctors might be excited about the promise of this treatment. However, to truly judge the effectiveness, there must be a comparison. The doctors find that children who received the standard treatment already available—not the new treatment—increase their lung function by 21% over the same time period. Had the test only measured the effectiveness of the treatment, it might have been expanded to more hospitals even though it is no more effective than the standard treatment already available. By setting up a comparison between two approaches, the doctors could identify the most effective approach available, not just an effective approach. A-ROI analysis should be constructed in the same way. If a district wants to expand a reading program across the district, conducting an A-ROI analysis of the three different reading programs in use at the elementary schools will help identify which approach is the best for students.

Key Outcomes

- When the district has completed Step 3 it should be able to answer:
 - Which students receive the program?
 - How will similar students be identified to include in a comparison group?
 - How will the comparison group be set up to measure the effects of the program?



3a. Decide to pursue a backward- or forward-looking analysis

Beginning at this stage, there are two different approaches a district could take when conducting an analysis: forward-looking or backward-looking.

Forward-Looking

A forward-looking analysis occurs when a district decides to analyze a program based on new data it will collect in the future. This might happen when a district wants to analyze a new program that is starting or use data from a future time period to analyze a program that currently exists.

Backward-Looking

A backward-looking analysis occurs when a district decides to analyze a current program using data from past years of the program.

There are tradeoffs to using each type of analysis, which are outlined in the diagram below. In many ways, a forward-looking analysis can take into account concerns and nuances better than a backward-looking analysis, making it easier to overcome potential objections. When designing a forward-looking analysis, districts can establish measures to ensure that the comparisons of outcomes are meaningful as well as organizing more efficient processes to collect accurate costs and outcomes data. The advantage of a backward-looking analysis is that a district can find results in a shorter time frame, because the analysis uses currently existing data and does not require future collection of data.

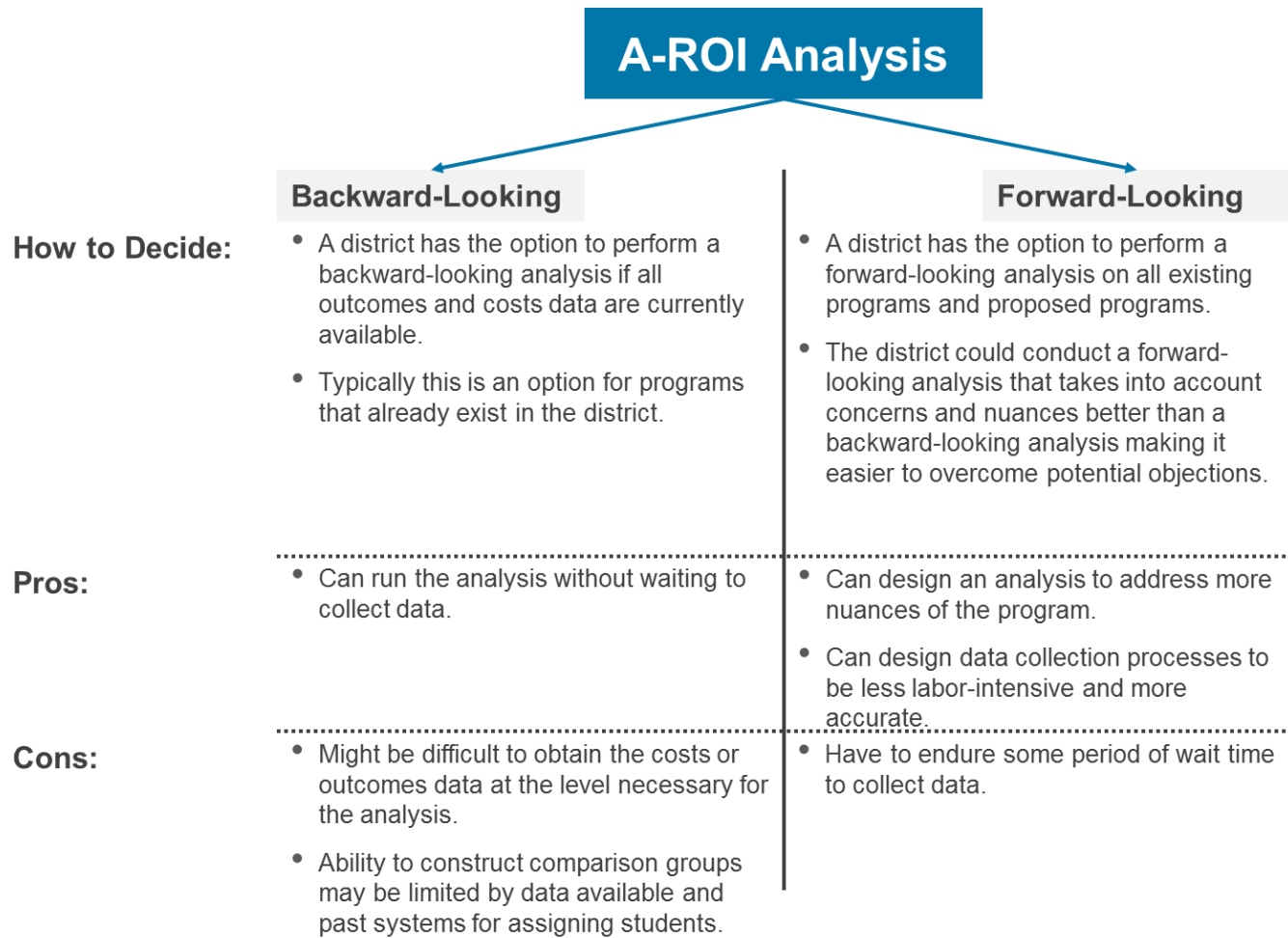
To determine if a forward- or backward-looking analysis would be more suitable for a district's analysis, consider the following questions:

Question	Yes	No
1. Does the district have all of the costs, outcomes, and segment data at the level necessary for the analysis?	<input type="checkbox"/>	<input type="checkbox"/>
2. Is it possible to identify exactly which students or schools took part in the program in each year?	<input type="checkbox"/>	<input type="checkbox"/>
3. Is it important to have results soon?	<input type="checkbox"/>	<input type="checkbox"/>
4. Can you identify any naturally occurring control groups in the past (see Step 3c)?	<input type="checkbox"/>	<input type="checkbox"/>

The more “yes” boxes checked, the more likely it is that a backward analysis will be a suitable approach. Though there is no hard-and-fast rule, it will be difficult to do a backward-looking analysis if questions 1 and 2 were not answered “yes.” Additionally, a forward-looking analysis can be conducted at any time during a program.



Tool 3.1: Comparing Forward- and Backward-Looking Analyses



3b. Identify students who receive the program

The next step is to understand how students are selected to take part in the program, and the criteria used to make this selection.

An effective A-ROI analysis is based on a comparison of two groups of students that are as similar as possible, except that one group receives a program and the other does not. To accomplish this level of comparison, it is critical to understand how students enter into the program of interest.

Typically, it is helpful to interview teachers and administrators involved in the program to fully understand how students are identified for the program. The identification process may look different for different programs.

Example

For a reading intervention program: students are selected based on a combination of DIBELS scores and teacher recommendations.

For a middle school math curriculum: all students enrolled in Algebra 1 at participating schools receive the curriculum.

For a targeted professional development program for teachers: students are impacted by the program if they are in a class with a teacher who chose to take part in the PD (in other words, there were no meaningful selection rules for the students to be impacted by the program).

It is important to distinguish between the formal process and the actual process for identification.

It is not uncommon to find a discrepancy between the formal policy for identifying students for a program and the process that is used in schools. This is important because it could indicate that the students in the program are not the same students for whom the program is intended. To understand both, it might be helpful to talk with the program or departmental director as well as teachers or practitioners that provide the program to see if there is variance between the formal and actual processes.

Backward-Looking

For backward-looking analyses, it is particularly important to distinguish between formal and actual identification processes.

Forward-Looking

For forward-looking analyses, this distinction is less important, because the district can create systems to ensure that only the students that should be in the program are placed in the program.



Example

For instance, the policy to enter students into an elementary reading program might weight DIBELS scores and teacher recommendations equally, but in reality students might only be entered into the program if their teacher makes a strong recommendation. This disconnect could lead to students with behavior problems—not necessarily reading problems—entering into a reading program that does not address the root cause of their academic struggles.

Use the table below to help guide your description of how students are identified for the program.

Formal Processes

How were students identified?	What is the benchmark for identification?	When does this process occur?
<i>Ex 1. DIBELS scores</i>	<i>180 composite score at beginning of 3rd grade</i>	<i>Students are identified in the fall of each year</i>
<i>Ex 2. All general education 8th grade students received this tech-based math program at Chamberlin MS and Smith MS</i>	<i>Being in a grade-level math course at a participating school</i>	<i>During course placement in the fall</i>

Actual Processes

How were students identified?	What is the benchmark for identification?	When does this process occur?
<i>Ex 1. Teacher recommendation</i>	<i>Teacher strongly advocates for student to enter the program</i>	<i>Throughout the year</i>



3c. Identify a comparison group

Constructing a meaningful comparison group is possibly the most important step in ensuring that an analysis will provide actionable results.

Creating an effective comparison group will help show how the program changed outcomes for students rather than just seeing raw outcomes data. There are four main types of comparison groups:

Types of Comparisons

Type of Comparison	Description	Example
Comparison to Similar Students	Comparing outcomes of otherwise similar students who were randomly assigned to either participate in a program or not.	Comparing math MAP scores for students at the district's two middle schools involved in the program to the students at the district's middle schools not involved in the program
Comparison to Baseline Data	Comparing the relevant outcomes of students in a program to previous outcomes for the same students or to students of the same teachers.	Comparing math MAP scores for students in the program to their scores before they entered the program or students at the middle schools before the program began
Comparison to Internal Standards	Comparing the outcomes of students in the program to a district performance standard.	Comparing the math MAP scores for students at the district's middle schools in the program to yearly targets to meet the district's performance standards
Comparison to External Benchmarks	Comparing the outcomes of students in a program to outcomes from a broader pool of students outside the district.	Comparing the math state test scores of 8 th graders at schools in the district's math program to statewide results



Using a randomized control group of similar students as a comparison model will provide the most rigorous results, although a non-random control group can be a feasible alternative.

Of the four types of comparisons outlined in the “Types of Comparisons” diagram, comparisons to similar students are the most effective at isolating the program effects when measuring outcomes. There are a variety of ways to construct a comparison of similar students, but the strongest model is a randomized control group.

Randomized Control Groups

A randomized control group is when a group of similar students is split at random, with one portion receiving a treatment program and the other portion receiving the “business-as-usual” services. A district usually will have to plan the creation of a randomized control group in advance, although in many cases it will only take a few minor tweaks to the usual student assignment process in order to accomplish this.

One of the most effective way to generate a randomized control group is to identify a large group of students who qualify for a program and use a lottery to select which students are placed into the program. This requires starting with an initial pool of students at least twice as large as the number of seats available in the program.

Example

For example, if a district has a set number of slots in a Reading Recovery program, the process would begin with identifying a list of students meeting the criteria for Reading Recovery with twice as many students as there are spaces. Then a lottery or random number assignment is used to determine which students will be placed into Reading Recovery. If the results for students in the program are better than for students not in the program, it is likely that Reading Recovery was the cause of the better results. This conclusion can be made because the students who received Reading Recovery were not significantly different than the students who did not. If the students had not been assigned randomly, it would be more difficult to draw conclusions about the cause.

As an alternative, students might be placed into Reading Recovery based on DIBELS scores until seats are filled, then the remaining students are placed into a separate reading program, even if their scores qualified them for Reading Recovery. This approach to assignment can still lead to significant findings, but because the assignment was not random, there may be other factors causing differences between the groups besides the program itself.

As this example illustrates, random assignment only requires a small extra step when designing the program and leads to much stronger results.

Non-Random Control Groups

Non-random control groups, although not as rigorous as randomized control groups, often occur naturally in districts. This can occur when groups of reasonably similar students are split, with one group receiving the program and the other group not receiving the program. These control groups leave a greater chance of external variables influencing the outcomes than a randomized control group, but in many cases can satisfy the needs of a district for an A-ROI analysis.



Naturally occurring examples of non-random control groups include:

- A group of schools volunteer to pilot a math program, but only some of the volunteers receive it in the first year.
- Elementary schools independently choose to use different reading programs.
- Some middle schools use a “team model” for staffing, while some use a “junior high model.”

Anytime a program is rolled out incrementally over multiple years with more schools or classrooms receiving it each year, there should be a control group to be found—though it is important to consider if the early implementers are led by principals or teachers who are generally more effective. If so, better results may not necessarily be attributable to the program.

Districts can (and should) require that all new programs be rolled out with a defined comparison group and plan to analyze impact.

The moment when a district is contemplating the roll out of a new district initiative, whether in a pilot group or district-wide, presents an ideal opportunity to set the stage for meaningful evaluations of program effectiveness in the future. If a district is launching a pilot, the district can define a comparison group of similar students within the district. If students are randomly assigned into the pilot or comparison group, then any difference in outcomes can be more strongly attributed to the program in question.

Even if a program is launched district-wide in the first year, the district can still identify a comparison group in another district or from national benchmarks. Though this comparison will be somewhat muddled by other factors and influences, defining the comparison at the outset will still help the district have a clearer point of view as to the success of the program.

When schools, teachers, or students have the choice to opt-in to a pilot program, a district should be cautious of using this group as a treatment group to compare against the rest of the population.

Typically, these “early adopters” are highly motivated and very willing to make adjustments to practice, which can skew the comparison. If a new science curriculum is piloted by the most enthusiastic science teachers, they will likely achieve good results with the new curriculum simply because they are strong teachers. This doesn’t provide conclusive evidence that the curriculum itself will be a positive change for all teachers.

Example

The examples below use the same three programs from Step 3a to illustrate how non-random (but still meaningful) comparison groups can be found:

For a reading intervention program: students within five points below the DIBELS identification benchmark that receive the program could be compared to students within five points above the benchmark that do not receive the program. This is not considered a randomized control group, but by comparing students just below the cutoff to those just above, the analysis will be comparing fairly similar groups of students.



For a middle school math curriculum: students in Algebra 1 at schools with the program could be compared to students in Algebra 1 at schools without the program. If there are concerns about schools with different student body makeup, it may be possible to construct comparison groups by identifying schools that have similar makeups.

For a targeted teacher professional development program: students in the classes of teachers that are participating in the program could be compared to both their MAP growth from prior years as well as the MAP growth of students in the classes of their teachers in prior years.

Revisit the benchmarks for success (outlined in Step 2c) to create measurable hypotheses about the effect of the program.

Once the design of the study is determined, the A-ROI team should make a number of measurable hypotheses about the effect of the program. To do this, the team should revisit the measure for success defined in Step 2, and set a target for how the results of the students in the program should compare to those receiving the “business-as-usual” alternative. A hypothesis can be a comparison of averages, a percentage difference, a higher number of students meeting a target, or any other measurable gauge of success that will be answered by the research design.

Example

The three examples from above are revisited here:

For a reading intervention program: The team has two measurable hypotheses about the effect of this program.

- Students receiving the program will show greater average growth in reading level over the course of a year, compared to the similar students not receiving the program.
- The increase in the percentage of students reading on grade level will be higher for the group in the program than for the group out of the program.

Once the two student groups are determined, testing these hypotheses will require fairly simple analysis of the student results.

For a middle school math curriculum: The team makes only one hypothesis for this program.

- Students in schools with the new curriculum will have higher growth in their Galileo scores during the course of the year, compared with students in schools with the old curriculum.

For a targeted teacher professional development program: The team makes two hypotheses for this program.

- Students in the classes of teachers who received the PD will show greater growth in their MAP scores than students of teachers who did not receive the PD.
- Students in the classes of teachers who received the PD will show greater growth in their MAP scores than the students of *those same teachers* in the previous year, before the teachers had the PD.

The second hypothesis is important because of the possibility that the teachers in the PD program have some selection bias to be above average teachers, in which case their students would have greater



growth than similar students even without the benefit of the PD. By comparing to the classes of the *same* teachers in a previous year, the study can more clearly isolate the effect of the PD.

This is a case where randomized selection of *teachers* to receive the PD would boost the strength of the analysis. Just as described above, the district could have identified a large pool of qualified and interested teachers, then randomly picked half to receive the PD and half to serve as a control group.



3d. Solicit input on the design

Before collecting data, this is an important chance to confirm the research design and the validity of the comparison group.

Soliciting input on the structure of the analysis from the stakeholders that have influence over the decision-making process in the district is important for three reasons:

1. Soliciting input allows a district to consult the participants in the program to catch any oversights in the methodology thus far. This will help increase the quality of the analysis.
2. Communicating the methodology moving forward and setting expectations about the results of the analysis are important. It is helpful to invest time in priming stakeholders to interpret the results in a meaningful way.
3. This is an opportunity to confirm that the data set the district plans on collecting is the right data set for the analysis.

Soliciting input at this step is particularly important because this will probably be the last point at which some stakeholders engage with the process until after the analysis is complete.

This is the time to begin including the person who will conduct the statistical analysis in the process so that he or she has a deep conceptual understanding of the content.

To ensure that the analysis is completed in the way that the district leadership intends, it is important to include the person who will run the statistical analysis in each stage of the process beginning in the design phase. Providing the person in charge of data analysis with a deep conceptual understanding of why and how the analysis is being structured is critically important to ensuring an accurate and actionable analysis.

