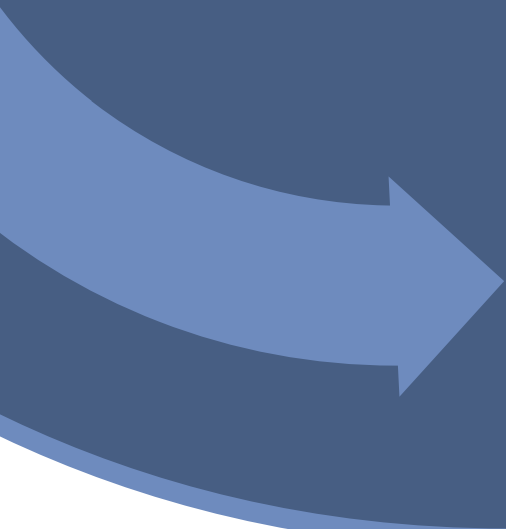




TEACHER **as** **RESEARCHER**

Professional Development Series

TOOLKIT

Contents

Introduction	3
Step 1: Select an Instructional Strategy	4
Instructional Strategy Selection Worksheet	
Step 2: Implement the Strategy	6
Class and Unit Selection Worksheet	
Assessment Planning Worksheet	
Implementation Planning Worksheet	
Step 3: Collect data on strategy implementation	14
Implementation Note Catcher	
Step 4: Analyze the data and reflect on the results	15
Reflection Worksheet	
Technical Documentation	20
References	22

Introduction

When it comes to conducting experiments on the effectiveness of specific instructional strategies to improve student learning in classrooms, teachers are best positioned to find out what works.

At Marzano Research, we have been working to rethink classroom research and empower teachers to test and adjust day-to-day instruction. We call this approach Teacher as Researcher.

To build the skills of teacher researchers, we have developed tools that support classroom-based experiments, helping teachers develop and test hypotheses about what works for their particular students and use results to inform real-time adjustments.

This toolkit provides includes a set of tools worksheets to help teachers engage in the Teacher as Researcher instructional improvement cycle and is organized by its four recurring steps:



Instructional Improvement Cycle

Step 1: Select an instructional strategy

Step 2: Implement the strategy

Step 3: Collect data on strategy implementation

Step 4: Analyze the data and reflect on the results

Flexibility and Benefits

The Teacher as Researcher instructional improvement cycle is meant to be flexible to your needs and circumstances. The instructional improvement cycle can be used by either individual teachers or 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.

Teachers can implement an 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 plan upcoming units (within 1 to 2 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.

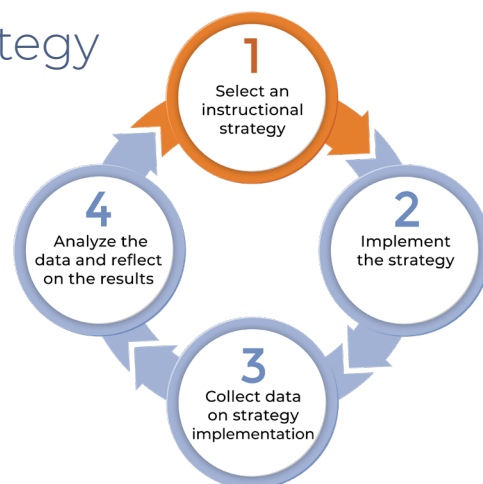
Teachers can implement the cycle as many times as they like to test multiple strategies or to gather additional information about one strategy from different classes of students. An important aspect of the process is that all the students in the cycle are taught by the same teacher—this means the instructional strategy itself, rather than differences among teachers, can be tested.

In short, the instructional improvement cycle leads to:

- deliberate application and study of one strategy and the effect it has on student learning
- efficiency through the intentional replication of what works for you and your students
- reflection on practice based on systematic data collection and analysis processes

Step 1: Select an Instructional Strategy

Use the *Instructional Strategy Selection Worksheet* to select a Route that will help identify a strategy that will best suit your classroom.



