Tools

May 2015

Instructional improvement cycle: A teacher's toolkit for collecting and analyzing data on instructional strategies

Trudy L. Cherasaro
Marianne L. Reale
Mark Haystead
Robert J. Marzano
Marzano Research

Summary

This toolkit includes three tools to help teachers use data from their classroom assessments to evaluate promising teaching practices:

1. A planning guide that introduces teachers to an instructional improvement cycle in which they compare the learning results from one group of students who receive a new classroom instructional strategy (experimental group) with those of another group of students who receive a traditional strategy (comparison group) using a scientific approach.
2. A preprogrammed Excel spreadsheet that allows teachers to compare the performance of students who receive the strategy (experimental group) with that of a similar group of students who do not (comparison group).
3. A reflection guide that provides information on how to interpret and reflect on the results.

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences
U.S. Department of Education

REL CENTRAL
Regional Educational Laboratory
At Marzano Research

U.S. Department of Education
Arne Duncan, *Secretary*

Institute of Education Sciences
Sue Betka, *Acting Director*

National Center for Education Evaluation and Regional Assistance
Ruth Curran Neild, *Commissioner*
Joy Lesnick, *Associate Commissioner*
Amy Johnson, *Action Editor*
Sandra Garcia, *Project Officer*

REL 2015–080

The National Center for Education Evaluation and Regional Assistance (NCEE) conducts unbiased large-scale evaluations of education programs and practices supported by federal funds; provides research-based technical assistance to educators and policymakers; and supports the synthesis and the widespread dissemination of the results of research and evaluation throughout the United States.

May 2015

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-12-C-0007 by Regional Educational Laboratory Central administered by Marzano Research. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This REL report is in the public domain. While permission to reprint this publication is not necessary, it should be cited as:

Cherasaro, T. L., Reale, M. L., Haystead, M., & Marzano, R. J. (2015). *Instructional improvement cycle: A teacher's toolkit for collecting and analyzing data on instructional strategies* (REL 2015–080). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Central. Retrieved from http://ies.ed.gov/ncee/edlabs.

This report is available on the Regional Educational Laboratory website at http://ies.ed.gov/ncee/edlabs.

Summary

This toolkit provides teachers with three tools to implement an instructional improvement cycle, which allows teachers to test a new instructional strategy using a scientific approach: a planning guide, an Excel analysis tool, and a reflection guide. The toolkit is intended for use by individual teachers to test and reflect on a strategy on their own or in teacher teams. The planning guide provides step-by-step instructions and worksheets to plan for testing an instructional strategy using the instructional improvement cycle. The Excel analysis tool is a preprogrammed Excel spreadsheet that compares the learning results from students who receive the strategy to the learning results of students who do not. The reflection guide provides information on how to interpret results and a worksheet to help teachers reflect on the results.

This toolkit was developed by Regional Educational Laboratory Central in collaboration with York Public Schools in Nebraska. Educators in York Public Schools were interested in creating tools and protocols to help teachers independently evaluate promising teacher practices.

Contents

Summary i

Introduction 1
What is this toolkit? 1
What is the instructional improvement cycle and how long does it take? 2
How can teachers use the instructional improvement cycle? 2
What does the toolkit include? 3
How can teachers use this toolkit? 5

Toolkit part I. Planning guide 7
Step 1. Select an instructional strategy 7
Step 2. Select classes 10
Step 3. Select a content assessment 12
Step 4. Determine when and how to use the instructional strategy 14

Toolkit part II. Excel analysis tool 14
Where to download the Excel analysis tool 14
How to use the Excel analysis tool 14

Toolkit part III. Reflection guide 15
Using results from the Excel analysis tool 15
What do your results mean? 16

Appendix A. Technical documentation A-1

References Ref-1

Boxes
1 Toolkit development 1
2 Example of the cycle 3
3 Benefits of randomization 3
I-1 Examples for selection routes 8

Figures
1 Instructional improvement cycle 2
2 Tools for implementation 4
3 Using the planning guide in part I to plan steps of the instructional improvement cycle 5
4 Implementation steps of the instructional improvement cycle 6
II-1 Screenshot from the Excel analysis tool 14
III-1 Interpreting effect sizes 16
A1 Screenshot of the Excel analysis tool A-1

Tables
I-1 Characteristics of group comparison tests 10
I-2 Assessment selection template 12

Worksheets

I-1 Strategy selection 8
I-2 Class description 11
I-3 Strategy and assessment implementation 13
III-1 Reflection 17

Introduction

With the advent of information management systems, most teachers have access to an abundance of student data. One U.S. survey revealed that almost all districts provide student information systems, and the majority (70 percent) of districts have systems that organize benchmark assessment data for analysis (Means, Padilla, & Gallagher, 2010). While many teachers use student data to track and determine student progress, few are able to use the data to establish which aspects of their teaching are working well or to evaluate new teaching practices (Means et al., 2010). Promising methods to help teachers use data to improve instruction include providing them with protocols or tools for analyzing data on student learning in order to inform their instructional practice (Gallimore, Ermeling, Saunders, & Goldenberg, 2009; Jaquith & McLaughlin, 2010; Means et al., 2010) and using data that are closely aligned with their instruction (Datnow, Park, & Kennedy-Lewis, 2013; Massell, 2001; Means, Chen, DeBarger, & Padilla, 2011).

This toolkit, developed by Regional Educational Laboratory (REL) Central in collaboration with York Public Schools in Nebraska, provides a process and tools to help teachers use data from their classroom assessments to evaluate promising practices. A description of how this toolkit was developed and tested with multiple groups of teachers is provided in box 1.

This toolkit provides a set of tools that can help teachers engage in an instructional improvement cycle of collecting and analyzing data on instructional strategies and student learning and adjust their instruction based on these data

What is this toolkit?