Example

The following story illustrates how an analysis can go wrong when the data analyst does not fully understand the context and goals for the work. One district conducted an analysis of a district-wide reading program, which had a goal of helping students make more than one year of growth during the school year. Upon first glance at the end of year data, it appeared that nearly 80% of the students in the program had not made one year of growth. The program appeared to be ineffective.

However, when district leaders probed more deeply into the analysis, they found that the person who conducted the data analysis did not include students who had benefited so much that they were successfully exited from the intervention after returning to grade-level reading ability. They were not “in the program” at the end of the year, so the statistics person did not include them in the analysis. When these students were included in the analysis, the results showed that the program was wildly effective, but the district was on the brink of reducing the program due to a miscommunication with the data analysis department.

Additionally, this is the best time to check for flaws in the research design. If the analysis uses a flawed design, it will be much more difficult to draw meaningful conclusions. The person with statistical expertise should check to confirm that the comparison groups are valid, the sample size is sufficiently large, and the data will be collected in a way that will allow the desired analysis to be done.



Mistakes to avoid

When Designing the Analysis:

- Not stating a testable hypothesis about how the measures of success should change as a result of the program.
- Comparing groups that have substantial differences that might obscure the impact of receiving the program.
- Using a comparison group with a sample size that is too small for the analysis to provide results with certainty.
- Treating opt-in or volunteer selection as a random assignment. A possible solution to this is soliciting volunteers for the pilot and only providing the program to some of the volunteers randomly chosen from that group.

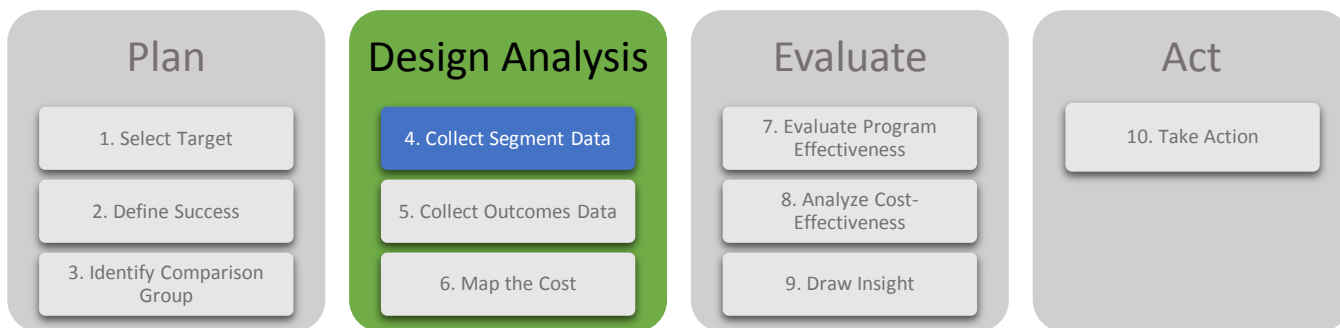
During a Forward-Looking Analysis:

- Not constructing a randomized control group (or at least a control group).

During a Backward-Looking Analysis:

- Beginning a backward-looking analysis and realizing later in the process that some of the data is not at the correct level of detail or is incomplete for the analysis.
- Not confirming that a sufficient control group exists before beginning a backward-looking analysis.





Step 4: Collect Segment Data

Introduction to Segmentation

The concept of student segmentation is new to many districts, so the overview below may be helpful in understanding the role that segmentation plays in the A-ROI process.

Typical program evaluation in districts seeks to answer the question “Does this program work?” The response is generally a blanket “good or bad” judgment about a program’s effectiveness. Unfortunately, the “good or bad” approach to program evaluation limits districts’ options for what to do with the program. If a program is bad, the district should eliminate it; if a program is good, the district should keep it.

With A-ROI, the analysis should move past “good or bad” and into a mindset of “good or bad *for which students*.” Many districts already look at test scores broken down by student sub-groups as defined by No Child Left Behind (NCLB), such as race, disability and poverty status. However, these groupings are overly broad and within each group there will be students who respond very differently to specific programs and interventions. To assess which students benefit most from a program, it is important to use segments that group students according to characteristics relevant to the program.

Student segmentation by key variables allows for more meaningful comparisons between programs. To compare the effectiveness of Program A to the effectiveness of Program B, segmentation allows an analysis to isolate how a particular segment of students performed in each program.

For example, a district might assess whether their elementary reading intervention is working well for students with special needs. To take this one step further, the district could segment the students with special needs according to the actual learning challenges that they face: dyslexia, ADHD, specific learning disability, etc. In addition, the students can be segmented further according to their reading level at the start of the year. Then within a given segment, the analysis can compare the results of students who received the intervention with those who did not. With these factors taken into consideration, the district can now draw much more sophisticated conclusions about who actually benefits from the reading program.



Key Outcomes

- After completing Step 4 a district will have:
 - At least one hypothesis about which student segments will benefit most from the program
 - At least one hypothesis about what the most important variables for determining success in the program are
 - The data that will be used to place students into each hypothesized segment

4a. Make a hypothesis about segments that will benefit most

It is important to hypothesize which student segments will receive the most benefit from the program and identify the key variables that might lead to their success.

The purpose of making a hypothesis at this stage is to guide the data collection process. Before collecting segment data, it is important to hypothesize which groups of students will benefit most from the program. This can be done by:

- Understanding the specific criteria for entering students into the program
- Identifying a few key variables that might influence how students respond to the program (see Tool 4a)
- Identifying possible student segments based on the key variables

The analysis of segments will go deeper in Step 7 when analyzing the effectiveness of the program, at which point the segments may be re-defined.

Choose the key variables based on the root causes of student successes and struggles.

When identifying the key variables, it might help to think about the root cause of the student's academic struggles or level of performance. In many cases students are selected for an intervention program based on test scores or grades, which do not reflect the root cause of why a student is struggling. The process of identifying key variables provides greater insight into the root causes of student performance.

Not all hypothesized student segments will be meaningful in the end, but generating hypotheses can guide a district as it determines which data might be important to collect for the analysis. If there are competing hypotheses about key variables, a district can collect data on each attribute and test their importance.



Example

For instance, imagine that students are placed into a double-time math intervention program due to low math scores. One possible segment in the program that might benefit significantly is students with high engagement that learn math slowly, since the root cause of their struggle in math was a lack of content understanding. Another possible segment is students with low attendance but high academic potential. Given that the root causes of these students' struggles are different, it is unlikely that the same intervention will impact the student segments in the same way.

If these were the hypothesized student segments, the district might want to collect attendance and homework completion data to measure engagement to help identify which students are in each segment. These data will help to understand the different root causes of each student segment's struggles.

Example

Typical Student Segmentation

<u>Student</u>	<u>Math Score</u>	<u>Current Segmentation</u>	<u>Intervention</u>
Student A	59	Special Education	All four students receive math intervention
Student B	60	English Language Learner	
Student C	61	Low-Income	
Student D	60	African American	

Proposed Student Segmentation

<u>Student</u>	<u>Math Score</u>	<u>Root Cause</u>	<u>Intervention</u>
Student A	59	Has difficulty conceptualizing fractions	Receives differentiated instruction with a highly effective math teacher
Student B	60	Is tuned out, seldom completes homework, and has substance abuse issues	Receives rigorous content and social work resources
Student C	61	Learns math very slowly	Receives extra time with math teacher on current material
Student D	60	Had ineffective 6 th - and 7 th - grade math teachers	Receives extra time with math teacher on current <i>and past</i> material



Tool 4.1: Possible Key Variables for Segments

Key Variables (Academic)	Potential Segments
Reading Ability	<ul style="list-style-type: none"> • On or above grade level • <1 year behind • >1 year behind
Subject-Specific Ability	<ul style="list-style-type: none"> • On or above grade level • <1 year behind • >1 year behind
Specific Special Needs	<ul style="list-style-type: none"> • Mild/Moderate Autism • Severe Autism • Mild/Moderate Speech and Language Disability • Severe Speech and Language Disability • ADD/ADHD/OCD • Non-ADD/ADHD/OCD Other Health Impairments • Orthopedic Impairment/Hearing Impairment/Visual Impairment/Deaf-Blindness • Multiple Disabilities/Traumatic Brain Injury • Emotional Disturbance • Mental Retardation/Developmental Delay
Specific ELL Status	<ul style="list-style-type: none"> • Continuous rigorous schooling, but little English • Continuous low-rigor schooling, but little English • Limited or inconsistent formal education • Continuous rigorous schooling, and moderate English proficiency <p>These can be tracked by knowing the country of origin for ELL students and measuring their DIBELS/TOEFL scores.</p>
Student Engagement	<ul style="list-style-type: none"> • Homework completion rate: <ul style="list-style-type: none"> ○ >90% ○ 70-90% ○ <70%
Extracurricular Involvement	<ul style="list-style-type: none"> • Member of two or more groups • Member of one group • Not involved in a group
Grade Retention/Promotion History	<ul style="list-style-type: none"> • Passed all grades on schedule • Repeated one or more years of school
Grade	<ul style="list-style-type: none"> • K-12
Participation in Related Programs	<p>For a dropout prevention program:</p> <ul style="list-style-type: none"> • Has a mentor with the Big Brother/Big Sister Mentorship Program <p>For a reading program:</p> <ul style="list-style-type: none"> • A student's parents provide an outside tutor on the weekend

Tool 4.1: Possible Key Variables for Segments, Cont.

Key Variables (Non-Academic)	Potential Segments
Age	<ul style="list-style-type: none">• Within expected age range for grade• Older than expected• Younger than expected
Attendance	<ul style="list-style-type: none">• >98%• 95-98%• <95%
Behavior	<ul style="list-style-type: none">• No detentions• 1-2 detentions this year• >2 detentions this year
Parent Engagement	<ul style="list-style-type: none">• Parent is a member of PTA• Parent attended one or more parent-teacher conferences• Parent did not attend a school event this year
Transience/Mobility	<ul style="list-style-type: none">• Changed schools at least once in the last 3 years• Changed schools in the last year• Remained at the same school for the last 3 years• Moved schools in the last year, but stayed within the district



4.1 Examples of Common Programs and Suggested Key Variables

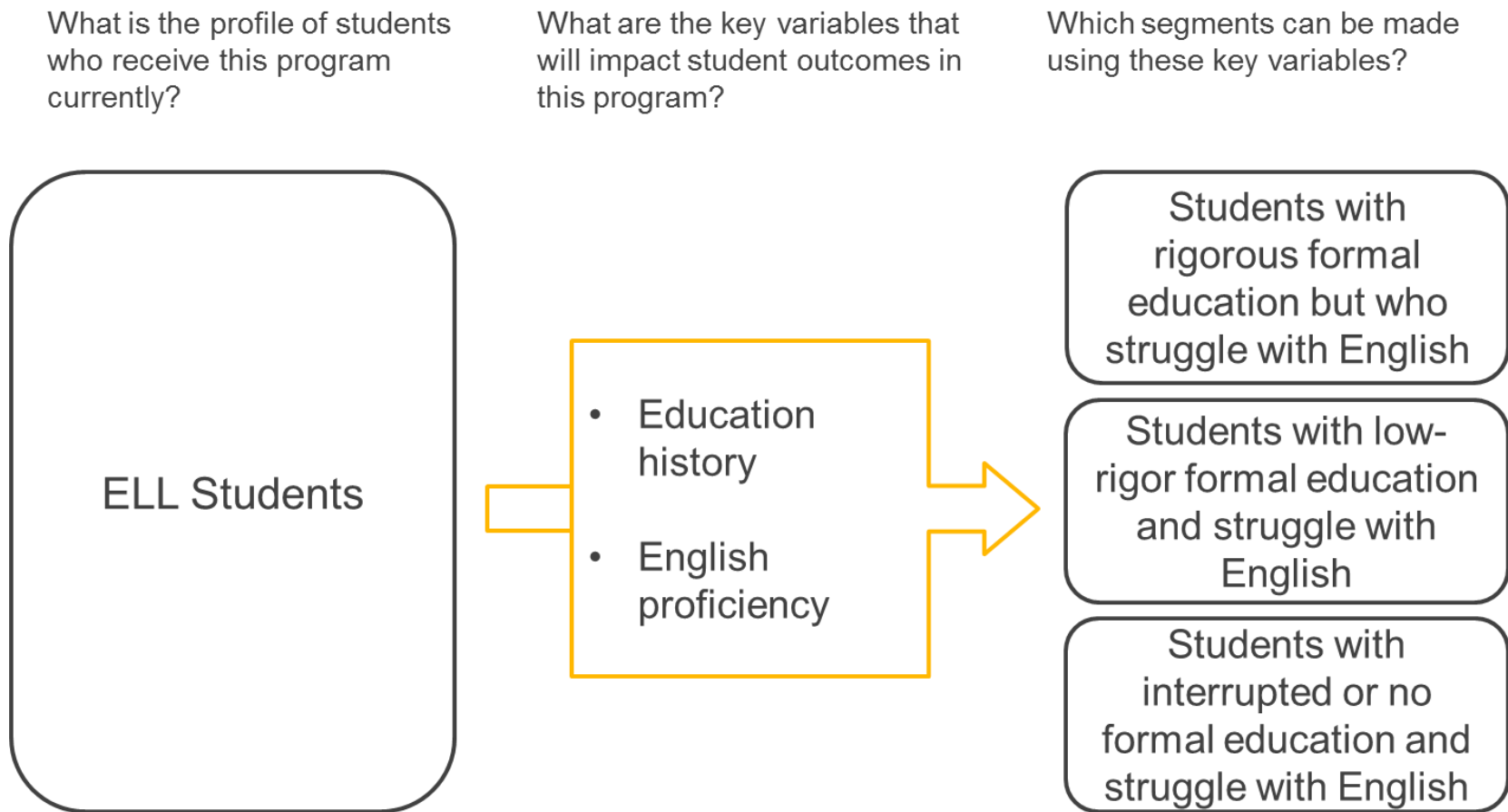
Type of Program	Suggested Key Variables
Math Curriculum	<ul style="list-style-type: none">• Math ability• Reading ability and/or ELA scores• Attendance
Reading Intervention for ELL Students	<ul style="list-style-type: none">• Reading ability• Math ability• Specific ELL segment
Co-Teaching Model	<ul style="list-style-type: none">• Specific special need• Math and reading ability• Behavior
Dropout Prevention Program	<ul style="list-style-type: none">• Attendance• Career aspirations• Educational attainment of siblings



Examples of Hypotheses and Segmentation

ELL Curriculum

Hypothesis: This program will be most effective for students who have content knowledge and ability from previous education but for whom English is a barrier to learning.



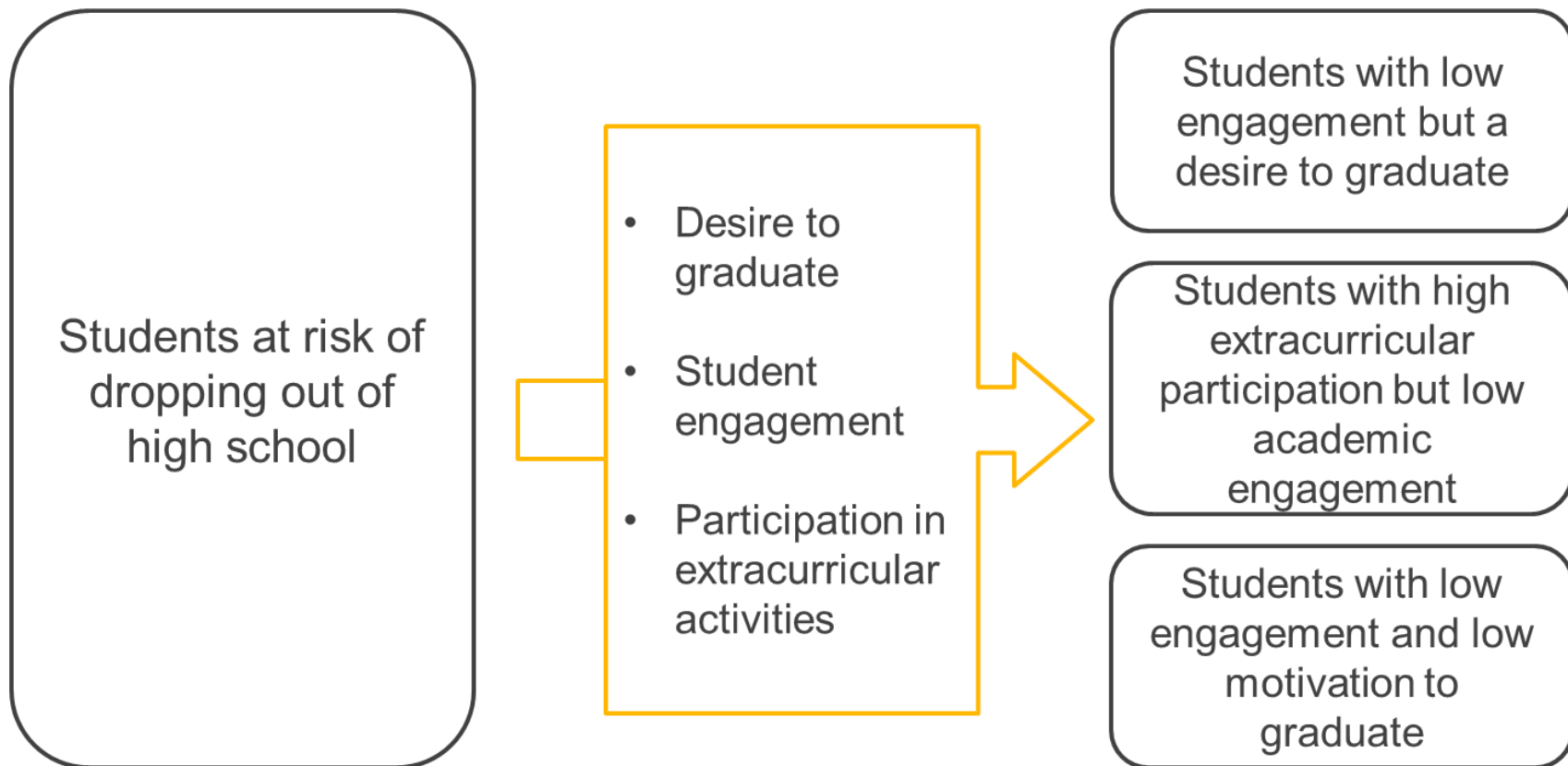
Dropout Prevention Program

Hypothesis: This program will be most effective for students who are unengaged in school but do not want to damage their future prospects by not graduating.