This toolkit provides a set of tools that can help teachers engage in an instructional improvement cycle of collecting and analyzing data on instructional strategies and student learning and adjust their instruction based on these data. The toolkit can be used either by individual teachers or by teacher teams. Individual teachers can implement the instructional improvement cycle to test a strategy and reflect on their practice, and teacher teams can test individual strategies and reflect on the results as a team. The toolkit consists of three tools—a planning guide, an Excel analysis tool, and a reflection guide. The tools provide teachers with guidance on how to deliberately apply and study one classroom strategy over the course of one instructional unit and how to systematically document and compare results in order to consider the strategy's effects on student learning. By following

Box 1. Toolkit development

In York Public Schools, as part of teachers' professional development experiences, teachers are asked to independently evaluate instructional strategies by identifying two classes with similar instructional units and then teaching one class using an innovative instructional strategy (experimental group) and the other class using customary instruction (comparison group). Teachers collect pre- and post-test data using teacher-developed assessments to evaluate whether the instructional strategy appears to work well.

Regional Educational Laboratory (REL) Central collaborated with York Public Schools to create protocols and tools for these teachers that could be self-executed without the need for external data analysis. REL Central worked closely with two cohorts of teachers in York Public Schools to iteratively develop and test a set of protocols and tools, which resulted in this toolkit. REL Central collected feedback from teachers and revised the toolkit to make the tools clearer and more precise.

the toolkit, teachers are guided to select an instructional strategy to test and then to implement the strategy in one of their classrooms during one instructional unit while implementing typical instruction with the same content in another classroom. Teachers test students in both classes before and after implementing the instructional strategy and enter the scores into an Excel spreadsheet. The spreadsheet is preprogrammed to automatically show results that provide information on how well the strategy worked. Teachers can use the information gathered through the instructional improvement cycle to reflect on and adjust their instruction to increase student learning.

What is the instructional improvement cycle and how long does it take?

The instructional improvement cycle includes four recurring components: selecting an instructional strategy, implementing the strategy, collecting data on strategy implementation, and analyzing the data and reflecting on the results (figure 1).

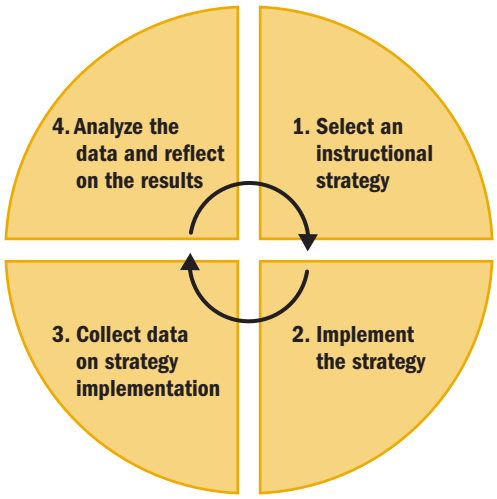
Teachers can implement the cycle as many times as they like to test multiple strategies or to gather additional information about one strategy from different classes of students

Teachers can implement the instructional improvement cycle over the course of one unit. Using the tools, teachers select an instructional strategy, implement the strategy in existing lessons, collect student learning data, and analyze and reflect on the results. Teachers can implement the cycle with very little change in their regular practice. As teachers are planning upcoming units (within one to two weeks of teaching the unit), they can incorporate the instructional improvement cycle by embedding the strategy into their upcoming unit and using planned content assessments as measures of student learning. An example of a high school English language arts teacher using the toolkit to implement the cycle is given in box 2. Teachers can implement the cycle as many times as they like to test multiple strategies or to gather additional information about one strategy from different classes of students.

How can teachers use the instructional improvement cycle?

This section walks through a scientific approach to implementing the cycle and testing a specific instructional strategy by comparing the performance of students who receive

Figure 1. Instructional improvement cycle



Source: Authors' compilation.

Box 2. Example of the cycle

A grade 9 English language arts teacher is interested in learning whether the use of nonlinguistic representations (for example, graphic organizers or concept maps) during vocabulary instruction will help increase student achievement. The teacher implements a nonlinguistic representation strategy in fourth-period English class (experimental group) and uses regular instruction during sixth-period English class (comparison group). Both classes have upcoming lessons with focused vocabulary time, with the learning goal that students can identify prefixes, roots, and suffixes to determine meanings of words. The teacher administered a previously developed vocabulary quiz to both groups of students as the pre-test before implementing the nonlinguistic representation strategy. The teacher administers the same quiz at the end of the two-week unit as the post-test.

the strategy with the performance of a similar group of students who do not receive the strategy and adjusting for differences in student performance that may have existed before implementing the strategy. An important aspect of the process is that both groups of students are taught by the same teacher; thus the instructional strategy itself, rather than differences among teachers, can be tested. Researchers use a more rigorous approach to test instructional strategies by assigning groups of students at random to either receive or not receive the strategy. However, because this approach is not feasible for teachers, the next best approach is to select similar groups of students for comparison. More information on the benefits of randomly selecting students—and caution about interpreting results when students are not randomly selected—is provided in box 3. Because scientifically testing the strategy requires two groups of students taught by the same teacher, the approach used in this toolkit is most applicable for teachers with multiple classes of students.

An important aspect of the process is that both groups of students are taught by the same teacher; thus the instructional strategy itself, rather than differences among teachers, can be tested

What does the toolkit include?

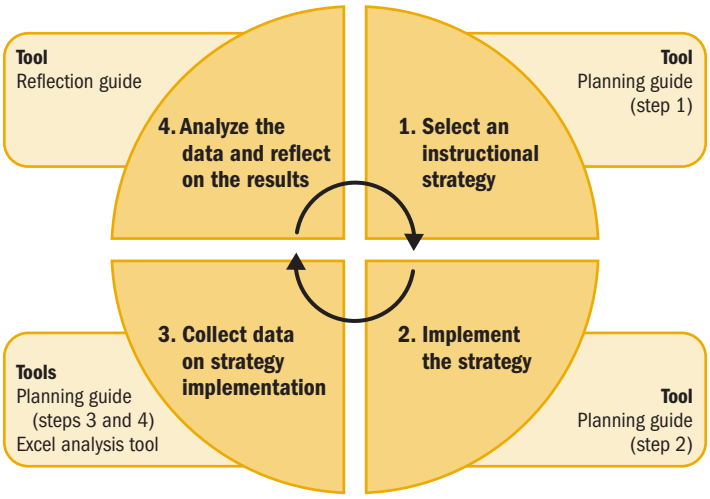
The toolkit includes three tools that teachers use to implement the instructional improvement cycle:

1. *Planning guide.* This tool provides step-by-step instructions and worksheets for teachers to plan to test a strategy using the instructional improvement cycle. The tool provides

Box 3. Benefits of randomization

The most rigorous way to determine whether an instructional strategy is effective is to conduct a randomized controlled trial in which students are assigned at random to either receive or not receive the strategy. Randomly assigning students to groups helps ensure that the groups are as similar as possible before the strategy is implemented in terms of both observable characteristics (such as test scores) and nonobservable characteristics (such as student motivation). When it is not possible to randomly assign students to groups, comparing the performance of existing similar groups of students is often used to examine whether a strategy is effective; however, this design has limitations because of the possibility that the groups differed in some way before the start of the study. For example, students who signed up for a first-period math class may be more motivated than students who signed up for a third-period math class. The analysis can account for some pre-existing differences between the groups by controlling for prior achievement, but it can never account for all the differences that could exist between the groups.

Figure 2. Tools for implementation



Source: Authors' compilation.

four steps to follow: select an instructional strategy, select classes, select a content assessment, and determine when and how to use the strategy.

2. *Excel analysis tool.* This tool analyzes student learning data to provide findings on how well the strategy worked. The analysis produces three results:
   - The baseline equivalence result tells whether the two groups of students had pre-existing achievement differences that could not be controlled for in the analysis, making the results uninterpretable because effects related to pre-existing differences cannot be ruled out.
   - The effect size tells how the achievement of students who received the strategy compares with the achievement of students who did not receive the strategy. Specifically, an effect size indicates, in a standardized way, how much better or worse on average students in the group that received the strategy did compared with students in the group that did not receive the strategy.
   - The confidence level tells whether a similar effect is likely to happen in similar classes.

3. *Reflection guide.* This tool provides instructions for teachers on how to use the results from the Excel analysis tool and a worksheet for teachers to interpret and reflect on the results. The worksheet guides teachers to think about how they implemented the strategy, their classroom context, the assessments used, and other factors that may have influenced the results, so they can plan for future instructional use of their strategy.

The corresponding tools for implementing each component of the instructional improvement cycle are identified in figure 2.

The tools for this toolkit are included here, can be downloaded as Word and Excel files from the Tools section of the REL Central website (http://www.relcentral.org/tools), or can be requested from REL Central by emailing relcentral@marzanoresearch.com.

How can teachers use this toolkit?

The three tools in this toolkit provide teachers with guidance to plan and implement the instructional improvement cycle to test an instructional strategy.

Planning to implement the instructional improvement cycle. Teachers follow the four steps in the planning guide to plan for testing an instructional strategy using the instructional improvement cycle (figure 3):

- Select an instructional strategy.
- Select classes.
- Select a content assessment.
- Determine when and how to use the strategy.

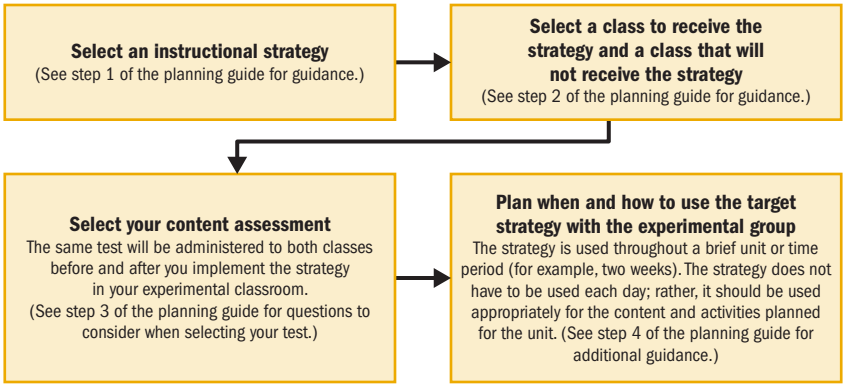
Implementing the instructional improvement cycle. To implement the instructional improvement cycle, teachers first determine whether their selected classes have baseline equivalence on the pre-test by entering their pre-test data into the Excel analysis tool. If the Excel analysis tool indicates that the classes do not have baseline equivalence, different classes should be chosen. If the classes have baseline equivalence, teachers should follow the remaining steps of implementation:

If the Excel analysis tool indicates that the classes do not have baseline equivalence, different classes should be chosen

- Teach content.
- Administer content assessment and record data (post-test).
- Interpret and reflect on results using the reflection guide.