What is the profile of students who receive this program currently?

What are the key variables that will impact student outcomes in this program?

Which segments can be made using these key variables?



Math Intervention Program

Hypothesis: This program will be most effective for students who are engaged in school and read on grade-level, but struggle in math.

What is the profile of students who receive this program currently?

Students with low math scores

What are the key variables that will impact student outcomes in this program?

- Student engagement
- Reading ability

Which segments can be made using these key variables?

Students who struggle in math but are grade-level readers with high engagement

Students who struggle with engagement, reading, and math

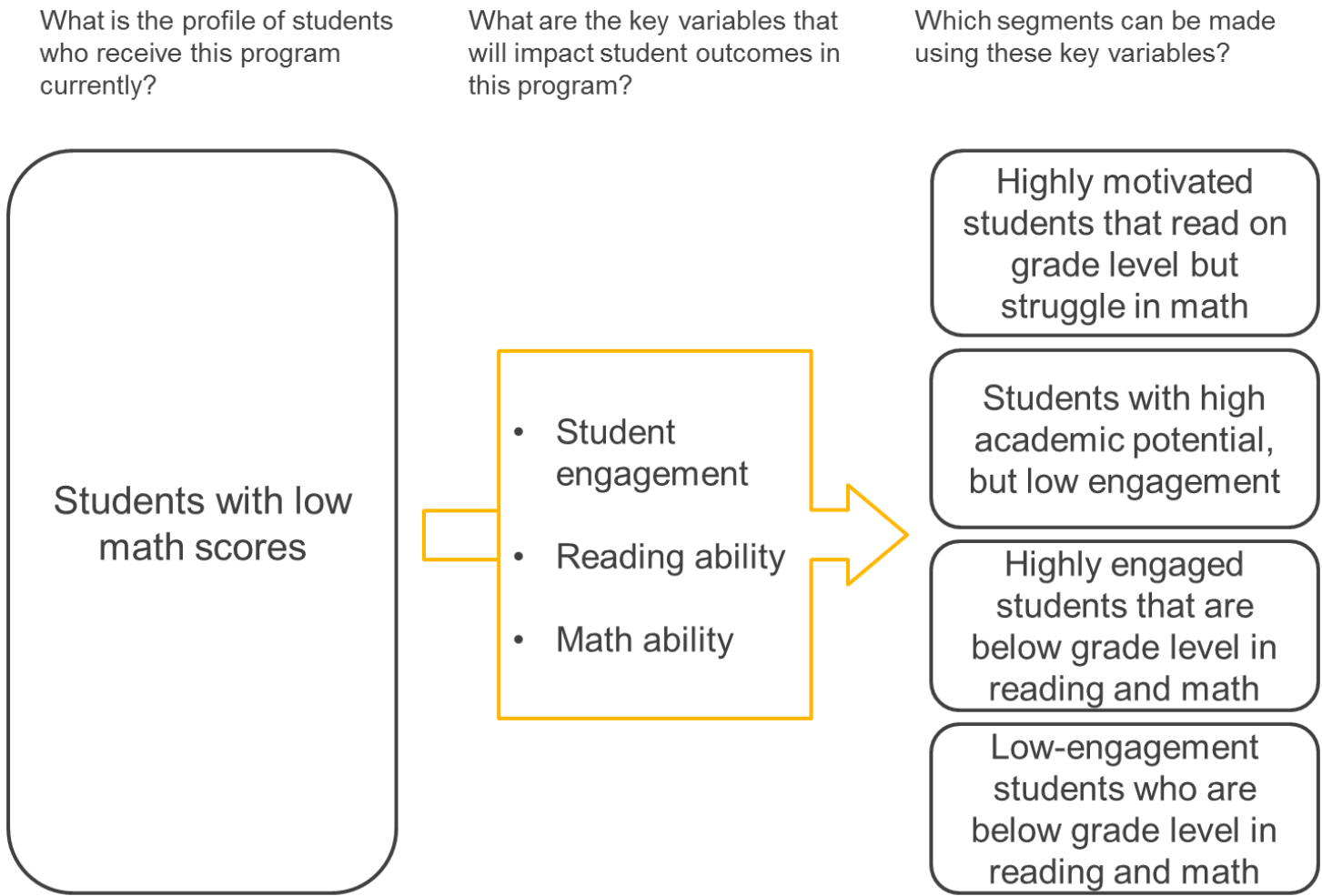
Students who struggle with math and reading but are highly engaged

Students who struggle with math and engagement but read on grade-level



Computer-Based Math Intervention

Hypothesis: This program will be most effective for students who are motivated to learn and are persistent on independent tasks, but struggle in math.



Tool 4.2: Student Segment Organizer

Use the diagram below to outline the hypothesized segments for the program being analyzed. Hypothesis:

What is the profile of students who receive this program currently?

Placeholder for student profile

What are the key variables that will impact student outcomes in this program?

Placeholder for key variables

Which segments can be made using these key variables?

Placeholder for student segments



4b. Collect segment data

Sometimes existing data on student characteristics can be insufficient for creating meaningful segments, so it is important to consider collecting new data to craft more relevant segments.

Districts collect a significant amount of student data, and much of it can be used to create meaningful student segments, such as segmenting by reading level. In many districts, the SIS contains a great deal of detailed information on student characteristics that could be used to identify student segments. There is a significant amount of data that districts do not collect, however, that could be equally important when identifying student segments, such as measures of student engagement.

DMGroup Tool

See the “Tool 4.3: Outcomes and Segment Data” in the A-ROI Workbook to see an example of how to set up the data collection tool.

Many districts avoid creating new processes for collecting data, fearing that it will be a drain on resources. Some data are simply too difficult to collect, but, in many cases, segment data can be captured without too much investment if the process is thoughtfully created. The additional segment data should have the potential to provide much greater insight into which students benefit from a specific program.

Each district will have to decide whether it is worth the effort to invest in collecting the segment data, but the more accurate the segment data is, the greater the possibility that the analysis will provide actionable insights. Possible methods for collecting the data include:

- Student interviews
- Parent or student surveys
- Additional fields in current data requests from teachers or principals

In some situations it is necessary to collect more segment data after beginning the analysis if the initial findings indicate that the segmentation is not meaningful.

The collection of segment data does not have to be completed exclusively before the analysis. It is usually easier for a district to collect all of the segment data at the outset, but if certain trends indicate that more specific segment data is needed, oftentimes it can still be collected, especially in a forward-looking analysis. As a district gains more experience identifying student segments, the likelihood that it will need to collect a second round of student data decreases.

Example

One district found tremendous variance in outcomes when analyzing a reading program. Upon further inspection, the leadership noticed that the growth for ELL students was wildly variable, indicating that “ELL students” was probably not a meaningful student segment. The district decided to collect more nuanced data in hopes it would shed light on the cause of the differing outcomes.

The data collection process did not require as significant of an investment as the district anticipated. Rather, one person in the superintendent’s office spent a few days going through the original files of each ELL student and recorded information about the student’s country of origin and his or her educational history. The necessary data was added to digital student profiles as it was collected, with the whole process taking less than one week.



The district did not retrieve data for every ELL student, but it did collect data for most, which is better than the alternative of not having any. The data provided new insight, which revealed that the country of origin determined how well a student responded to the program. Students that had received rigorous education in their home country, for instance in Japan, responded much more favorably to the program than students who did not receive rigorous education in their home country, for instance Senegal. This allowed the district to target the reading program to the students who were likely to benefit from it, while developing a different option for the students who were not.

Additionally, the district then added two questions to the survey sent to parents of new ELL students as a mechanism to collect the data in the future, so that the district can place students in the program that will be most effective for their needs.

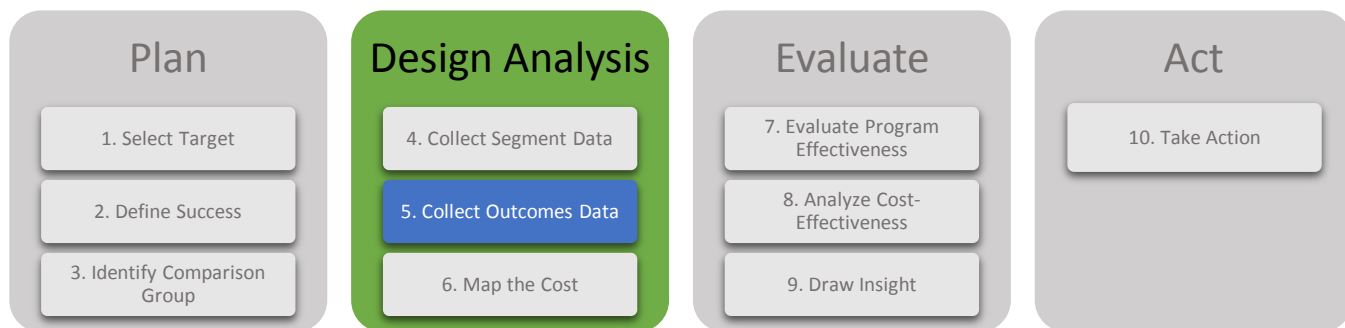
It is helpful to collect data to capture which teachers students had and which schools students attended to allow for classroom-level and school-level comparisons during the analysis.

Student segment data provides answers to “For whom did the program work?” but student-level outcomes data that can be matched to specific teachers and schools can provide an answer to “Why did the program produce the results that it did?” By analyzing the variance in outcomes based on teachers or schools, districts can gain greater insight into how the implementation of the program influences its effectiveness, protect against acting on a “false positive” result, and also identify pockets of excellence in the district that could shed light on how to fix the program.

This type of data may not be of critical importance for every district or program evaluation, but if it is available, it is worth collecting because it may be useful during analysis.

Similar strategies can be used to collect background and implementation data as were outlined in Step 4b when describing the collection process of segment data.





Step 5: Collect Outcomes Data

Collecting the most meaningful outcomes data is critical to fully understanding the effectiveness of a program. When Major League Baseball managers began using data analysis to inform their coaching decisions, some noticed that the most common metrics used to gauge player performance were not the most important. For instance, they found that one of the most important predictors of winning games was how often a team had players on base. To track the new outcome, managers began collecting data on how often each of their players reached base, called on-base percentage. This was a significant shift from the traditional metric, batting average, which measures how often a player gets a base hit, only one of many ways players can reach base. By collecting new types of data to capture a more meaningful outcome measurement, these coaches were able to make more informed coaching decisions. A similar approach can be taken during A-ROI analyses. If a reading program is provided to students across the district in different ways but there is no data about group size or time on task, investing the time to collect that data will bolster the accuracy and depth of the analysis. If the district has a dropout program but no cohort data to illustrate the annual dropout rate, collecting the new data will help shed light onto how well the program is working.

.....

Key Outcomes

- After completing Step 5, a district will have:
 - The data that will be used to measure the effectiveness of the program
 - The data that will be used to determine the fidelity of implementation and levels of service

5a. Collect readily available data and ensure that it is fully understood

Collect readily available outcomes data and engage the district data manager to understand in detail what is included in the data and what still needs to be collected, even if it seems straightforward.

Similar to the segment and background data, some of the outcomes data will already be tracked in the student information system (SIS). If that is the case, the data can typically be collected through a data pull from the central SIS or through a data request from individual schools.

However, caution needs to be exercised. District leaders can easily fall into the trap of requesting a particular data set or analysis from someone in the district data department who is not attuned to the reasons for the request. In many districts, the central office data department is isolated from decisions made about instruction. This leaves room for different assumptions to be made, details to be lost in translation, and incorrect conclusions to be drawn.

Example

As described earlier, one district wanted to analyze the effectiveness of its reading intervention program, given the substantial cost of the program. The superintendent requested an analysis of how students in the program fared relative to their peers who were not in the program. The initial results from the data department indicated that the interventions were completely ineffective, and nearly all of the students in the program showed less growth than similar students who were not in the program.

However, later in the conversation the superintendent discovered that the analysis provided by the data department did not include students who experienced enough growth to achieve grade-level reading ability and exit the interventions over the course of the year. When the students who achieved grade-level ability and exited the program were included in the analysis, the program was found to be wildly successful; the majority of students returned to grade-level ability in less than one school year. This program was nearly cut based on the initial (and incomplete) analysis, although it was not incorrect from the perspective of the data department. The data department did provide data on the students who were labeled as being “in the program” when the data request was submitted. However, they did not fully understand what the Office of Instruction wanted to measure (e.g., data for students who received the program in the last year), and the Office of Instruction did not communicate clearly their intent. This allowed room for distinct and isolated departments to make different assumptions about the same data, leading to critical details nearly being lost and incorrect conclusions being drawn.



Cleaning and quality checking the outcomes data is critical to ensuring that the analysis is accurate.

Raw data is rarely ready for analysis without significant quality checks and data cleaning. Overlooking the need to clean the data and conduct quality checks may lead to an analysis that provides an incorrect result. Some strategies to begin the data cleaning process include, but are not limited to:

- Checking the data for duplicate entries
- Verifying that all schools that should be included are present
- Selecting a few sample rows and checking that the data for that row are correct
- Common sense checks around total entries or scores (e.g., if there are 350 students in a school and the outcomes data provides a list of 600 students, some investigation is probably needed to determine why there is such a discrepancy)



5b. Create processes to collect new outcomes data

If the analysis requires data that is not currently collected, then the team doing the analysis faces a choice.

It may be possible for the team to collect the data as a one-time effort. However, the new data will most likely be relevant for analyses in the future, so it may be worthwhile to invest the time and effort to create data collection processes now so that data will be automatically collected going forward.

The processes for collecting this data are different for forward- and backward-looking analyses:

Forward-Looking

Usually this is simpler than a backward-looking analysis since the process can be created and implemented in advance. New processes could include adding a required field in the SIS student data profiles, administering surveys, or conducting participant interviews, depending on the intended outcomes of the program.

Example

For instance, a district might plan to implement a reading program that provides intensive instruction for the first half of the upcoming school year. If a district traditionally conducts DIBELS testing only in the fall and the spring, it may want to include more testing dates in the first half of the year to measure growth over the course of the program to get the most accurate measurements of its effects.

Backward-Looking

A backward-looking analysis may require significantly more effort to capture all of the needed data. The data might be stored in difficult-to-access databases, or some data might not exist. For instance, some districts might only have data recorded for the end-of-year assessments going back multiple years, but no beginning-of-year data, making it difficult to calculate annual growth to capture the effects of a program.

Example

A district might want to analyze its technology-based math program, which began three years earlier. It is possible that the district does not have student-level data from before the start of the program stored in its central SIS. To collect the data, the district might have to search through the paper files at each school to record the past outcomes data for each student. It is typically possible to capture the data needed, but it will generally be much more time-intensive than the process for a forward-looking analysis.

The other option for the district is to conduct a forward-looking outcomes analysis, which could be paired with a cost analysis that spans the lifetime of the program. A district does not lose much by taking this approach; it would still understand if the program is effective and how much it has invested in the program. The downside is that the analysis will have a longer timeline, as with any forward-looking analysis.



5c. Collect implementation data

Collecting data about how the program was implemented will be important for calculating the cost-effectiveness of the program and may require the creation of new processes for collection.

The two types of implementation data are data related to level of service and data related to the fidelity of implementation. There is some overlap between them, but both types of data are important for understanding how a program was delivered to students. Level of service data describe key factors in implementation, such as minutes of service provided, number of students in a group, or number of staff in the class. Fidelity check data illustrate how faithfully the program was delivered to the way it was intended to be delivered.

It is important to understand how a program's services are provided to students for two reasons. First, understanding group size and the time staff spend providing services to students are necessary pieces to conducting the cost and cost-effectiveness analyses.

Second, collecting data around the level of service provided to students can shed light on whether the program is being implemented as it was intended. For instance, if a school only has one reading teacher and 50 students in Reading Recovery, the principal might make the decision to provide the program in groups of three students instead of one-on-one. Knowing this fact about implementation can help explain why the program was or was not effective.

Examples of level of service data include:

- Group size
- Frequency of services
- Average length of services (hours/week)
- Duration in program (months)

Collecting data about fidelity of implementation will be important in determining if a program was actually responsible for the outcomes that were recorded.

Data about fidelity of implementation is critical to confirming that the program is driving the outcomes that were observed. It is not uncommon to find a discrepancy between what programs the central office believes are being used in the classroom and what teachers are actually using. If the program is not being used in the classroom, it cannot be the main cause of student outcomes, positive or negative.