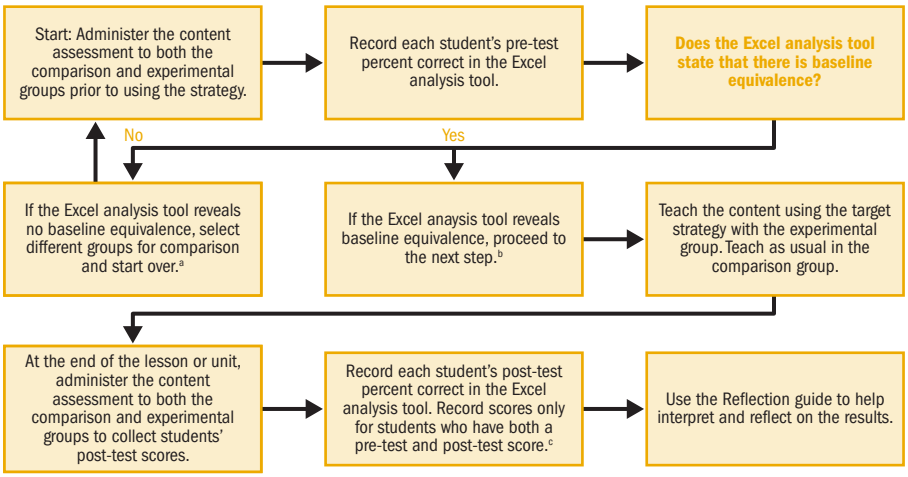
All implementation steps are presented in figure 4.

Figure 3. Using the planning guide in part I to plan steps of the instructional improvement cycle



Source: Authors' compilation.

Figure 4. Implementation steps of the instructional improvement cycle



a. "No baseline equivalence" means that the two groups have existing achievement differences that are too large to account for in the analysis; thus you will not be able to tell if the results are related to the existing differences or to the strategy implemented.

b. "Baseline equivalence" means that the groups do not have existing significant differences in achievement that cannot be controlled for in the analysis; thus the results will not be affected by existing achievement differences.

c. If you were unable to administer the post-test to some students, remove their pre-test scores and double-check that you still have a "yes" in the baseline equivalence box.

Note: Try to make sure that every student who took the pre-test also takes the post-test, if necessary by giving makeup exams to students who were absent. If some students do not take the post-test, this may change the baseline equivalence and increase the risk of having groups with pre-existing achievement differences, making your results uninterpretable.

Source: Authors' compilation.

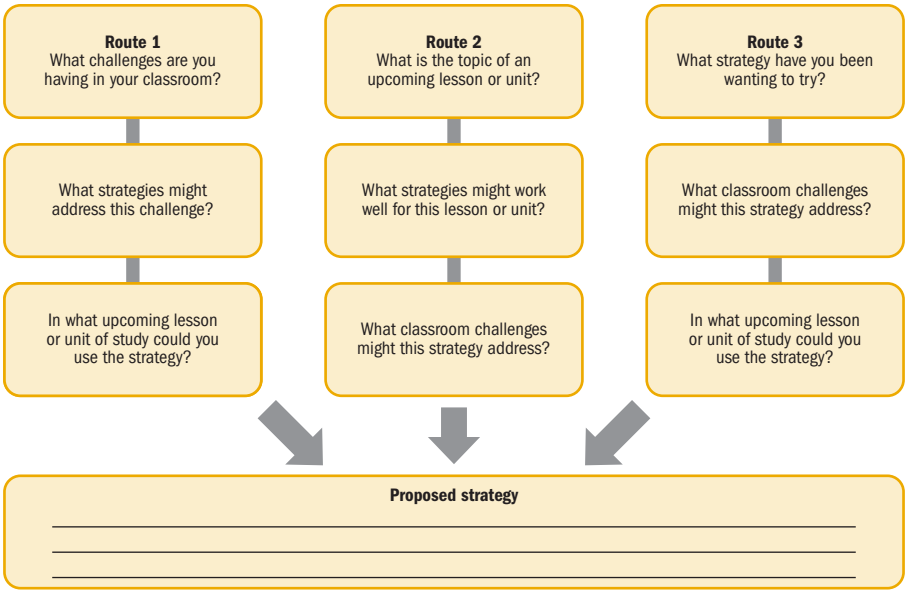
Toolkit part I. Planning guide

This guide is the first of three parts of the toolkit. To test how well a strategy works, complete steps 1–4 and the corresponding worksheets. For background information on the instructional improvement cycle, please consult the introduction to this publication.

Step 1. Select an instructional strategy

Select a strategy that you want to test by completing worksheet I-1, which presents questions to consider as you determine the instructional strategy for your project. You can start by thinking about a strategy to address current challenges in your classroom (route 1), identifying a strategy that aligns with an upcoming lesson (route 2), or selecting a strategy you want to try and aligning it to challenges and upcoming lessons (route 3). Once you choose a starting point, answer the questions in the boxes for your chosen route. An example of each route is shown in box I-1.

Worksheet I-1. Strategy selection



Source: Authors' adaptation of the roadmap provided by Jill Johnson, Education Service Unit 6, Milford, Nebraska.

Box I-1. Examples for selection routes

Route 1

Challenge. A high school math teacher saw that students were not doing well on homework that required them to solve problems.

Strategies to address the challenge. The teacher decided to incorporate more problem-solving activities into daily instruction, rather than just giving these problems as homework.

Upcoming lesson or unit. The teacher had an upcoming algebra unit on solving simple radical equations using a variety of methods.

Proposed strategy. The teacher included time for students to problem solve during class and provided them with a list of prompts to ask as they solved the problem.

Route 2

Upcoming lesson or unit. A teacher had an upcoming science unit on structure and function and cause-and-effect relationships.

Strategies for lesson or unit. The teacher decided that this would be a good unit in which to implement some inquiry-based strategies by having students conduct small experiments to examine these relationships.

Challenge. The teacher realized that students had a hard time developing hypotheses and using a scientific approach to solve problems.

(continued)

Box I-1. Examples for selection routes *(continued)*

Proposed strategy. The teacher had students conduct experiments in small groups in which they identified a hypothesis about cause and effect, tested the hypothesis by recording and reflecting on data, and generated explanations and new hypotheses.

Route 3

Strategy to try. In conversation with a subject-area team member, a teacher learned about a paired reading strategy that seemed to be successful. The teacher decided to try this strategy.

Challenge. Seeing that students were not very engaged, the teacher realized that this might be related to too much class-time lecturing. The teacher decided that implementing a grouping strategy might allow students to interact with the material and increase their level of engagement.

Upcoming lesson or unit. The teacher had an upcoming lesson on identifying and summarizing the main idea.

Proposed strategy. The teacher used a paired reading strategy in which students take turns reading and then provide praise for correct reading, provide error correction to partners, and summarize and discuss the main idea.

Step 2. Select classes

Select two classes to participate in the instructional improvement cycle. A new strategy will be tested by comparing achievement in one class that receives regular instruction methods (comparison group) with achievement in another class that receives the new strategy (experimental group; table I-1). Select two classes that target the same content (in which you will teach the same lesson or unit) and have a similar makeup (for example, gender, eligibility for the school lunch program, special education status, or English learner status). Using groups with similar traits helps address the limitations of this design by limiting any pre-existing differences. Complete worksheet I-2 to describe the two classes and the lesson or unit that you chose.

Table I-1. Characteristics of group comparison tests

Characteristic	Description
Number of classes	Two different classes.
Implementation	Class 1 gets instruction using the new strategy. This group serves as the experimental group. Class 2 gets regular instruction with the same content (same lesson or unit). This group serves as the comparison group.
Content pre-test and post-test administration	The groups take the same content assessment prior to the lesson or unit (pre-test) and the same assessment at the end of the lesson or unit (post-test).
Question addressed	How well does a group of students that receives the target strategy achieve compared with a group of students that does not receive the strategy?
Limitations	Without random assignment, the results on the content assessment may be a result of pre-existing differences between the two student groups rather than a result of using the strategy.

Source: Authors' compilation.

Worksheet I-2. Class description

Class 1: experimental group (the group for which the strategy will be used)
Number of students
Course name or subject[a]
Topic[a]
Grade[a]
Demographics (for example, percentage of students eligible for the school lunch program, special education students, or English learner students)
Lesson or unit learning goals (that is, what do you want students to know and be able to do at the end of the unit)[a]
Target strategy tested (for example, advanced organizers or exit slips)
Other strategies used, if any (for example, standard strategies you are using, such as discussion)
Class 2: comparison group (the group for which the strategy will not be used)
Number of students
Course name or subject[a]
Topic[a]
Grade[a]
Demographics
Lesson or unit learning goals[a]
Strategies used (that is, standard strategies you are using, such as discussion)

a. Must be the same for both groups.

Source: Authors' compilation.

Table I-2. Assessment selection template

Learning objective	Number of questions in the content assessment
1.	
2.	
3.	

Step 3. Select a content assessment

Select a content assessment that you will administer as a pre-test and post-test to both groups. The same content assessment is administered twice in each class, once prior to strategy implementation and again following strategy implementation. The content assessment can be one that you developed or one found in a textbook or other resource. The content assessment should cover content from the entire unit. Consider the questions below when selecting or designing the content assessment.

- Are the questions on the content assessment clearly aligned to the learning objectives identified in the lesson or unit? Does the content assessment have enough questions to fully capture student learning and to show change over time? The assessment should have enough questions to cover each of the learning objectives. A simple table can be used to identify the number of questions related to each objective (see table I-2 for a template).
- Are the content assessment questions at an appropriate difficulty level to allow for increases in learning to be captured from before instruction on the lesson or unit begins (pre-test) to after instruction on the lesson or unit is complete (post-test)? Do not make the assessment so difficult that most students will not be able to answer questions correctly during the post-test because this will make the gains look smaller than they are. Nor should you make the assessment so easy that students are getting most questions right on the pre-test, because this also will make the gains look smaller than they are. If you have used the assessment in the past, you might consider previous results as an indicator of difficulty.

Step 4. Determine when and how to use the instructional strategy

In the previous steps you decided on the instructional strategy that you would test, the classes that would be used to test the strategy, and the assessment that would be used as a pre-test and post-test. These decisions should have led you to choose a general timeframe for when you would implement the instructional strategy, including when you would administer the pre-test and post-test based on the length and timing of the lesson or unit that you chose in step 2. You will now need to consider how the instructional strategy will be implemented so that it is appropriately aligned to the content and activities in the unit. For example, if using a paired reading strategy, you would need to determine when students would have material that is appropriate to use in the pairs and the length of time that students would spend in the pairs. Complete worksheet I-3 to help you plan for implementation.

Worksheet I-3. Strategy and assessment implementation

Questions to consider
What is the name of the lesson or unit?
What dates will you teach the lesson or unit?
What date will you administer the pre-test?
What date will you administer the post-test?
How often will you use the strategy (for example, daily, after each reading passage)?
How much time will you allot for implementing the strategy each time it is used (for example, 10 minutes in groups, 15 minutes of problem solving)?

Source: Authors' compilation.

Toolkit part II. Excel analysis tool

The Excel analysis tool (part II of this toolkit) will help you analyze the results of the instructional strategy you developed in the planning guide (part I of this toolkit). Follow the steps below to conduct your analysis.