Examples of fidelity check data include:

- Classroom observations to record changes in teacher practice with program-specific "look-fors"
- Records of instructional time provided to students
- Surveys of students or examination of student work artifacts to determine if they have learned the central ideas that teachers should emphasize if implementing the program with fidelity

Example

An example of the important role implementation data plays in A-ROI analyses can be seen through the analysis of a reading program. It would be possible to roll out a reading program with textbooks and significant professional development, with the initial analysis indicating that the program caused significant growth in students' reading ability. However, if classroom observations were conducted as fidelity checks, they might show that very few teachers were actually using the concepts or textbooks from the program. If very few teachers actually implemented the program, then the district might determine that a factor other than the program is the driver for student growth.

Mistakes to avoid

When Collecting Outcomes Data:

- Make sure the outcomes data is at the student-level, not aggregated by classroom or school.
- Not collecting enough fidelity check data to determine if the program was implemented as planned.
- Making assumptions about the dataset without verifying, including which students were included or excluded, how the outcomes were calculated or collected, and many other details.

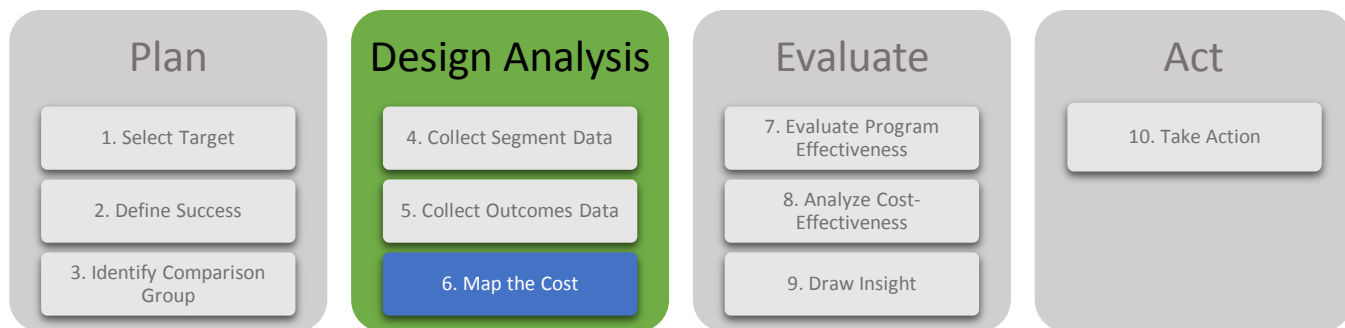
During a Forward-Looking Analysis:

- Creating a process to collect new data that captures aggregate data at the classroom- or school-level rather than at the student-level.

During a Backward-Looking Analysis:

- Engaging in a backward-looking analysis and not using student-level outcomes data.
- Beginning a backward-looking analysis and realizing that fidelity check and implementation data will be difficult to capture accurately.





Step 6: Map the Cost

Introduction to A-ROI Cost Accounting

The cost accounting presented in this section is unlike typical cost accounting in districts in five major ways:

It is not bound by compliance requirements.

Most school budgets are structured to comply with state reporting, so that revenues and expenditures must match to the penny. The cost accounting presented here is different. It is based on a need to understand the resources allocated to a specific program; this requires more nuance than finding the sum of line items in the budget, but less precision than matching revenues and expenditures to the penny.

Personnel costs are not fixed costs; staff time invested counts toward the cost of the program.

Districts typically view personnel costs as fixed costs, which limits leaders' understanding of cost and the management options available to them. DMGroup's approach to cost accounting treats staff time as a resource allocated to the program, which can be adjusted and reallocated however the district sees fit. Staff time includes the time general education teachers spend teaching students, the time a paraprofessional spends supporting a student, and time spent in professional development.

Expenditures from all sources, including grants, are counted.

Many districts keep separate budgets for the general fund, Title I, IDEA, and so on, sometimes with distinct accounting systems for each due to compliance requirements. The cost accounting approach presented here treats all dollars the same and does not exclude items paid for by different funding sources.

It takes more nuance than including or excluding each line item in the budget.

Many of the costs, especially investments of time, will not be distinct line items in the traditional school budget, but must be captured in the A-ROI cost accounting approach. For instance, instructional coaches might dedicate a portion of each week to supporting teachers of a reading intervention. In this case, the full line item of "Instructional Coaches" should not be included in the cost calculation for the reading intervention. However, to get a true total cost of the program, a

portion of the line item should be included based on an estimate of the time investment of the one instructional coach who worked on the program.

It is much more important to estimate the large costs accurately than to include all of the small costs.

Typical school accounting efforts, because they are forced to meet exact standards for compliance, focus on finding all of the costs that are included in the budget, no matter how small. The cost accounting approach outlined in this section focuses on estimating large costs accurately rather than including all of the minute costs. For instance, ensuring that the cost of a teacher who spends half of her time teaching reading is included in the program cost is much more important than working hard to determine whether 64 or 65 textbooks were purchased for the program.

There are three main steps in gathering the cost data before a more nuanced cost analysis can begin. These steps are described in more detail throughout Step 6:

1. Map out all the types of costs for the program
2. Collect the needed cost data for cash and time investments
3. Calculate the total cost of the program

Key Outcomes

- After collecting the costs data, a district will have:
 - All cash investment data and the information to reasonably estimate the cost of staff time investments
 - An estimate of the total cost of the program
 - Additional data for each expenditure that will help provide greater insight into how the program uses resources



6a. Map out all the types of costs associated with the program

To ensure a solid estimate of costs, it is helpful to conceptually outline all of the different costs associated with the program before collecting and calculating the data.

Many districts do a good job of capturing the cash investments associated with programs. Typically, these are included in the district budget database and can be accessed when necessary. However, the major costs of implementing a program are often tied up in staff time investments, not cash purchases. Many districts do not include time investments of staff in their cost calculations for a program. However, typical districts spend about 75% of their resources on staff, so the ways that staff use their time is an incredibly important piece of a district's resource allocation.

Many of the different costs are housed in different databases—or are not housed in a database at all—and need to be collected differently. Creating a list of the types of costs incurred from the program will help ensure that no data are overlooked during collection. It might be useful to distinguish between cash and time investments to help structure the identification process. From there, articulating the methods that will be used to collect each type of data will make the collection process much more efficient.

The “Collecting Costs Data” tool can help guide this process. Create a list of all costs associated with the program. Use the list below as a starting point, and create a list of cost categories relevant to the selected program. Next, identify how the cost will be collected or estimated for each category. When collecting the cost data, ensure that the data is complete for each type of expenditure.

Time Investments	Cash Investments
<u>Staff Time on Direct Service/Instruction</u> <ul style="list-style-type: none">• Instruction/Direct Service Time• Identifying Students for Entry into Program	<u>Materials</u> <ul style="list-style-type: none">• Technology• Textbooks and Curriculum
<u>Staff Time on Training</u> <ul style="list-style-type: none">• Professional Development Time• Staff Collaboration Time	<u>Operations</u> <ul style="list-style-type: none">• Facilities Rental or Construction• Food• Student Transportation
<u>Administrator Time Dedicated to the Program</u> <ul style="list-style-type: none">• New Administrative Positions Created for the Program• Portions of Existing Administrative Positions Dedicated to the Program	<u>Fees and Stipends</u> <ul style="list-style-type: none">• Non-PD Consultant Fees• Professional Development Fees• Staff Travel Stipends• Stipends for Additional Service• Substitute Teachers

Example

A district implemented a reading intervention program that included significant upfront professional development for all of the teachers involved. When calculating how much that professional development cost, the district initially pulled the consultant fees, the travel and food for the training, and the cost of the substitute teachers who covered classrooms during the session. However, including all of these costs does not capture the total cost of the professional development



sessions. The costs do not account for the time the teachers spent in the session. For a day-long session of PD (~6 hours) with ten teachers, this is a substantial investment of teacher time, on top of the additional cash investments to put on the training.

It is critical that the district include cost estimates for the staff's time investments in addition to the cash investments made for the program.

Because the purpose of this cost analysis is guiding management decisions rather than compliance, cost accounting for A-ROI analyses must also count staff time invested in the program as part of the total cost.

A suggested method for estimating the cost of time investments are outlined below, although there are a number of ways to estimate this value:

1. Identify all positions (if using average salaries) or individuals (if using specific salaries) that will invest time in the program
2. For each position or individual, find the total annual compensation, including salary and benefits
3. Calculate their hourly rate by using the formula: (annual compensation) / (contract days*hours per day)
4. Use the data collected on total time spent on the program for this position or person and multiply their time invested by the hourly rate



Cost Categories	Start Up Expenditures	Ongoing Expenditures
Staff Time on Direct Service	<ul style="list-style-type: none"> N/A 	<ul style="list-style-type: none"> Reading teacher time on direct instruction of the program Special education teachers time on direct instruction of the program Paraprofessional time supporting students in the program
Staff Time on Training	<ul style="list-style-type: none"> Teacher time for kickoff training 	<ul style="list-style-type: none"> Teacher time for annual training session Teacher time for monthly meetings to evaluate progress of program
Administrator Time Dedicated to Program	<ul style="list-style-type: none"> Administrator time for kickoff training 	<ul style="list-style-type: none"> Administrator time for monthly meetings to evaluate progress of program
Fees and Stipends	<ul style="list-style-type: none"> Read180 training fees Driving travel reimbursement Substitute teachers to cover the classes of teachers in the training 	<ul style="list-style-type: none"> Read180 training fees Driving travel reimbursement Substitute teachers to cover the classes of teachers in the training
Materials	<ul style="list-style-type: none"> N/A 	<ul style="list-style-type: none"> N/A
Operations	<ul style="list-style-type: none"> N/A 	<ul style="list-style-type: none"> N/A

Even if the data cannot be found, it is almost always possible to create a reasonable estimate to represent the costs incurred.

When hard data and records are not available, there are a variety of ways to calculate costs with reasonable accuracy, even with imperfect data. Unlike in budget and reporting documents, it is sufficient for an A-ROI analysis simply to have a “close enough” estimate of costs, even if every dollar and cent is not tracked. The goal is to understand how much of different resources are being used to try to achieve certain student outcomes. Even when there are no data available about certain costs, it is still better to estimate them than to leave them out entirely. The strategy for estimating depends on the type of expense:

- **Time spent on instruction:** If the time is not tracked, estimate the time spent in a typical day or typical week (using interviews with teachers to inform the estimate), then multiply by the number of days or weeks in the school year.
- **Time spent on meetings or ongoing training:** Use a school calendar or log of meeting agendas to count the number of meetings that occurred in a year. If they were not tracked, estimate how often the meetings occur (e.g. every three weeks) and how long they typically last.
- **Planning time:** Only count the paid time spent on planning. Begin by estimating the total amount of planning time for a teacher, taking the planning time provided in one day and multiplying by the number of school days. Then take some percentage of that to represent the portion of

planning time that is spent on activities related to the program (this estimate is informed by teacher interviews, and could be as high as 100%).

Including staff time investments will offer a truer picture of the program's true costs, but a district will need to decide on a case-by-case basis whether the time investments represent costs that can be repurposed.

If a district decides to reduce class sizes in grades K-2, the biggest investment will be an increase in staff devoted to teaching those grades. If a district initiates a co-teaching program, the additional special education teachers will be a major investment. There are many other examples of programs where the staff time involved is one the largest investments, if not the largest.

Because of this, it is important to factor in costs for the time that staff is spending on the program. This should include all time staff spend on the program, including core instruction, extended time, additional staff support and professional development time. However, once all these costs are captured, the district may decide on a case-by-case basis to treat some costs as fixed, rather than as incremental costs triggered by the program.

For example, when a district purchases a new reading program, the total investment in implementing the program includes the cost of materials and training, but also the hours that teachers spend in professional development – and even the time teachers spend teaching the curriculum. When viewed with this lens, the reading program looks like a very large investment by the district. It is important to present this view, so that leaders have a true view of the time and funding invested in reading. Nonetheless, the district could reasonably decide to treat the core instructional time as a fixed investment, since the teachers would spend this time teaching reading regardless of the curriculum or program. The final reported cost of the program would include time spent in PD, but not the core classroom time.

As a guideline, there are a few types of staff time investment that should usually be treated as incremental and counted in the total cost of the program:

- Extended time, either by extending the whole school day or dedicating a greater portion of the day to a particular topic
- Staff beyond the general education teacher providing core instruction (e.g., paraprofessionals, social workers, special education teacher, etc.)
- The program explicitly calls for class size reduction (which triggers an increase in staffing)



Tool 6.1: Cost Map

Use the table below to articulate the different types of costs associated with the program. The goal at this stage is simply to list out the different types of expenditures, NOT to list the actual amounts.

Cost Categories	Start Up Expenditures	Ongoing Expenditures
Staff Time on Direct Service		
Staff Time on Training		
Administrator Time Dedicated to Program		
Fees and Stipends		
Materials		
Operations		



6b. Collect data and estimate costs

It is important that every cost listed in the Cost Map has an estimate attached to it, even if it requires collecting or estimating data that is not easily available.

Some of the costs may be very simple to estimate. Materials costs or trainer fees might be collected from an invoice, stipends might be clearly tracked and labeled in the budget system, and transportation costs might be billed from the operations department.

However, the main budget software might aggregate some data in a way that makes it difficult to collect costs with the level of granularity necessary for A-ROI analysis. To calculate all of the cost items, it may be necessary to use multiple data sources and estimation techniques, including:

- Data from main budget software
- Substitute teacher records or new positions created from HR software
- Transportation or facilities costs from the operations department
- Stipends for extra instructional time or PD sessions from grants
- Time investment estimates from interviews and schedule requests

Example Obtaining accurate time estimates is an important part of the costing process, and can be a new way of thinking in many districts. The example below illustrates possible approaches to estimate the costs of professional development for a technology-based math program.

Backward-Looking The professional development stipends were recorded with very little description of what type of training was associated with each expenditure. Conducting interviews with program participants, however, might help identify the dates of the trainings and attendance estimates. This data could make it possible to match up each of the program training sessions with the stipends paid out for them and also to estimate the cost for staff time spent at the trainings.

Forward-Looking The costing process for a forward-looking analysis should be simpler. It might include setting up a system for tracking teacher and administrator attendance at each training and a program code in the budget software to track stipends related to the program of interest. With this process set up, the data should be easily accessible in the future.

Each expenditure should be labeled with additional information to allow cost analysis based on category, time, and location.

A-ROI analyses can provide a deeper understanding of a program's cost-effectiveness if additional data about each cost is collected, for example:

- Date and/or program year
- Location or site
- Category (e.g., textbooks, PD, etc.)



Typically, additional information about time investments will help accurately estimate their costs:

- Attendees of the event and their positions
- Average or specific compensations of each position (salaries and benefits)
- Length and frequency of the events

These pieces of additional data for each expenditure will help a district understand how much of the costs have already been spent, how much of the costs it will incur by continuing the program, and if there is any variance in costs for the program across schools or years.

DMGroup Tool

See the A-ROI Workbook to find an example of how to set up the cost data collection tools.



Example

For example, imagine a reading program in its second year. The program involved a significant amount of upfront professional development, continuing professional development, and materials costs. Examples of tables that could be used to collect the cost data are shown below.

For cash investments:

Expenditure	Date	Year	Category	School/Site	Annual Cost (\$)
Annual Training Consultant Fee		2013	Professional Development Fees	Bissell Elementary	\$ 1,750
Travel to Annual Training		2013	Staff Travel	Bissell Elementary	\$ 500
Annual Training Consultant Fee		2014	Professional Development Fees	Bissell Elementary	\$ 1,750
Travel to Annual Training		2014	Staff Travel	Wilcox Elementary	\$ 500
Annual Training Consultant Fee		2015	Professional Development Fees	Wilcox Elementary	\$ 1,750

For time investments:

Activity/Meeting/Event	Date	Year	Category	School/Site	Position	Frequency Per Year	# of Attendees	Length (Hours)	Hourly Rate	Annual Cost (\$)
Annual Teacher Training		2013	Professional Development	Wilcox Elementary	Teacher	1	25	4	\$ 41	\$ 4,135
Annual Teacher Training		2014	Professional Development	Wilcox Elementary	Teacher	1	25	4	\$ 41	\$ 4,135
Annual Teacher Training		2016	Professional Development	Wilcox Elementary	Teacher	1	25	4	\$ 41	\$ 4,135

Armed with this level of information about the costs associated with the program, the district is in much better position to fully understand what has been invested and will have more efficient processes in place to track this type of data in the future. By tagging each expense with the additional information identified in the table, the district will be able to identify variances between schools and years to identify trends in how resources are being invested in the program.



6c. Calculate the total and per-student cost

While more detailed cost calculations can be saved until Step 8, at this stage it is a good idea to calculate a rough estimate of the total cost and total cost per student.

Once all the costs are mapped and estimated, sum the individual items to obtain an estimate of the total cost of the program. Be clear about whether this cost estimate includes only costs already spent, or also includes costs projected into the future.

To find the per-student cost, estimate the number of students for the time period covered by the cost estimate. If the costs include future projections, then the number of students should include projected number of students served in the future as well.

Capturing the cost per specific student will allow a district to compare the cost and cost-effectiveness of different implementation approaches within a program.

Additionally, collecting the data to calculate the cost *per specific student* can provide insight into how effective the program was with varying duration, group size, or other implementation factors. For instance, programs that have students entering and exiting over the course of the school year or are implemented with different service levels across the district will have costs that vary between students. For more information on duration see Step 8; for service levels, see Step 5.

Mistakes to avoid

When Collecting Costs Data:

- Only counting dollars spent on materials and stipends, and neglecting the investment of time spent on training, delivering instruction, planning and other related activities.
- Only collecting data that is in the budget or student information system rather than all of the data that is necessary to provide a complete cost of the program.
- Not collecting additional information, such as location of expenditure, with each data point.
- Leaving the cost of personnel benefits out of staff costs data collection.
- Asking someone who does not have a deep understanding of the intent of the analysis and the structure of the program to pull cost data.

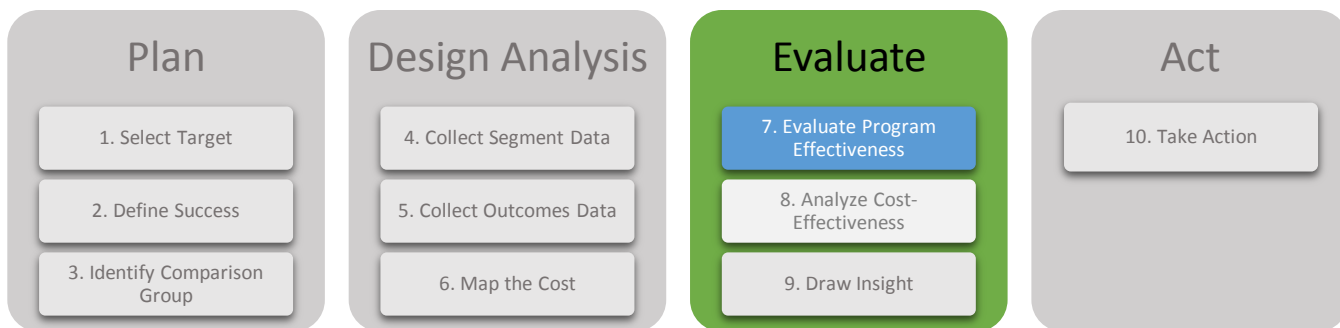
During a Forward-Looking Analysis:

- Not creating a process to track cash and time investments over the course of the analysis on the program level, school level, and even classroom level (when possible).

During a Backward-Looking Analysis:

- Deciding to conduct a backward-looking analysis and realizing the costs data cannot be disaggregated in a way that allows the district to analyze costs on a program level.





Step 7: Evaluate Program Effectiveness

To determine the true impact of a program, the analysis must take into account and control for external factors as much as possible. A car company is analyzing the efficiency of different approaches to producing the same car model across its different factories. It finds that the most productive factory is significantly more efficient than the others and, at this factory, the cars are assembled in a slightly different order. The CEO decided to have all of the factories switch their methods to emulate the most efficient factory, in an effort to increase the productivity of the company as a whole. However, even after the other factories shifted their practice, very little increase in performance was realized. The CEO had believed the efficiency was due to the order of assembly, but did not account for the skills of the workers and leadership of the plant managers, which was the driving force behind efficient car production. A-ROI analysis, when executed effectively, is a way to control for the impact of external factors on the outcomes of the program being analyzed.

Key Outcomes

- After evaluating the program's effectiveness, a district will understand:
 - If any segments benefited from the program
 - If students in the program received a greater benefit than students who were in a different program or did not receive services
 - If any particular segments performed significantly better than others
 - If the implementation of the program, including duration and levels of service, effected the outcomes for each segment

Disclaimer: This section is not a substitute for a textbook on statistical analysis. The steps below serve as a guide for someone who already has the skills to calculate means and standard deviations, perform simple regressions and interpret the results. Beyond the statistical skills, the analyst must also have a thorough understanding of the program and its goals, or work with someone who does.



7a. Assign students to segments and determine level of service

The data collected in Step 4 can be used to match each student to a defined segment.

Many districts look at test scores broken down by student sub-groups as defined by No Child Left Behind (NCLB), such as race, disability, and poverty status. However, these groupings are overly broad and within each group there will be students who respond very differently to specific programs and interventions. To assess which students benefit most from a program, it is important to use segments that group students according to characteristics relevant to the program. Step 4 of this manual provides guidance on how to create segments based on relevant characteristics.

Once these segments are determined, the available data should be used to assign students into segments, for both the students who received the program and those who did not. This will enable analysis of outcomes on a segment-by-segment basis, to determine how the outcomes of similar groups of students are impacted by the target program.

The data set is ready for deeper analysis once every student in the data set has both outcomes data and a segment.

Example

One way to approach assigning students to segments can be seen in the A-ROI Workbook, on tab “3. Outcomes & Student Data (Ex),” which provides an illustrative example for a reading intervention program. Data was collected both for students who are in the program and for students not in the program. The hypothesis was that students who only struggled with the content knowledge of reading—in other words, students with high engagement but struggle with reading—would benefit most from this program. To test the hypothesis, the district chose to segment students further based on the type of struggles they experienced in school. Using this approach, the district created segments based on the following criteria, with the assumption that all students in the dataset struggle with reading:

- Struggle w/ Engagement, Reading, Math: students with reading and math ability scores below the median, with low engagement
- Struggle w/ Reading: students with reading scores below the median, with high engagement and with math scores above the median
- Struggle w/ Math and Reading: students with reading and math scores below the median, but with high engagement
- Struggle w/ Engagement: students with low engagement, but who are not lacking ability (e.g., high test scores, but low grades)

The implementation data can be used to match up students with the level of service they received.

If the implementation data collected includes class size and frequency of service provided, a district can gain greater understanding of the relationship between the level of service provided and the outcomes achieved. This is important because if the district determines that greater investment could lead to greater outcomes, it might be beneficial to increase the investment for the program. Likewise, if the



level of service delivery does not influence the level of outcomes, then greater investment will not better serve students.

Example

A district might find that a student segment consists of ELL students with low rigor previous education. This segment has 100 students. Of these 100 students, seventy of them received reading instruction in groups of six for five hours per week. The other thirty students received reading instruction in groups of three for five hours per week. Matching each individual student to the level of service he or she received allows the district to analyze the cost-effectiveness of the program even within a segment.

Use the fidelity check data that was collected to determine if the program is being implemented in the way it is intended.

It is important to analyze how closely the program's actual implementation aligns to its intended implementation. For instance, if a district is analyzing a reading intervention curriculum that calls for 90 minutes of instruction each day, the district might want to organize random walk-throughs of classrooms to determine fidelity of implementation, checking for two specific indicators: if the curriculum is being used at all and if the curriculum is being used for the appropriate amount of time prescribed. If a program is not being implemented with fidelity and is found to be ineffective, it is possible that the poor implementation led to poor results, not necessarily the program itself.

The district might find that the curriculum was only used consistently in one-quarter of classrooms that were supposed to be using it. If that is the case, the district may determine that no conclusion can be made about the effectiveness of the program, so the key finding of the analysis is that the program is not being implemented effectively.



7b. Analyze the effect of the program on specific segments

Compare results for students within the same segment, so that the key difference is whether the student was in the program or not. Comparing across segments does NOT demonstrate the effect of the program.

It is crucial to analyze the difference in outcomes between students who received the program and students who did not, *within the same segment*.

Example

For example, if a dropout prevention program is believed to benefit students who hold a job outside of school, then the analysis should compare the dropout rate between the following two groups, where students with jobs is a segment:

- Students with jobs who participated in the dropout prevention program
- Students with jobs who DID NOT participate in the dropout prevention program

This is different from comparing the effect of the program across different student segments, e.g. comparing dropout rates for students with jobs to dropout rates for students without jobs. Though this analysis could be helpful, as it provides information about what types of students are most likely to drop out *in general*, it does not provide information about the effects of the dropout prevention program.

Setting up the analysis in a way that compares students across segments introduces variables that may influence the outcomes but are not part of the program. For instance, imagine that after comparing the dropout rates of students with jobs who were in the program to students without jobs who were in the program, a district found that the students with jobs were significantly less likely to drop out of high school. There are two plausible explanations that are not at all related to the program that this cross-segment analysis cannot rule out:

- It is possible that the types of students that seek out jobs are also the types of students that are likely to be perseverant and get back on track to graduate high school, leading to a lower dropout rate.
- It is also possible that having a job is the influence that made the students more likely to graduate high school, not the dropout prevention program.

Use the measures that were initially chosen in Step 3 when setting the benchmark for success.

A rigorous program evaluation identifies the key measures at the outset, so that there is a clearly agreed-upon measure of success. The analysis at this stage should compare the average measure for students who received the program to the average measure for similar students in the comparison group who did not receive the program.

Depending on the team's level of sophistication with statistics, this analysis could take many forms. A simple comparison of averages tells part of the story, showing how the program impacted outcomes for a given segment of students. This analysis does not address questions of statistical significance, but it does give an initial indication of the results in each segment.



Example

The table below shows how a (simplified) analysis might look for a program designed to raise reading scores.

Segment	Average not in program	Average within program	Impact due to program
Segment A	54%	58%	+4%
Segment B	14%	24%	+10%
Segment C	67%	62%	-5%

A more robust analysis should also take into account the sample size within each segment and standard deviation (if scores are spread over a range) in order to determine whether the results can reasonably be attributed to the program itself, or occurred due to chance.

Be prepared for the possibility of finding zero impact or very small impact from the program.

Sometimes districts will conduct an analysis, only to find that the program did not produce results that were significantly better or worse than the alternatives. In other words, the program is not driving results, positively or negatively. It is important to remember that this is a very reasonable, insightful, and actionable result of the analysis. If a district finds that the student outcomes for a program are not significantly different than the alternative, then it knows that the program is not a comprehensive solution to students learning a specific skill. In this case the district must investigate why the program is not delivering the expected results or reevaluate its strategy to serve students with that particular need.

Control for external effects that might influence the outcomes of the program, such as variance in teacher or principal effectiveness.

Oftentimes the outcomes of a program are influenced by external factors unrelated to the program design, such as varying levels of teacher effectiveness or principal capacity. It is important to structure the analysis to account for the influence of external factors. One simple strategy to control differences in school leadership is to check for growth in student outcomes in a subject unrelated to the program. For example, if a district is analyzing a math program and finds growth in math scores at the schools implementing the program, at first glance this seems to indicate an effective program. On the other hand, if the ELA scores at these schools show similar growth, that is an indicator that external factors (e.g. principal effectiveness) might be driving improvement at the school across all subjects, and that the math program may not be the cause.



7c. Investigate other possible segmentations

If the initial segmentation shows no difference in outcomes across the students segments, it may be worthwhile to re-assess the segmentation.

The initial analysis in Step 7b used the segments developed in Step 4, but these segments were essentially an educated guess. Now that data has been collected, cleaned and organized, it is possible to use statistical approaches to identify a meaningful segmentation of students.

Once the data about outcomes and student segments have been loaded into a statistical analysis software package, someone familiar with the software should take the following steps:

1. Initial check of the effect on the hypothesized segment (Step 7b)
2. “Guess and check” approach to test effect on other segments
3. Use of regression or other statistical methods to determine factors that correlate with success in the program

“Guess and check” approach should be based on logical links between student characteristics and program design.

If the initial segmentation showed roughly equal impact across segments, then it may be possible to identify a more meaningful segmentation. The “guess and check” approach is just what it sounds like. The analyst working with the data may have a reason to believe that students with certain characteristics could have better outcomes, so he/she performs the analysis to check. A regression can be helpful to identify characteristics that correlate with better outcomes.

Keep in mind that when analyzing a number of student segments it is possible to find “false positives” in the data, so it is important to do a logic check to make sure the attributes used to construct the segment are meaningful for the program.

Statisticians have learned that as the number of distinct segments increases, the probability that the results show a segment with significant outcomes due to random chance increases as well (this is called the Multiple Inferences Problem). This means that, especially if a district analyzes many different segments, it is possible that the analysis will find one segment that experienced a substantial effect from the program, when the variance in outcomes was actually driven by random chance. A similar phenomenon occurs when the sample size of a segment is particularly small, whether due to creating many segments or having a small population of a particular student type.

Example

One district found that students with last names starting with a letter in the first half of the alphabet experienced greater effects from a dropout prevention program than students whose last names were in the second half of the alphabet. The district determined this finding to be a false positive, because the last names of students do not have a logical link to whether they drop out of high school or not. Rather, the district continued to analyze the data and found that the program worked significantly better for students with higher attendance rates, which had a much more logical link to whether students dropped out.



Tool 7.1: Segment Identification Process

1. Test Initial Segmentation

Conduct an analysis to determine if the initial segments created in Step 4 proved to be meaningful.

3. Run Regression Analysis

Especially if no meaningful segments have been found, it can be useful to conduct a regression analysis to identify any significant correlation between potentially key variables and outcomes. If trends are identified, return to Guess & Check to confirm the importance of the new segments.



2. Guess & Check

Use the remaining segment data collected to analyze if the program had meaningful effects on other segments of students.

4. Draw Insight

Use the three types of analysis to draw insight into which student segments benefited the most from the program, if any.



7d. Define the relationship between investment level and outcomes

Another important analysis is to determine if students who received higher levels of service experienced greater effects from the program than students who received lower levels of service.

It is possible that a program provides students with a variety of service levels. Some programs might produce better outcomes when they are implemented with smaller group size or more time per week, but some may not. This variance could be unintentional, arising naturally from schools or classrooms making different decisions or it could be intentional to provide certain students with more intensive services from the program. Creating intentionally varied levels of service within a segment could be an effective method to evaluate whether the level of service effects the outcomes of the program. (See Step 3 for strategies on how to design the program to generate more conclusive results.)

An analysis of outcomes and levels of service will be more likely to provide conclusive results if it is performed within a segment.

Example Imagine a reading intervention program that serves a variety of student segments, including ELL students with low rigor of previous education. Within this student segment, thirty students receive the program in groups of three students and seventy students receive the program in groups of six students for the same amount of time per week.

Number of Students	Group Size	Avg. Time Per Week	Outcomes
			(Compared to Students Not Receiving the Program)
30 students	3	5 hours	+6%
70 students	6	5 hours	+6%

In the scenario described above, there is no discernible relationship between an increased investment (e.g., smaller group sizes) and increased student outcomes within the student segment.

An analysis of outcomes and levels of service will be less likely to provide conclusive results if it is performed across segments.

Example Another example is a math intervention program targeted at students in the lower quartiles of math ability. Built into the program is a stipulation that services are more intensive for the lowest ability students. The variance in service delivery and outcomes is outlined below:

Student Segment	Group Size	Avg. Time Per Week	Outcomes
			(Compared to Students Not Receiving the Program)
First Quartile, Math	5	5 hours	+18%
Second Quartile, Math	10	2.5 hours	+8%

The results from this program indicate that there might be a correlation between increased investment and increased outcomes, but it is impossible to tell whether the increased outcomes are due to the more intense intervention, or simply due to the students' starting in the first quartile.

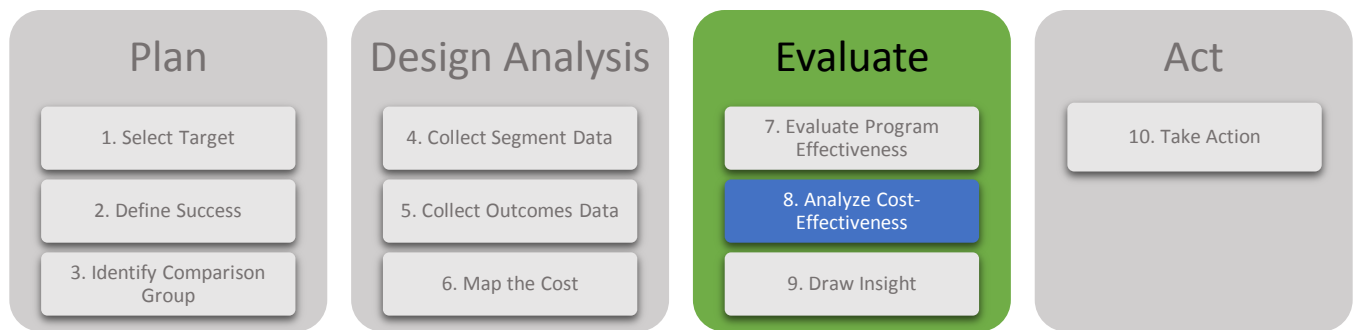


Mistakes to avoid

While Evaluating Program Effectiveness:

- Trying alternate analysis approaches to show the program had an impact, when the results seem to indicate no impact. “No impact” is an important finding.
- Analyzing too many segments at once, which could lead to a false positive.
- Analyzing segments that are based on attributes without a logical connection to the outcomes of the program.





Step 8: Analyze Cost-Effectiveness

Part of understanding what does the most good for students in a context of limited resources is the cost-effectiveness of a program. A pharmaceutical researcher might be testing the effectiveness of two different vaccines for the flu. A major hospital network is planning to buy one of the two vaccines. The researcher finds that Vaccine A is slightly more effective than Vaccine B, but it is twice as expensive. This means that by choosing the cheaper treatment, Vaccine B, the hospital could provide services of a similar quality to twice as many patients for the same cost. Although it is slightly less effective, Vaccine B is much more cost-effective, allowing more patients to receive the treatment. In a world of limited resources, both effectiveness and cost-effectiveness must be incorporated in the decision-making process. The situation in public school districts is similar. For instance, Reading Recovery, although it has seen a high-level of effectiveness in some districts, is quite expensive on a per-student basis, and can typically only be provided to a small number of students even if many students could benefit from extra reading instruction.