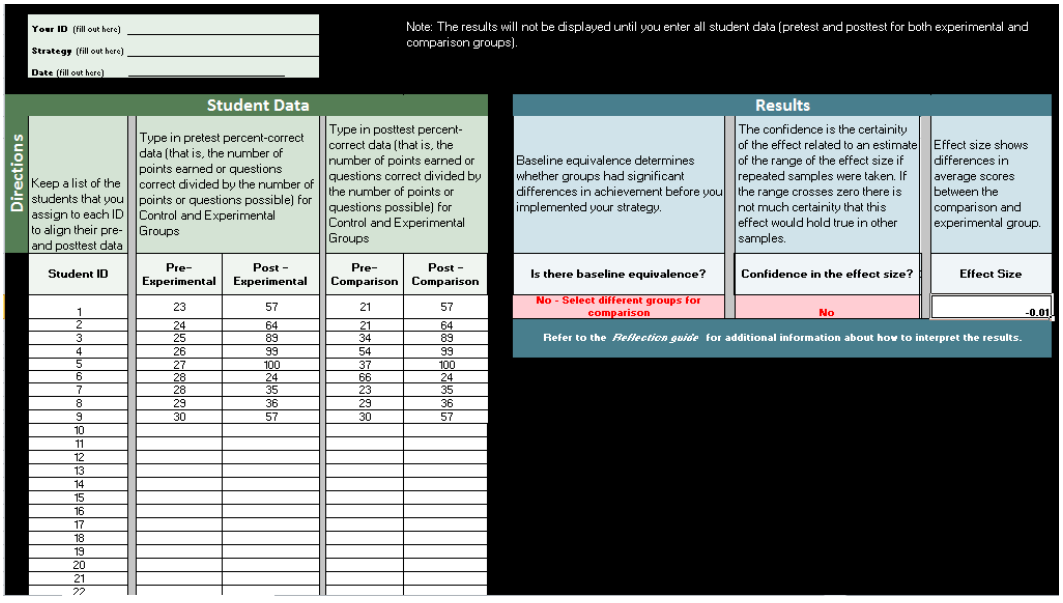
Where to download the Excel analysis tool

The Excel analysis tool is an Excel file available for download on the Regional Educational Laboratory Central website (http://www.relcentral.org/tools). Figure II-1 shows a screenshot of the data entry and results sections of the Excel analysis tool. You will use the Excel analysis tool at two points in your study: to enter the pre-test data and check for baseline equivalence and to enter post-test data in order to calculate the results.

How to use the Excel analysis tool

Enter the pre-test and post-test scores for the experimental and comparison classes into the appropriate student data boxes (highlighted in green). Once the test scores are entered, the baseline equivalence, confidence, and effect size will show in the results box (highlighted in blue). Guidance on how to interpret the results is provided in worksheet III-1 in the next section. Appendix A provides the formulas that are used in the Excel analysis tool.

Figure II-1. Screenshot from the Excel analysis tool



Your ID (fill out here) _____
Strategy (fill out here) _____
Date (fill out here) _____

Note: The results will not be displayed until you enter all student data (pretest and posttest for both experimental and comparison groups).

Student Data

Directions		Type in pretest percent-correct data (that is, the number of points earned or questions correct divided by the number of points or questions possible) for Control and Experimental Groups		Type in posttest percent-correct data (that is, the number of points earned or questions correct divided by the number of points or questions possible) for Control and Experimental Groups	
Keep a list of the students that you assign to each ID to align their pre- and posttest data	Student ID	Pre–Experimental	Post – Experimental	Pre–Comparison	Post – Comparison
	1	23	57	21	57
	2	24	64	21	64
	3	25	89	34	89
	4	26	99	54	99
	5	27	100	37	100
	6	28	24	66	24
	7	28	35	23	35
	8	29	36	29	36
	9	30	57	30	57
	10				
	11				
	12				
	13				
	14				
	15				
	16				
	17				
	18				
	19				
	20				
	21				
	22				

Results

Baseline equivalence determines whether groups had significant differences in achievement before you implemented your strategy.	The confidence is the certainity of the effect related to an estimate of the range of the effect size if repeated samples were taken. If the range crosses zero there is not much certainity that this effect would hold true in other samples.	Effect size shows differences in average scores between the comparison and experimental group.
Is there baseline equivalence?	**Confidence in the effect size?**	**Effect Size**
No - Select different groups for comparison	No	-0.01

Refer to the *Reflection guide* for additional information about how to interpret the results.

Toolkit part III. Reflection guide

In the planning guide (part I of this toolkit), you developed a plan to test your instructional strategy. In the Excel analysis tool (part II of this toolkit), you analyzed the effects of the strategy. In this third and final part of the toolkit, you will interpret your results and determine what your next steps should be.

Using results from the Excel analysis tool

After entering your data into the Excel analysis tool, look at the baseline equivalence, effect size, and confidence boxes in the results section. These data provide information about how the implementation of an instructional strategy may relate to student achievement in a given situation.

As you enter the post-test data, delete the pre-test scores of any students who did not take the post-test. After doing that, double-check that you still have a "yes" in the baseline equivalence box (first box). If you do not, the students in the two classes had differences in achievement prior to implementing the strategy that cannot be accounted for in the analysis. If this is the case, your results are not interpretable because the effects may be related to pre-existing differences.

The effect size (last box on the results screen) compares the scores on average from the comparison and experimental groups to show how much larger (or smaller) the average score is in the class where your new strategy was used compared with a class where it was not used. A positive effect size means that the students who received the strategy did better on average, while a negative effect size means that the students who did not receive the strategy did better on average. The larger the effect size, the larger the difference between the groups. For example, an effect size of 0.50 indicates a larger effect than an effect size of 0.01.

The confidence in the effect size (middle box) shows the level of certainty in the effect size by estimating what the range of the effect size would be if the study were repeated many times with different groups of students (confidence intervals). The reported effect size describes only the effect on your group of students, whereas the confidence shows what you could likely expect if you used the strategy again on another group. If there is no confidence, it is unclear whether this strategy would work again. The Excel analysis tool requires a confidence level of 90 percent (Steiger, 2004), which means that it is 90 percent sure that the effect size would fall within a certain range if the study were repeated many times. If the range crosses zero (goes from negative to positive; for example, –0.1 to 0.5), it is uncertain whether the intervention would have a positive or a negative effect if implemented with a different group of students. Thus the results receive a "no" for confidence in the effect size. If the range does not cross 0, the results receive a "yes" for confidence in the effect size.