Key Outcomes

- After completing the cost-effectiveness analysis, a district will understand:
 - How much the district has invested in the program already
 - How much money the district could recover if it discontinued the program
 - The variance in costs across schools and/or segments and how the implementation of the program relates to that variance
 - The relationship between increased investment and outcomes

Disclaimer: If the program was shown to be ineffective for all student segments in Step 7, it might not be necessary to conduct a more detailed cost analysis. The total cost (calculated in Step 6) is the most important value for programs that do not sufficiently influence outcomes for any students. Programs that are considered ineffective typically should not be continued, so only the costs that have been already paid are relevant to decision-making, as it can help districts understand how much was actually invested in the program. In the future, district leaders can use this information to inform decisions about spending on similar programs.

8a. Calculate costs at a more detailed level

To obtain a clear understanding of costs, the district can use the data it collected to analyze the costs on a variety of different levels, including total cost and per-student cost.

The two primary costs are:

- **Total cost:** includes all cash and time investments over the lifetime of the program
- **Per-student cost:** captures the value of the resources that the district is investing in a specific student during a given time period to increase achievement

The total cost can be calculated by adding all of the costs for the program, both past and future. This number can help a district understand the scope of investment it is making in a specific program.

The per-student cost can be calculated by finding the total cost and dividing it by the estimated number of students served over the lifetime of the program (past and future). The per-student cost provides insight into how much it costs to serve each student in the program, as it is structured currently, and is valuable for comparisons with other programs.

There are additional ways to analyze costs that could provide insights into how resources are used to implement a program. It is important to view the cost of the program from different perspectives to identify variance in cost that could be connected to variance in outcomes. Some possible ways to view the costs are:

- Total vs. per-student
- Sunk costs vs. future costs
- By school
- By program year
- By category of expenditure
- By cash investments vs. time investments

DMGroup Tool

See the A-ROI Workbook to find examples of how to analyze the program's cost.

If there is wide variance in cost across schools or years, that could be an indication that the implementation was inconsistent. This could skew the outcomes of the program, and should be investigated further to ensure that the program was implemented with fidelity.

When deciding whether to continue a program, it is critical to distinguish between sunk costs and future costs.

Sunk costs are costs that, at the point of the analysis, cannot be recovered, refunded or repurposed by the district. This includes past professional development to support the specific program, teacher time for instruction that has occurred, and any materials that have already been purchased.

Future costs are costs that can be repurposed by the district in the future if the program is discontinued. These include planned professional development or staff collaboration time for the program and anticipated purchases of materials.

When facing decisions during the analysis of a program, there are two important questions to ask about the cost of the program:

Question	How to answer
Was the investment in the program a good decision? Should I fund a similar program in the future?	Look at the total lifetime cost of the program, or total per student cost, and compare to other available options
If the program is ineffective, how many resources could the district repurpose if the program was discontinued in the future?	Look only at future costs (neglecting sunk costs), to determine the funds that would be freed up if the program were canceled.

Understanding the sunk and future costs of the program will provide insight into the first question, while understanding the projections for future costs will shed light on the latter question.

Example

Different programs have different cost structures, so disaggregating the total cost by year can provide meaningful insight. For instance, a technology-based math curriculum might require significant startup cash investment for iPads and professional development for teachers on differentiating instruction with technology. After those startup costs are incurred, however, it will only require a small amount of ongoing staff time for teacher training and collaboration.

Technology-Based Math Program, Year 2 (\$2,640,000 total)

	Startup	Year 1	Year 2	Year 3	Total
Future	\$0	\$0	\$50,000	\$50,000	\$100,000
Sunk	\$2,490,000	\$50,000	N/A	N/A	\$2,540,000

Contrast that with a multi-year targeted professional development program around the Common Core. This professional development program includes significant costs both at the start of the program and each year, as every middle school teacher in the district is required to attend a two-hour training and a PLC meeting each month with a focus on the Common Core.

Common Core Professional Development Program, Year 2 (\$2,640,000 total)

	Startup	Year 1	Year 2	Year 3	Total
Future	\$0	\$0	\$800,000	\$800,000	\$1,600,000
Sunk	\$240,000	\$800,000	N/A	N/A	\$1,040,000

It is likely that the decision-making process following the cost analysis of these two programs would take a different path. Because nearly all of the costs are recoverable in the future, the district could repurpose significantly more funds by discontinuing the professional development program than the math program even though the programs' total costs are equal.



8b. Analyze cost-effectiveness

Cost-effectiveness has two main components: general program effectiveness and cost-effectiveness compared to alternatives.

The district should begin analyzing the cost-effectiveness of the program by asking two questions:

- For which students was the program effective?
- Was the program more cost-effective than the alternatives?

After evaluating the program's effectiveness for each student segment in Step 7, a district will have an understanding of the answer to the first question: whether the program worked and for whom.

To answer the second question, the district can gain insight into the tradeoffs by comparing the program in its current form to its alternatives. Alternatives could be other efforts that serve similar purposes, variations on the level of service of the same program, or not providing the service to students.

Analyzing the cost-effectiveness of the program and its alternatives is a chance for the district to evaluate whether the investment in the program was worth the outcomes produced.

Calculating the per-student cost of different implementations of the program can inform district decision-making about the best format of the program to use going forward.

Districts can use data about how services were provided to students to discover relationships between the cost and the effectiveness of the program. Elaborating on the student categorization in Step 7d, the district can calculate the cost differences based on variance in levels of service to individual students.

Example

Imagine analyzing the cost-effectiveness of the same reading program for the segment of ELL students with low rigor previous education from Step 7d. The program effectiveness analysis indicates that students in the program experience greater growth than similar students not in the program. Upon further cost-effectiveness analysis, the district finds significant variance of implementation across schools but no correlation between higher spending and better outcomes.

Group Size	Time Per Week	Avg. Cost Per Student Per Week	Outcomes
			(Compared to Students Not Receiving the Program)
3	5 hours	\$55.00	+6%
6	5 hours	\$27.50	+6%

*Avg. Teacher Compensation: \$50,000/year (\$33/hour)

The district finds that the variance in service levels arose because one middle school implemented the program with an average group size of 3 students, while another middle school had an average group size of 6. The data indicate that a doubling of the investment in the form of smaller group size does not improve overall outcomes for students. Because this was a comparison of students within a segment, the students in each group have similar characteristics, so it is reasonable to conclude that there is no relationship between increased investment and increased outcomes.



This type of understanding of cost-effectiveness allows the district to discover potential savings in the way the program is implemented without decreasing the quality of the services to students.

Example

In some cases, the program may be intentionally designed to provide different services to students, according to their ability level or need. For example, a math intervention for elementary students might provide support for 5 hours per week to students with scores in the 1st quartile, and only 2.5 hours per week to students in the 2nd quartile.

In this situation, it will be difficult to determine how the amount of time provided to students is linked to the outcomes achieved, because the students receiving 5 hours are quite different from those receiving 2.5 hours (different incoming math level). This illustrates why it would be preferable to compare students within the same segment receiving different interventions: to show a clear connection between the intervention provided and the resulting outcomes.

If the analysis generated results like the example in the table below, there may be many questions about what led to the difference in outcomes. Did the extra hours make a difference? Did the lower-level students show more growth because they were starting with lower scores? Did the program target skills that were more appropriate to students in the 1st quartile? Did group size have an effect? The analysis does not provide a conclusive answer.

Segment	Group Size	Time Per Week	Avg. Cost Per Student Per Week	Outcomes
				(Compared to Students Not Receiving the Program)
First Quartile, Math	5	5 hours	\$66.00	+18%
Second Quartile, Math	10	2.5 hours	\$16.50	+8%

*Avg. Teacher Compensation: \$50,000/year (\$33/hour)

Though many questions are still unanswered, the results do seem to indicate that students in the program did better than similar students who did not receive the program. With these promising initial results, the district may want to consider designing a more careful test to determine the effect of group size and time provided. For the next year, they could randomly assign students in the 1st quartile to two different implementations of the program, and then analyze whether the implementation had an effect on student outcomes. This is an example of how A-ROI analysis becomes part of a cycle of continuous inquiry and improvement.

If students can take part in the program for differing amounts of time based on achievement, then tracking the amount of time spent in the program (duration) is key. Also, the cost to serve a student will differ by student based on the period of time in the program, and the calculation should be adjusted to reflect this.

Duration can be a meaningful measure for intervention programs if students are placed into and taken out of them based on test scores. If all students remain in a reading intervention program until they achieve grade-level reading ability, then the duration in the program is the most meaningful measure of effectiveness. If one segment averages 7 months in the program and another segment averages 2



months, the program was more effective for that second segment. For programs that students are placed in for a set amount of time, such as a full school year, duration will be a less helpful metric of success.

To calculate cost using duration for individual students, use the total cost of the program to find its monthly cost per student. Then, use the data about when individual students enter and exit the program to find how many months they received services. When multiplied together, these values should provide a cost estimate of the services students actually received, rather than allocating costs to students evenly across the program.

Duration of Services and Costs

Specific Student	Duration of Receiving Services from Program	Actual Cost of Services
Student A	x2.0 months	\$250
Student B	x1.0 months	\$125
Student C	x0.5 months	\$63

*Cost per month: \$125 per student

Comparing the sunk and future costs could influence decision-making about the future of the program.

Different cost structures of programs, in some instances, may affect how a district decides to move forward with the program. If future costs were not included in the initial cost estimation (Step 6c), now is the time to estimate the future costs of the program.

Some programs, particularly those that have already begun, have significant upfront costs and low recurring costs each year. If the district has already invested most of the costs, the district might feel less urgency to end the program if it is producing student outcomes that are comparable to its alternatives.

Other programs may have low upfront costs but require significant time investment from staff each year. In this case, the district might want to reallocate the staff time to programs that are producing greater results for students, in which case there would be greater urgency to reduce the program.

Example

In an ideal world, a district would eliminate programs that are found to be ineffective for students, while only keeping programs that produce significant student gains. However, this can be very difficult in practice. For instance, imagine that the technology-based math program from the example in Step 8a produced 3% growth for students in the program, and students who were not in the program averaged 5% growth. By most measures, a program with these results would be considered generally ineffective.

Technology-Based Math Program, Year 2 (\$2,640,000 total)

	Startup	Year 1	Year 2	Year 3
Future Costs	\$0	\$0	\$50,000	\$50,000
Sunk Costs	\$2,490,000	\$50,000	N/A	N/A

	Growth Year 1
Students in the program	3%
Students not in the program	5%

However, once the district has found these results, it faces a difficult decision. Should the district repeal the new program and revert to the old program for all students? Should the district investigate why the new program isn't working as expected, and continue for another year while fine tuning the practices? Or should the district keep both programs in place for now, while searching for a more effective practice that could replace both in the future?

Different districts might make different decisions, depending on the context. The decision will depend on how difficult or demoralizing it might be for teachers to switch programs again, after already going through the challenges of implementation.

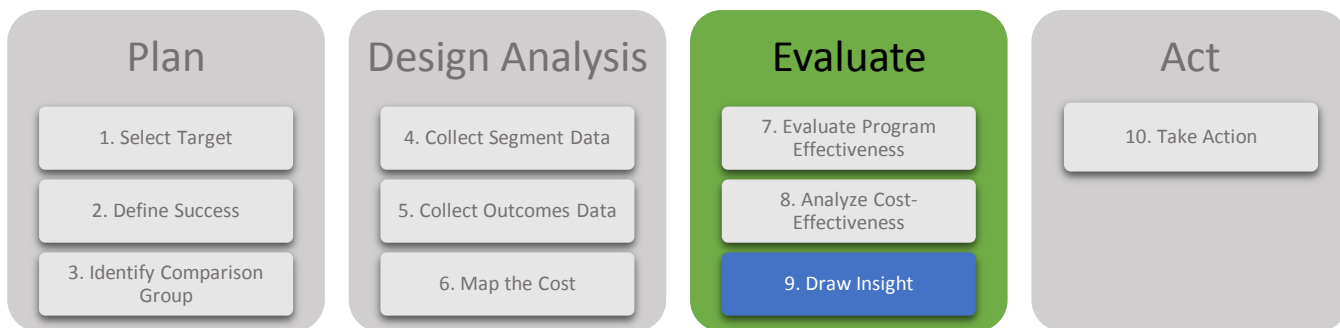
Regardless, the analysis still produces two useful findings for the district:

- The district still needs an effective approach to teaching students math, since neither of these programs are effective in their current form.
- Investing in this program did not deliver the expected results, so if the district sees a similar program proposal in the future, the leadership should be cautious about believing the promised results.

Mistakes to avoidWhen Analyzing Cost-Effectiveness:

- Comparing the cost-effectiveness of programs with different outcomes measures.
- Analyzing the cost and cost-effectiveness only at the school- or general student-level rather than the student segment-level.
- Not distinguishing between past and future costs when determining if the program should be continued or not.



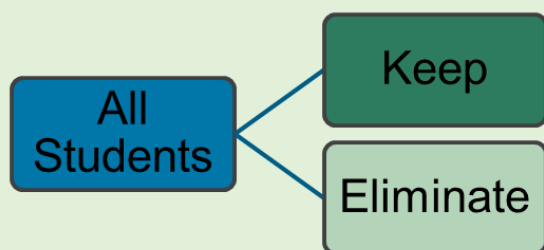


Step 9: Draw Insight

Introduction to Drawing Insight with A-ROI

Many districts currently conduct program evaluation to answer a simple question: should we keep or eliminate this program for our students? A-ROI analysis, by providing a more nuanced understanding of program effectiveness, gives districts the opportunity to be more strategic and creative in determining the future of a program.

Typical Options for the Future of a Program (Without A-ROI Analysis)



DMGroup is introducing two major shifts away from the traditional thinking on how to draw insight based on program evaluation using A-ROI analysis:

1. There are more choices than just keeping or eliminating a program based on the results of an A-ROI analysis.
2. There can be different decisions about the future of a program for different student segments.

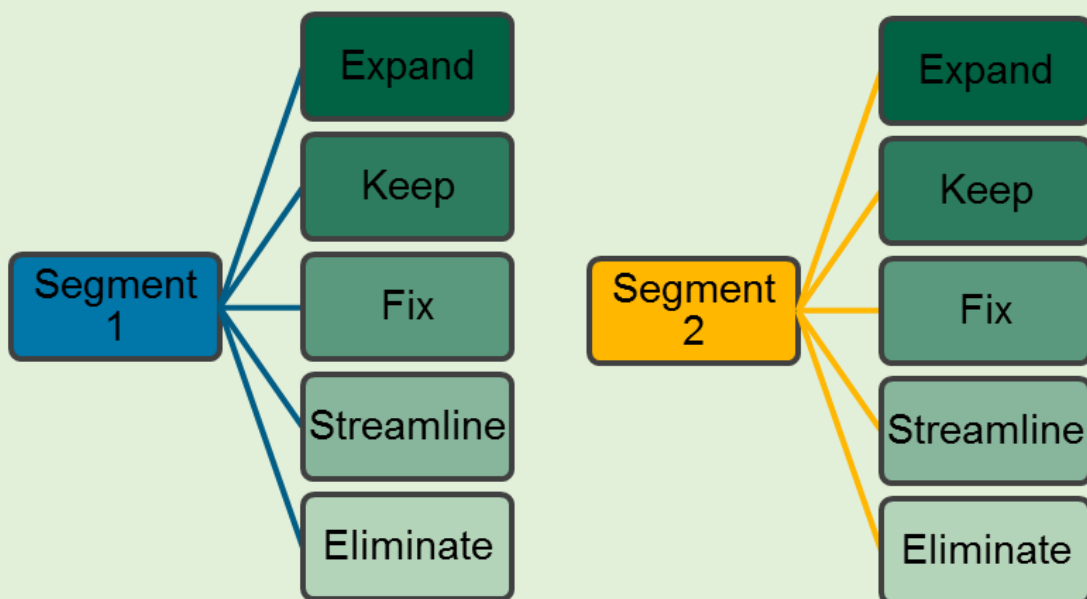
Options regarding the future of a program are outlined below. These options apply not only to programs as a whole, but should be considered with regard to segments within a program.

- **Expand:** if a district finds that the program is particularly cost-effective, it may want to expand the program to all students in the segment and even similar segments if they are not already receiving it.
- **Keep:** if a district finds that the program is cost-effective, but all of the students in the segment already receive the program, the district may choose to keep the program as is. Or, if a program is effective but costly, the district may choose to keep it as, until they find an alternative that is more cost-effective for those students.



- **Fix:** if a district finds that only some pieces of a program are effective, it may seek to fix the ineffective pieces. Particularly if the district finds that implementation was weak or the levels of service were variable, the district may seek to modify the current program delivery without expanding or reducing it.
- **Streamline:** if a district finds that pieces of a program are ineffective, it may seek to cut out those pieces but keep the effective portions. Alternately, if a program is found to be effective but costly, the district may want to reduce the program and explore alternatives that might be more cost-effective or effective.
- **Eliminate:** if a district finds that a program does not work for a certain segment, it may want to stop providing the program to a particular student segment altogether.

Proposed Options for the Future of a Program



Key Outcomes

- After drawing insight from the analysis, a district should understand:
 - For which student segments the program was effective or not
 - For which student segments the program was more effective and cost-effective than alternative programs
 - How the numbers and jargon from the statistical analysis translate into practical implications for the future of the program
 - Recommendations for the future staffing, design, and budget of the program

9a. Articulate key findings from analysis

When communicating key findings, it is important to translate the statistical jargon of the analysis to common language with an emphasis on practical implications.

The precursor to taking action is ensuring that the district leadership and key stakeholders understand the key findings and recommendations from the analysis. To accomplish this, it is important to communicate the key findings in language that is understandable to someone without a statistical background or detailed knowledge of the program. Some key questions can help guide this process and keep the focus on driving to actionable recommendations, although these may not be applicable to all programs.

DMGroup Tool

See “Tool 9.1: Draw Insight Summary” in the A-ROI Workbook to find an example of how to frame the program’s outcomes and costs to draw insight.

Key Considerations

- Was the program effective for any students or student segments? Was the program ineffective for any students or student segments?
- What level of certainty does the analysis provide for each of the findings?
- How does the cost per student affect the outcomes for each segment?
- Were the outcomes or costs variable across student segments?
- Were the services provided as intended?
- Was the student population served aligned with the intended population?
- Did the level of service, such as time per week or group size, influence outcomes?

Example

Key findings for an elementary reading program might look like this:

Segments (Based on Initial Reading Ability)	Effect Size	Level of Certainty	Cost Per Student
1 st Quartile (Low)	Small	High	\$3,500
2 nd Quartile	Large	Medium	\$2,900
3 rd Quartile	Small	High	\$2,900

A key step in drawing insight is transitioning from statistical jargon to an understanding in practical terms. Translating the information in the table above into practical terms might sound like this:

“Key Findings: The reading program is most effective for students who are slightly below grade level average, based on their reading scores. Of these students, some received the services in groups of three, some in groups of six, but there was no discernible difference in outcomes based on the difference in group size. The program does not benefit students who are above grade level or far behind grade level. The program costs approximately \$3000 per student, though it is slightly more expensive for students with the lowest initial reading level (for whom it is not effective).”



Example

Some districts task the data analyst who crunched the numbers with writing the report. He or she is deeply knowledgeable about the statistical methods that were used in the analysis, and perhaps, not surprisingly, it is not uncommon for a program evaluation report written by a data analyst to read as follows:

“The regression analysis indicated that the group “students with behavior issues” has a 0.2 coefficient relative to the control group. However, this finding is not statistically significant, as it has a p-value of 0.13.”

What this actually means is that students with behavior issues performed 20% better than the control group, but something about the analysis did not meet standards for certainty (a p-value of 0.05 or lower). However, the precise statistical language makes it difficult for a layperson to identify the important piece of the finding: that the program is producing much better results than the alternative.

While a p-value of 0.13 might not meet academic standards of statistical significance, it still means that there is an 87% chance that the findings are valid. The report-writing team needs someone who deeply understands how the analysis relates to district priorities and strategies. With this pragmatic voice on the team, the finding for the same program might read:

“The program analysis found that if a student struggled with behavior issues, that student was likely to have better results if they participated in the program, compared to similar students who did not.”



9b. Generate a short list of recommendations for staffing, program design, and budget

The report should provide district leadership with a few actionable recommendations stemming from the key findings to help move toward impactful change.

Translating the key findings into a list of 2-3 actionable recommendations for how to move forward with the program will allow the district leadership to use the results effectively. It provides a clear roadmap for tangible change based on the analysis.

Some districts have program evaluation departments that produce a list of a dozen interesting findings and recommendations for each program analysis, and, in an ideal world, the districts would have the resources to implement all of the recommendations. Unfortunately, given the context of school districts today, it is difficult to implement a plan that has ten or twelve different pieces. It is generally more effective to identify the 2-3 highest leverage areas and make recommendations to change those.

Key Considerations for Drafting Recommendations

Not all programs will require addressing each of these considerations, but districts can use this list as a starting point when drafting recommendations of how to proceed with program design, staffing, and budgeting for the program.

Program Design

- Process to determine which students receive the program (e.g., entry criteria)
- Time per week of services
- Average group size
- Fidelity of implementation of services

Staffing

- Change in the number of students who will receive the program (based on modified entry criteria)
- Change in overall staff based on the change in the number of students receiving the program
- Reallocation of staff to meet the needs of new students in the program

Budget

- Cost or savings based on the change in overall staff
- Cost of additional PD for new staff
- Materials for new students in the program
- Change in the total cost and per-student cost of the program based on the overall change in students in the program



Example

This can be applied to the elementary reading example. Based on the “Key Findings” articulated in Step 9a, a district could continue to make recommendations for the program design and staffing of the program that reads as follows:

“Key Findings: *The reading program is most effective for students who are slightly below grade level average based on their reading scores. The program does not benefit students who are above grade level or far behind grade level. The program costs approximately \$3000 per student, though it is slightly more expensive for students with the lowest initial reading level (but for whom it is not effective).*

Recommendations: *The district should stop entering students from the first and third quartiles of reading ability into this program. It should prioritize placing students in the second quartile into the program, unless a less costly program of similar effectiveness is identified. The change in student enrollment in the program could result in 4.0 FTE available to be reallocated from this program to more high-leverage areas. The staffing change could result in \$200,000 to repurpose.*

Conclusion: *More effective and cost-effective alternatives should be explored for all three quartiles, and, in the meantime, the district should reduce the reach and budget of this program.”*



9c. Publish a report that highlights the key findings

Publishing a report explaining the methodology, key findings, and recommendations can provide transparency, which can help generate buy-in from key stakeholders.

Creating a tangible and accessible report can help districts communicate about the A-ROI process. The report does not have to be a long and detailed account, but should outline in clear language how and why the analysis was conducted as well as the key findings and recommendations. To provide context for multiple audiences, the report should also include an explanation of how A-ROI and the recommendations around the specific program fit in with larger district priorities.

The report should include the following:

- Summary of the key findings and recommendations (1-2 pages)
- Goals of the analysis
- Methods
- Intended outcomes of the program
- Actual results of the program

This is a document that can be presented to district leadership, key stakeholders, and even published on the district website to provide a window into the decision-making process in the district.

It is important to have a multi-disciplinary team develop the report and to highlight recommendations for district leadership at the front of the report.

Statisticians tend to write reports that emphasize statistical methods; business managers emphasize costs; and educators emphasize instruction. The report for an A-ROI analysis needs to be balanced, practical, and in common language. Having a team of individuals with different skill sets will help achieve this balance.

It is especially helpful to have a person or people on the team who possess the following skills:

- Deep understanding of the methods used in the analysis
- Deep understanding of how the district will take action based on the analysis
- Ability to communicate complex concepts concisely and clearly

Do not bury key findings in a comprehensive list of results. Prioritize based on the quality of the findings and the degree of impact rather than presenting a long list of findings and recommendations.

When writing the final report, it can be tempting to share a long list of all the findings for all segments. However, it is important to prioritize. There should be a short list of high-impact findings at the start, focusing on findings that affect the largest student segments, findings that could lead to significant shifts in funding or program design, and findings that demonstrate very large effects.



Mistakes to avoid

When Drawing Insight:

- Only looking for areas where the program worked to identify opportunities for expansion or only looking for areas where the program did not work to identify opportunities for reduction.
- Evaluating the program in a binary “good or bad” framework.
- Using too much jargon when communicating the results.
- Sharing findings with no tangible recommendations.
- Creating a “laundry list” of findings or recommendations.





Step 10: Take Action

Improving outcomes is dependent on the leadership's ability to ensure that practices are improved based on findings from the analysis. A city mayor takes office with a plan to revamp the local government's internet presence, allowing citizens to pay fees and process forms online. He hires a software company to work with his deputy director of technology to develop the online infrastructure for each department. In the mayor's first monthly address he communicated the goals of the project to his constituents. Six months after he took office, the mayor tasked his chief of staff with collecting a status update on the new online tools. His chief of staff discovered that the project was not any closer to being implemented than it was when the mayor took office. How could that be? It was such a priority for the mayor, how did the implementation falter so mightily? The person in charge of the project, the deputy director of technology, had not been in contact with the mayor or anyone from his office in months. She did not have access to the mayor's cabinet meetings. Simply, she did not possess the clout necessary to lead a major change within the organization, even though she was motivated and capable, and the mayor had endorsed the plan. The same is true for A-ROI analysis. A district might establish a team of data analysts to conduct A-ROI analyses, but if they do not have regular access to the leadership and decision-making processes, tangible change will be difficult to accomplish. Once the report is concluded, the superintendent or another senior instructional leader must take on the responsibility for deciding which of the recommended changes to implement and for managing the implementation process. Using A-ROI to improve the quality of services to students must be a priority for the district leadership.

.....

Key Outcomes

- After taking action, a district will have a:
 - Statement that articulates specifically what will change based on the findings of the report
 - Timetable for implementing the changes
 - Clear description of the implications for future decisions on similar programs



10a. Share findings with district leaders

It is important to get a district-level instructional leader to share the findings of the report and communicate the need for action.

Typically, the author(s) of the report will not be in the top tier of the district leadership, so it is important to introduce the report and the recommendations to a district-level instructional leader before engaging other stakeholders. It is important that the message of change come from an instructional leader with clout in the district (often a superintendent, assistant superintendent or chief academic officer).

Key stakeholders who must approve or implement any changes to the program should be reminded of the student-centered vision that motivates this analysis.

Just as it was important when introducing the concept of A-ROI, emphasizing the positive impact on students is a critical piece to generating support for the changes. Ideally, some of the steps taken when selecting the program and designing the analysis will have primed the key stakeholders to interpret the results of the analysis in a meaningful way. Typically, it is helpful to communicate—at a high-level—and remind all about the motivation for the analysis, the process for selecting the program, the methods of the analysis, the key findings, and the recommendations put forth from the analysis. Reminding stakeholders of this context helps them understand and interpret the results of the analysis, even if they are unfamiliar with the process of A-ROI or the content of the program.

Possible stakeholders that will need to be involved in the approval process might include, but are not limited to:

- School board or school committee
- Director of the program
- Principals and/or teachers affected by possible changes

In addition to the formal approval processes, such as school committee votes, the district should focus on generating support through informal approval processes, such as introducing the idea at a meeting with all of the affected principals. A major piece of successful implementation is ensuring that the vision and the plan created at the central office is translated effectively down through the principals and into the classroom.

Example

Imagine there is a reading program that is shown to have modest increases in student outcomes. A finding from the analysis hypothesized that a major reason the program was not more effective is because the entry criteria for the program was based on teacher recommendation, leading to many grade-level readers who struggle with behavior to be placed into a program designed for students who are behind grade-level in reading. In presenting this information, it is helpful to clearly articulate the goal for the program, its intended effects and target population, and then compare them to its current effects and population.

Goal of the Program	To teach students to read effectively.
----------------------------	--

	Population	Effects
Intended	Students with low reading ability	All students who are behind in reading will return to grade-level
Current	Some students with low reading ability, many grade-level readers who act out in class	Modest improvement of reading ability when averaged for the group in the program

Recommendation	Do not place students who are proficient readers in the program. Analyze how the program fares when it is provided to the target population.
-----------------------	--



10b. Make explicit decisions about which recommendations to pursue

Once the report is published by the A-ROI team, the district leadership should decide which recommendations to pursue.

Without an explicit description of what will change in the future to improve student outcomes based on the findings from the report, it can be difficult to communicate a plan to implement the changes. There is a real danger in investing significant time and energy in analyzing a program, identifying a few actionable findings, and then never taking the necessary steps to realize the change promised by the recommendations.

It is important for the district leadership to decide and explicitly state the changes that can be expected in the program's staffing, design, and budget.

10c. Create a timetable for change and work plan to initiate changes

Once the district leaders decide on which recommendations to pursue, it is important to take steps toward implementing those changes by engaging principals and other district players.

It is possible to craft a plan at the district-level that is perfectly aligned with strategic priorities only to see the implementation falter. To avoid a misstep during implementation, districts can take actions that will move them toward realizing their plan soon after approval. A few strategies to support a successful implementation might include:

- Creating a work plan, including a timeline for major pieces of the work to be completed
- Assigning a high-capacity individual to oversee the implementation process across the district
- Creating a process to track the types of students entering a program at each school
- Refocusing the efforts of some staff members to or away from a particular program

Example

Using the same example as Step 10a, a district found that many students entering into an elementary reading program were referred by their teachers due to behavior issues rather than an inability to read. These students were on grade-level in reading but acted out in class, causing them to be misidentified as low readers. These students did not benefit from the intervention. On the other hand, students who were behind in reading realized significant gains over the course of the program. Thus, the district wanted to limit the reading program to students who struggle with reading and to provide a more appropriate behavior intervention to students who read well but act out in class.

To do so, the district created a rule that only students under a certain DIBELS score could enter the program, and the district charged one of its highest-energy central office workers to oversee the new process at each school. If any school referred students with high DIBELS scores to the program, the principal would work with the person overseeing the program to address the problems causing the discrepancy. Additionally, the Chief Academic Officer offered a timeline for summer work to ensure that the district was prepared to implement these changes on the first day of school in the coming year.



10d. Assess the impact of changes

Once a clear path forward has been set, the district should create the necessary systems to monitor the effect of the changes.

To embed a culture of continuous improvement, the district can develop systems to monitor the progress and effectiveness of programs. Developing a team that is dedicated to conducting A-ROI analyses to determine what works in the district is a solid foundation for these systems. They will help to ensure that the changes made based on the initial analysis of this program are driving better student outcomes. This team can revisit the same program a year later and write a brief update to the report to document the improvements that have occurred.

Mistakes to avoid

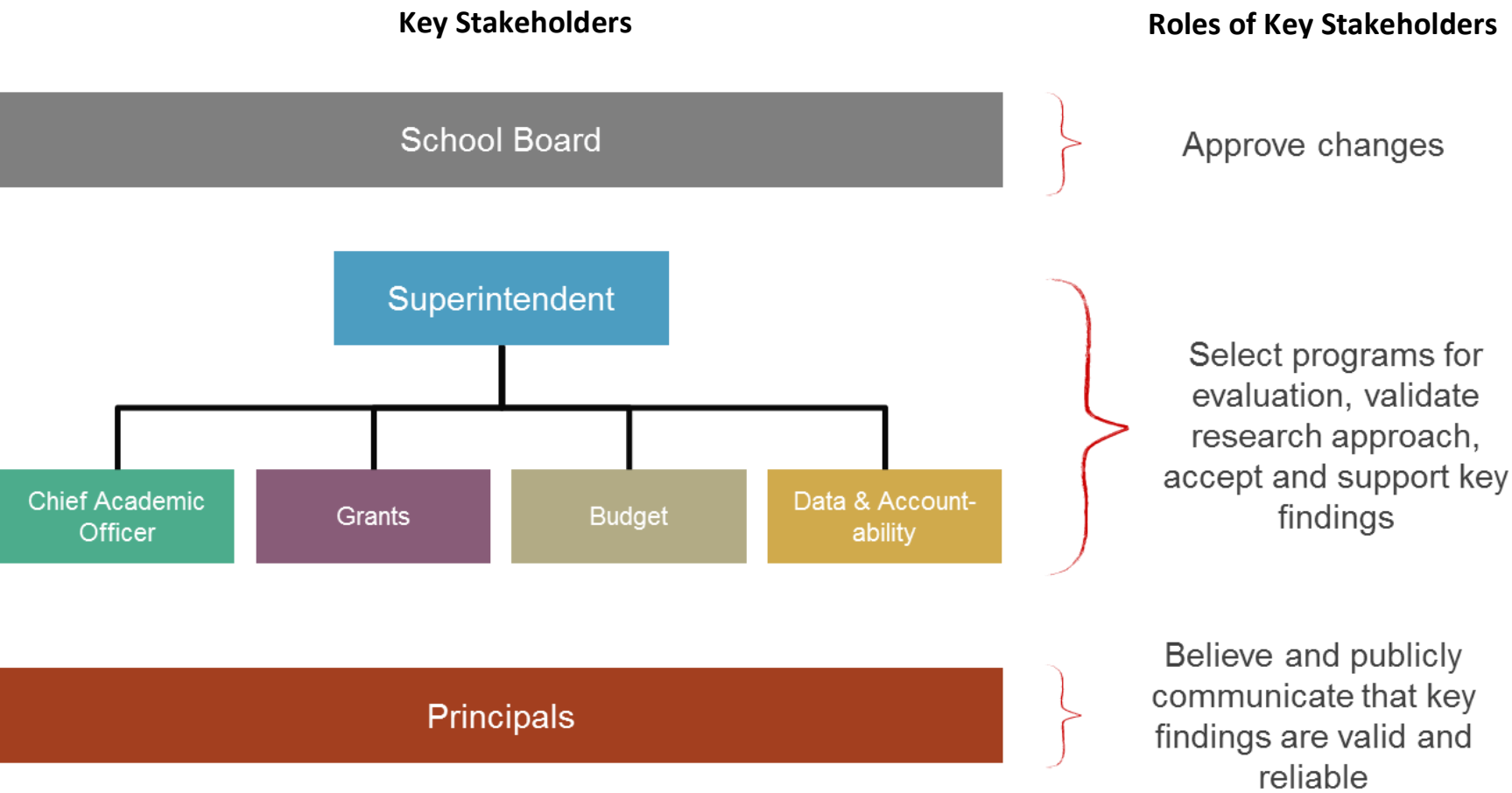
When Taking Action:

- Presenting the decision made by the district leadership as a firmly embedded plan before other key stakeholders have a chance to engage with the recommendations.
- Thinking and communicating about the future of the program in a “keep or eliminate” framework rather than utilizing the six possible options laid out in Step 9.
- Not creating an accountability mechanism to ensure that implementation is effective.
- Treating an A-ROI analysis as a one-and-done process, rather than a continuous cycle of improvement.