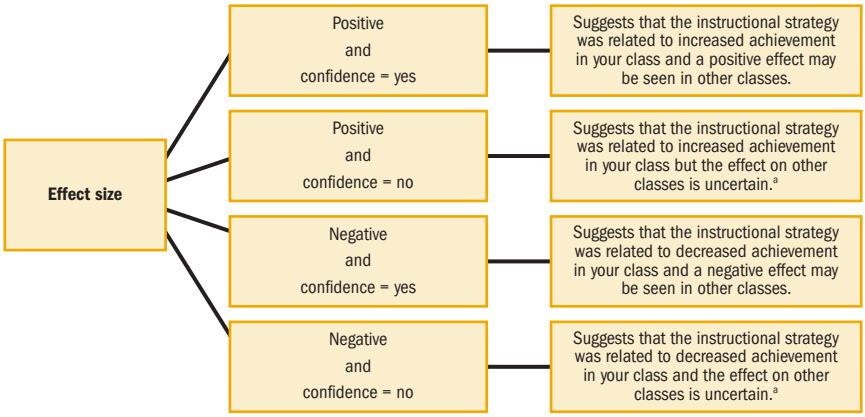
While the results from one study cannot be used to conclusively state whether a strategy is effective, they do give some information about how that strategy worked in your setting. Formulas for the analyses are provided in appendix A.

What do your results mean?

Once all your scores are entered, the Excel analysis tool provides an effect size and information on the confidence in the effect size. Your effect size can be either positive or negative, and there can be confidence in your effect size or no confidence in your effect size. Figure III-1 can help you interpret the four possible effect size results from the Excel analysis tool.

Regardless of the magnitude of effect size (closer to either 1 or –1) or whether it was positive or negative, remember to consider and reflect on other information you have when interpreting the results. Respond to the questions in worksheet III-1 to reflect on your results. Refer to worksheet I-2 in the planning guide when reflecting on results to remind yourself of the context of your groups.

Figure III-1. Interpreting effect sizes



a. If there is no confidence, interpret the results cautiously because they cannot be distinguished from a finding of no effect.

Source: Authors' compilation.

Worksheet III-1. Reflection

1. Consider the results

If you had a positive effect size and the confidence interval was "yes" …

Why do you think the strategy worked well for this group of students?

What else was happening that may have helped the strategy work well (for example, another strategy or resources that you were using that may have also helped)?

If you had a negative effect size and the confidence indicator was "yes" …

Why do you think the strategy did not work well for this group of students?

What else was happening, other than the strategy, that may have caused the result (for example, students were disengaged)?

If you had either a positive or negative effect size and the confidence interval was "no" …

Do you think the strategy worked well with your class? What evidence do you have of this (for example, did you see an increase in student engagement)?

Do you think that the strategy did not work well with your class? What evidence do you have of this? What else was happening, other than the strategy, that may have caused the result (for example, a long break in the unit)?

2. Consider implementation

Do you feel you had a good understanding of how to implement the strategy? If not, what additional information did you need to implement the strategy?

Do you feel you implemented the strategy the way in which it was intended to be implemented? If not, what adjustments did you make?

Do you feel the strategy you chose was appropriate for the class or lesson? Why or why not?

3. Consider the characteristics of the assessments used

Were the content assessment questions clearly aligned to the learning objectives identified in the lesson or unit?

Were enough content assessment questions available that related to each of the learning objectives, indicating that the assessment captured the breadth of knowledge students needed to understand from the unit?

Was there enough time between the pre-test and post-test for students to learn the material?

Were the questions at an appropriate difficulty level to allow for increases in learning to be captured from pre-test to post-test?

4. Consider next steps: based on your responses to the reflection questions, what steps do you plan to take?

If you had positive results, will you:

Implement the strategy the same way in other classes?

Implement the strategy for a longer period of time?

Combine the strategy with other strategies?

If you had negative results, will you:

Adjust the strategy and try it again in the same or another class?

Test the strategy with different classes?

Adjust your assessments to better reflect your content?

Implement the strategy for a longer period of time?

Try a new strategy?

If you had positive or negative effect size results with no confidence, will you:

Look at other evidence of whether the strategy worked?

Test the strategy again in another class, because you believe the strategy worked well?

Adjust the strategy or try a different strategy because you do not think the strategy worked well?

Appendix A. Technical documentation

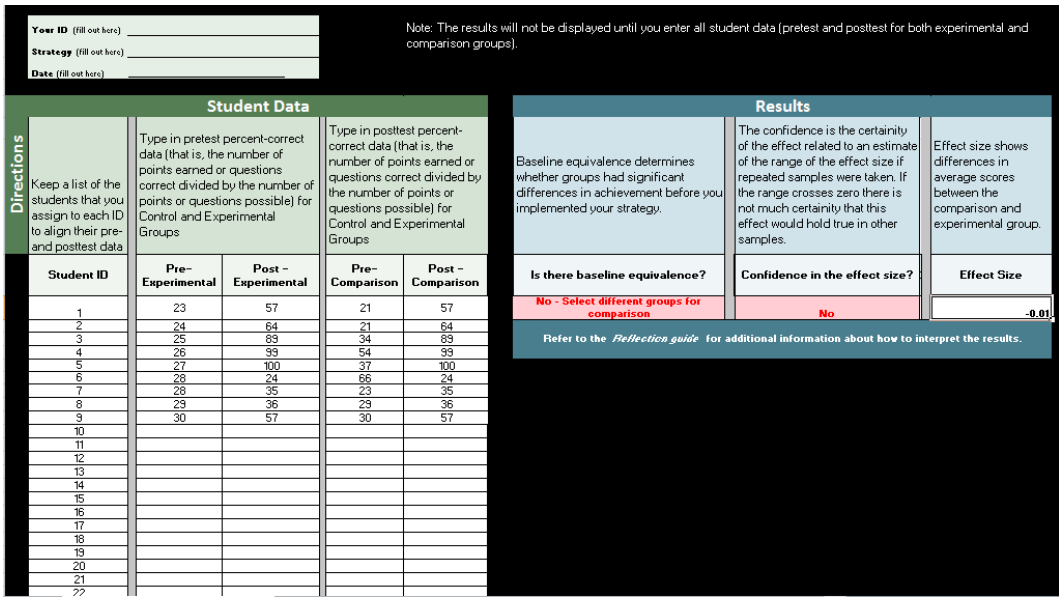
This appendix provides a screenshot of the Excel analysis tool (figure A1) and details the statistical formulas used in it to compute the results: baseline equivalence, effect size, and effect size confidence.