Appendix A: Toolkit for A-ROI Analysis

Tool 1.1: Roles of Key Stakeholders

Communicate with these key stakeholders with the understanding that a successful A-ROI analysis requires each of them to fulfill the role listed below.



Tool 1.2: Program, Strategy, and Effort Organizer

Instructions: List all major programs currently in use in your school or department as well as any programs that you plan on launching in the near future. Major programs can be a variety of items, including, but not limited to, programs aligned with district priorities, programs funded by Title I, programs serving the district's neediest students, and programs that require significant investment of staff time or dollars.

Curriculum Initiatives	Instructional Strategies	Support For Teachers	Non-Academic Programs
<i>e.g., Read 180, Elementary world language initiative</i>	<i>e.g., Co-teaching; additional reading block for struggling readers; class size reductions; alternative schools</i>	<i>e.g., Teacher mentorship program; focused PD initiative</i>	<i>e.g., Middle school social worker program; Autism inclusion program; Parent engagement initiative</i>

Tool 1.3: Program Selection Rubric

(Electronic version available)

List the high-potential programs, strategies or efforts, then score each of them.

Program or Strategy	Aligned to strategy?	Direct impact on learning?	Large reach or plans for expansion?	Significant investment of staff time?	Significant investment of cash?	Politically feasible to change?	Data available?	Large number of students?	Comparison group?	Uncertain effect?	Total
Example: <i>Technology based math program</i>	X	✓	✓	✓	✓	✓	X	X	✓	✓	7
1.											
2.											
3.											
4.											
5.											



Tool 2.1: Logic Model Template

Begin by writing the intended outcomes of the program, then describe the important actions that take place in the program to create those outcomes for students. Finally, identify the necessary inputs that will be required for those actions to take place.

Intended Outcomes

Actions

Inputs

Tool 2.2: Identifying Outcome Measures

List each intended outcome and a high quality metric to measure the specific impact of the program.

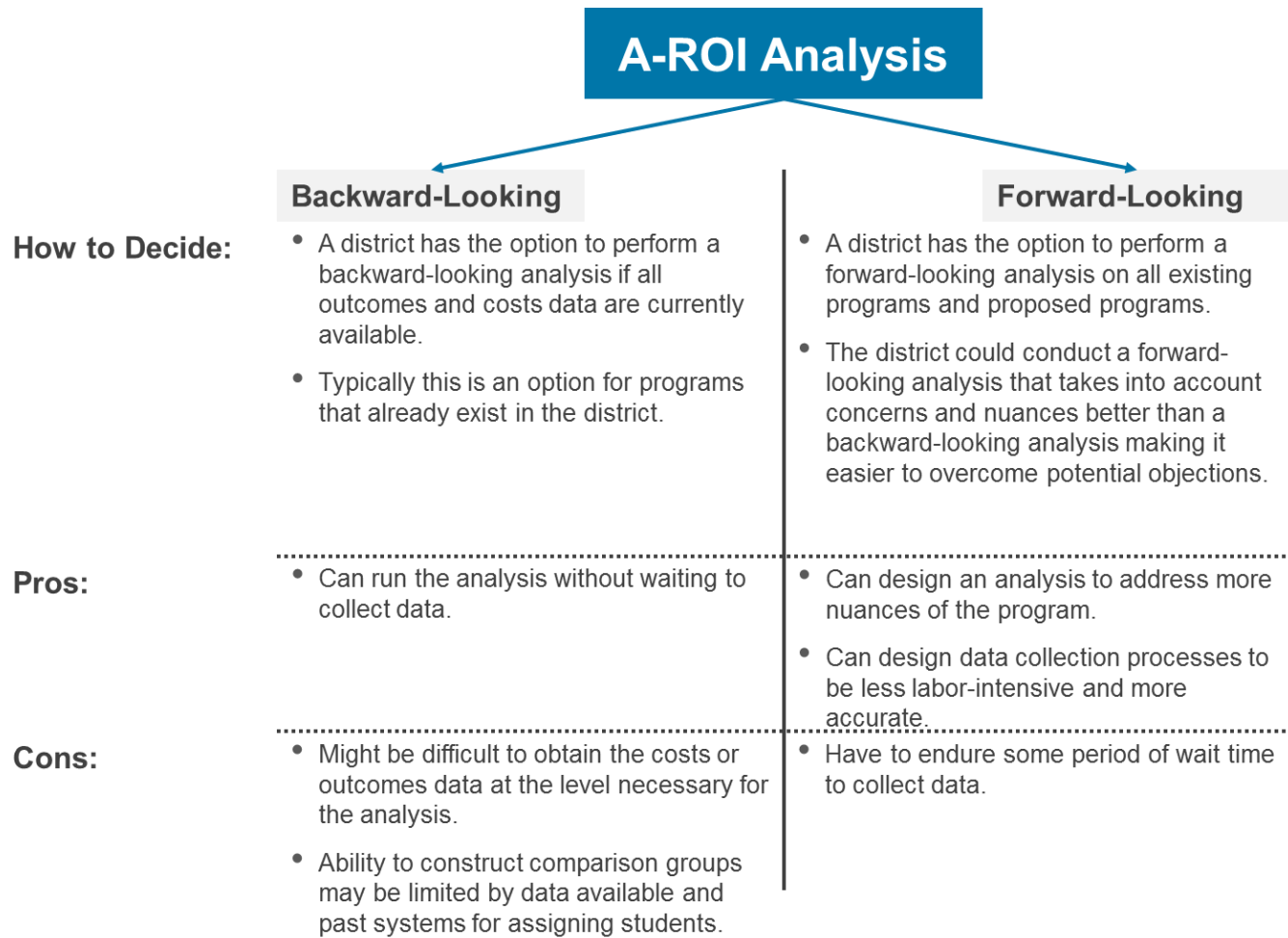
Intended Outcome	Metric
<u>Ex 1.</u> Increased reading ability	DIBELS scores
	ELA state test scores
<u>Ex 1.</u> Increased enjoyment of reading	Number of minutes read outside of school or number of pages read outside of school
<u>Ex 2.</u> Increased use of higher-order thinking (HOT) questions for MS math teachers	Count of HOT questions asked by teachers from random observations by instructional coaches
<u>Ex 2.</u> Increased understanding of mathematical concepts	MS math state test scores in particular strands

Tool 2.3: Benchmark for Success Template

Using your information from “Identifying Outcome Measures” for the first two columns, set a specific benchmark for anticipated gains for each outcome to help measure the success of the program.

Intended Outcome	Data Needed	Anticipated Gains
Ex 1. Increased reading ability	DIBELS scores	>1 year of growth in 1 year on DIBELS
	ELA state test scores	15% growth in # of students proficient/advanced
Ex 1. Increased enjoyment of reading	Records of outside of school reading logs	>50% increase in minutes read outside of school for struggling readers
Ex 2. Increased use of higher-order thinking (HOT) questions for MS math teachers	Count of HOT questions asked by teachers from random observations by instructional coaches	>3x of questions from teacher are HOT
Ex 2. Increased student understanding of mathematical concepts	MS math state test scores	>15% growth in # of students proficient/advanced in specific strands

Tool 3.1: Comparing Forward- and Backward-Looking Analyses



Tool 4.1: Possible Key Variables for Segments

Key Variables (Academic)	Potential Segments
Reading Ability	<ul style="list-style-type: none"> • On or above grade level • <1 year behind • >1 year behind
Subject-Specific Ability	<ul style="list-style-type: none"> • On or above grade level • <1 year behind • >1 year behind
Specific Special Needs	<ul style="list-style-type: none"> • Mild/Moderate Autism • Severe Autism • Mild/Moderate Speech and Language Disability • Severe Speech and Language Disability • ADD/ADHD/OCD • Non-ADD/ADHD/OCD Other Health Impairments • Orthopedic Impairment/Hearing Impairment/Visual Impairment/Deaf-Blindness • Multiple Disabilities/Traumatic Brain Injury • Emotional Disturbance • Mental Retardation/Developmental Delay
Specific ELL Status	<ul style="list-style-type: none"> • Continuous rigorous schooling, but little English • Continuous low-rigor schooling, but little English • Limited or inconsistent formal education • Continuous rigorous schooling, and moderate English proficiency <p>These can be tracked by knowing the country of origin for ELL students and measuring their DIBELS/TOEFL scores.</p>
Student Engagement	<ul style="list-style-type: none"> • Homework completion rate: <ul style="list-style-type: none"> ○ >90% ○ 70-90% ○ <70%
Extracurricular Involvement	<ul style="list-style-type: none"> • Member of two or more groups • Member of one group • Not involved in a group
Grade Retention/Promotion History	<ul style="list-style-type: none"> • Passed all grades on schedule • Repeated one or more years of school
Grade	<ul style="list-style-type: none"> • K-12
Participation in Related Programs	<p>For a dropout prevention program:</p> <ul style="list-style-type: none"> • Has a mentor with the Big Brother/Big Sister Mentorship Program <p>For a reading program:</p> <ul style="list-style-type: none"> • A student's parents provide an outside tutor on the weekend

Tool 4.1: Possible Key Variables for Segments, Cont.

Key Variables (Non-Academic)	Potential Segments
Age	<ul style="list-style-type: none">• Within expected age range for grade• Older than expected• Younger than expected
Attendance	<ul style="list-style-type: none">• >98%• 95-98%• <95%
Behavior	<ul style="list-style-type: none">• No detentions• 1-2 detentions this year• >2 detentions this year
Parent Engagement	<ul style="list-style-type: none">• Parent is a member of PTA• Parent attended one or more parent-teacher conferences• Parent did not attend a school event this year
Transience/Mobility	<ul style="list-style-type: none">• Changed schools at least once in the last 3 years• Changed schools in the last year• Remained at the same school for the last 3 years• Moved schools in the last year, but stayed within the district



Tool 4.2: Student Segment Organizer

Use the diagram below to outline the hypothesized segments for the program being analyzed. Hypothesis:

What is the profile of students who receive this program currently?

Placeholder for student profile information

What are the key variables that will impact student outcomes in this program?

Placeholder for key variables

Which segments can be made using these key variables?

Placeholder for student segments

Tool 6.1: Cost Map

Use the table below to articulate the different types of costs associated with the program. The goal at this stage is simply to list out the different types of expenditures, NOT to list the actual amounts.

Cost Categories	Start Up Expenditures	Ongoing Expenditures
Staff Time on Direct Service		
Staff Time on Training		
Administrator Time Dedicated to Program		
Fees and Stipends		
Materials		
Operations		



Tool 6.3: Cost Data Collection Templates

For cash investments:

Expenditure	Date	Year	Category	School/Site	Annual Cost (\$)

For time investments:

Activity/Meeting/Event	Date	Year	Category	School/Site	Position	Frequency Per Year	Number of Attendees	Length (Hours)	Hourly Rate	Annual Cost (\$)

Tool 7.1: Segment Identification Process

4. Run Regression Analysis

Especially if no meaningful segments have been found, it can be useful to conduct a regression analysis to identify any significant correlation between potentially key variables and outcomes. If trends are identified, return to Guess & Check to confirm the importance of the new segments.

2. Test Initial Segmentation

Conduct an analysis to determine if the initial segments created in Step 4 proved to be meaningful.



3. Guess & Check

Use the remaining segment data collected to analyze if the program had meaningful effects on other segments of students.

5. Draw Insight

Use the three types of analysis to draw insight into which student segments benefited the most from the program, if any.



Appendix B: Introductions to A-ROI Concepts

Introduction to Segmentation

The concept of student segmentation is new to many districts, so the overview below may be helpful in understanding the role that segmentation plays in the A-ROI process.

Typical program evaluation in districts seeks to answer the question “Does this program work?” The response is generally a blanket “good or bad” judgment about a program’s effectiveness.

Unfortunately, the “good or bad” approach to program evaluation limits districts’ options for what to do with the program. If a program is bad, the district should eliminate it; if a program is good, the district should keep it.

With A-ROI, the analysis should move past “good or bad” and into a mindset of “good or bad *for which students*.” Many districts already look at test scores broken down by student sub-groups as defined by No Child Left Behind (NCLB), such as race, disability and poverty status. However, these groupings are overly broad and within each group there will be students who respond very differently to specific programs and interventions. To assess which students benefit most from a program, it is important to use segments that group students according to characteristics relevant to the program.

Student segmentation by key variables allows for more meaningful comparisons between programs. To compare the effectiveness of Program A to the effectiveness of Program B, segmentation allows an analysis to isolate how a particular segment of students performed in each program.

For example, a district might assess whether their elementary reading intervention is working well for students with special needs. To take this one step further, the district could segment the students with special needs according to the actual learning challenges that they face: dyslexia, ADHD, specific learning disability, etc. In addition, the students can be segmented further according to their reading level at the start of the year. Then within a given segment, the analysis can compare the results of students who received the intervention with those who did not. With these factors taken into consideration, the district can now draw much more sophisticated conclusions about who actually benefits from the reading program.

Introduction to A-ROI Cost Accounting

The cost accounting presented in this section is unlike typical cost accounting in districts in five major ways:

It is not bound by compliance requirements.

Most school budgets are structured to comply with state reporting, so that revenues and expenditures must match to the penny. The cost accounting presented here is different. It is based on a need to understand the resources allocated to a specific program; this requires more nuance than finding the sum of line items in the budget, but less precision than matching revenues and expenditures to the penny.

Personnel costs are not fixed costs; staff time invested counts toward the cost of the program.

Districts typically view personnel costs as fixed costs, which limits leaders' understanding of cost and the management options available to them. DMGroup's approach to cost accounting treats staff time as a resource allocated to the program, which can be adjusted and reallocated however the district sees fit. Staff time includes the time general education teachers spend teaching students, the time a paraprofessional spends supporting a student, and time spent in professional development.

Expenditures from all sources, including grants, are counted.

Many districts keep separate budgets for the general fund, Title I, IDEA, and so on, sometimes with distinct accounting systems for each due to compliance requirements. The cost accounting approach presented here treats all dollars the same and does not exclude items paid for by different funding sources.

It takes more nuance than including or excluding each line item in the budget.

Many of the costs, especially investments of time, will not be distinct line items in the traditional school budget, but must be captured in the A-ROI cost accounting approach. For instance, instructional coaches might dedicate a portion of each week to supporting teachers of a reading intervention. In this case, the full line item of "Instructional Coaches" should not be included in the cost calculation for the reading intervention. However, to get a true total cost of the program, a portion of the line item should be included based on an estimate of the time investment of the one instructional coach who worked on the program.

It is much more important to estimate the large costs accurately than to include all of the small costs.

Typical school accounting efforts, because they are forced to meet exact standards for compliance, focus on finding all of the costs that are included in the budget, no matter how small. The cost



accounting approach outlined in this section focuses on estimating large costs accurately rather than including all of the minute costs. For instance, ensuring that the cost of a teacher who spends half of her time teaching reading is included in the program cost is much more important than working hard to determine whether 64 or 65 textbooks were purchased for the program.

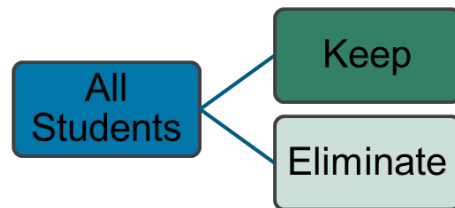
There are three main steps in gathering the cost data before a more nuanced cost analysis can begin. These steps are described in more detail throughout Step 6:

1. Map out all the types of costs for the program
2. Collect the needed cost data for cash and time investments
3. Calculate the total cost of the program

Introduction to Drawing Insight with A-ROI

Many districts currently conduct program evaluation to answer a simple question: should we keep or eliminate this program for our students? A-ROI analysis, by providing a more nuanced understanding of program effectiveness, gives districts the opportunity to be more strategic and creative in determining the future of a program.

Typical Options for the Future of a Program
(Without A-ROI Analysis)



DMGroup is introducing two major shifts away from the traditional thinking on how to draw insight based on program evaluation using A-ROI analysis:

1. There are more choices than just keeping or eliminating a program based on the results of an A-ROI analysis.
2. There can be different decisions about the future of a program for different student segments.

Options regarding the future of a program are outlined below. These options apply not only to programs as a whole, but should be considered with regard to segments within a program.

- **Expand**: if a district finds that the program is particularly cost-effective, it may want to expand the program to all students in the segment and even similar segments if they are not already receiving it.
- **Keep**: if a district finds that the program is cost-effective, but all of the students in the segment already receive the program, the district may choose to keep the program as is. Or, if a program is effective but costly, the district may choose to keep it as is, until they find an alternative that is more cost-effective for those students.
- **Fix**: if a district finds that only some pieces of a program are effective, it may seek to fix the ineffective pieces. Particularly if the district finds that implementation was weak or the levels of service were variable, the district may seek to modify the current program delivery without expanding or reducing it.
- **Streamline**: if a district finds that pieces of a program are ineffective, it may seek to cut out those pieces but keep the effective portions. Alternately, if a program is found to be effective but costly, the district may want to reduce the program and explore alternatives that might be more cost-effective or effective.
- **Eliminate**: if a district finds that a program does not work for a certain segment, it may want to stop providing the program to a particular student segment altogether.



Proposed Options for the Future of a Program