Baseline equivalence

To compute baseline equivalence, the Excel analysis tool calculates an effect size for pre-test differences between comparison and experimental groups using a formula suggested in the What Works Clearinghouse handbook. This formula calculates the difference between the experimental and comparison group pre-test means divided by the pooled standard deviation (What Works Clearinghouse, 2013, p. 22). In the formula below, $\bar{X}_E$ and $\bar{X}_C$ are the adjusted pre-test means for students in the experimental and comparison groups, n_E and n_C are the student sample sizes, and s_E and s_C are the student-level standard deviations of the pre-test. To determine whether baseline equivalence is acceptable, the What Works Clearinghouse guidelines are used: if the effect size differences between experimental and comparison groups are between 0.00 and 0.25, adjustments are made and baseline equivalence is met; if the effect size difference was greater than 0.25, baseline equivalence is not met (What Works Clearinghouse, 2013, pp. 15–16).

$$g = \frac{\bar{X}_E - \bar{X}_C}{\sqrt{\frac{(n_E - 1)\, s_E^2 + (n_C - 1)\, s_C^2}{n_E + n_C - 2}}}$$

Figure A1. Screenshot of the Excel analysis tool



Effect size

The Excel analysis tool uses an effect size formula based on guidance from the What Works Clearinghouse handbook. This formula, known as Hedges' *g*, calculates the difference between the adjusted experimental and comparison group post-test means divided by the pooled standard deviation while accounting for small sample sizes (derived from the unadjusted post-test standard deviations; What Works Clearinghouse, 2013, p. 22). In the formula below, $\bar{X}_E$ and $\bar{X}_C$ are the adjusted post-test means for students in the experimental and comparison groups, n_E and n_C are the student sample sizes, and s_E and s_C are the student-level standard deviations of the post-test.

$$g = \frac{\bar{X}_E - \bar{X}_C}{\sqrt{\frac{(n_E - 1)\, s_E^2 + (n_C - 1)\, s_C^2}{n_E + n_C - 2}}}$$

The formula from Lipsey and Wilson (2001, p. 49) is used to calculate the standard error for each effect size. In the formula below, n_E and n_C are the student sample sizes in the experimental and comparison groups, and g is the effect size. Using the standard error (computed with the formula below), the confidence interval is then computed using $g \pm (1.645 \times SE_g)$.

$$SE_g = \sqrt{\frac{n_E + n_C}{n_E n_C} + \frac{g^2}{2(n_E + n_C)}}$$

$$SE_d = \sqrt{\frac{2(1 - r)}{n} + \frac{d_g^2}{2n}}$$

Effect size confidence

To determine the significance of effect sizes, the confidence intervals are reviewed using a 90 percent confidence level. If the confidence intervals cross zero, the effect size is not significant and is therefore labeled as no confidence.

References

Datnow, A., Park, V., & Kennedy-Lewis, B. (2013). Affordances and constraints in the context of teacher collaboration for the purpose of data use. *Journal of Educational Administration, 51*(3), 341–362. http://eric.ed.gov/?id=EJ1006319

Gallimore, R., Ermeling, B. A., Saunders, B., & Goldenberg, C. (2009). Moving the learning of teaching closer to practice: Teacher education implications of school-based inquiry teams. *Elementary School Journal, 109*(5), 537–553. http://eric.ed.gov/?id=EJ844058

Jaquith, A., & McLaughlin, M. (2010). A temporary, intermediary organization at the helm of regional education reform: Lessons from the Bay Area School Reform Collaborative. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *Second international handbook of educational change* (pp. 85–103). London: Springer International Handbook of Education.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Massell, D. (2001). The theory and practice of using data to build capacity: State and local strategies and their effects. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states* (pp. 148–169). Chicago: University of Chicago Press.

Means, B., Chen, E., DeBarger, A., & Padilla, C. (2011). *Teachers' ability to use data to inform instruction: Challenges and supports.* Washington, DC: U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. http://eric.ed.gov/?id=ED516494

Means, B., Padilla, C., & Gallagher, L. (2010). *Use of education data at the local level: From accountability to instructional improvement.* Washington, DC: U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. http://eric.ed.gov/?id=ED511656

Steiger, J. H. (2004). Beyond the *F* test: Effect size confidence intervals and tests of close fit in the analysis of variance and contrast analysis. *Psychological Methods, 9*(2), 164–182. http://eric.ed.gov/?id=EJ685055

What Works Clearinghouse. (2013). *What Works Clearinghouse procedures and standards handbook, Version 3.0.* Washington, DC: U.S. Department of Education, Institute of Education Sciences. http://eric.ed.gov/?id=ED544775

The Regional Educational Laboratory Program produces 7 types of reports

Making Connections
Studies of correlational relationships

Making an Impact
Studies of cause and effect

What's Happening
Descriptions of policies, programs, implementation status, or data trends

What's Known
Summaries of previous research

Stated Briefly
Summaries of research findings for specific audiences

Applied Research Methods
Research methods for educational settings

Tools
Help for planning, gathering, analyzing, or reporting data or research