Instructional Strategy Selection Worksheet

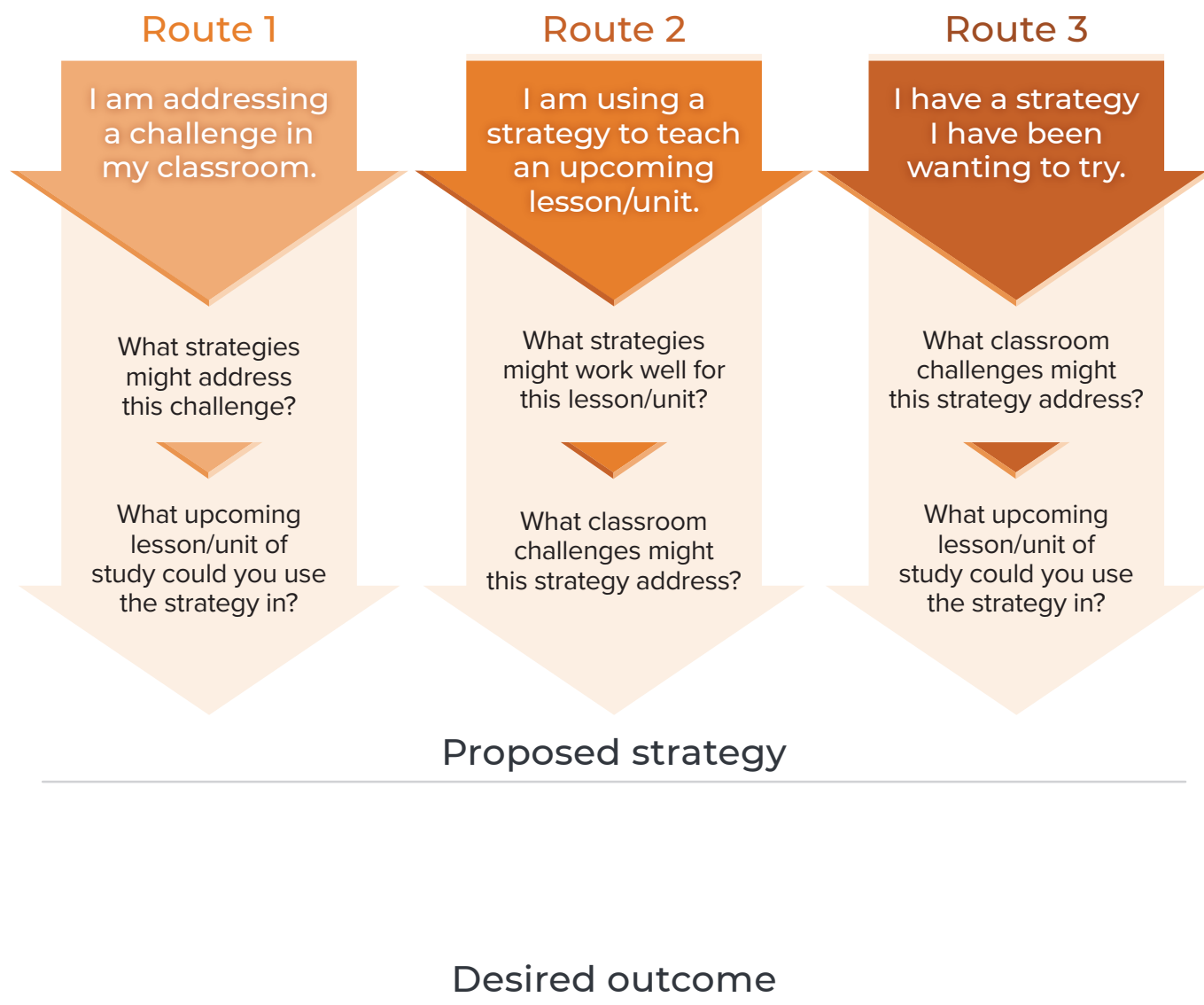
Use the prompts in this worksheet to select an instructional strategy for your Teacher as Researcher instructional improvement cycle.

Not everything that you do in a classroom should be considered an instructional strategy. Use the criteria below to help you determine whether the activities you are thinking about should be considered as strategies.

- You can describe the steps you take to execute the strategy.
- The strategy has an identifiable beginning and end.
- You can describe the outcome the strategy should have on students' processing of information, understanding of specific content, execution of specific procedures, or their behavior.

Selection Guidance

The figure below describes three potential pathways to identify and select an appropriate strategy. You can choose a current challenge in your classroom (Route 1), identify a strategy that aligns with an upcoming lesson (Route 2), or select a strategy you want to try and align it to challenges and upcoming lessons (Route 3). After selecting a strategy, clearly describe the desired outcome of the strategy.



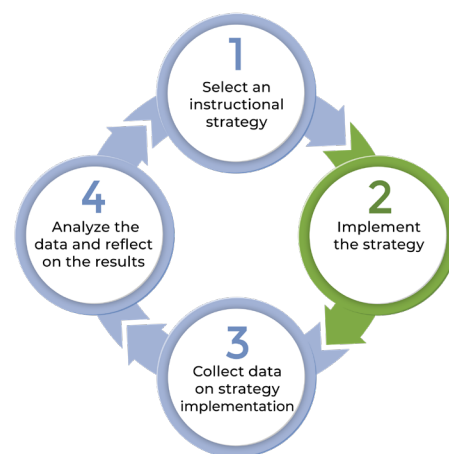
Step 2: Implement the Strategy

Once you have selected a strategy, you need to decide whether you are using a **single-group design** or a **multi-group design**. Select a study design that is most appropriate for your classroom(s) using the descriptions provided below.

Use the *Class and Unit Selection Worksheet* to determine where and when to implement the instructional improvement cycle.

Use the *Assessment Planning Worksheet* to select your content test. The same test will be administered twice (preassessment and postassessment) for the multiple-group design or three times (preassessment, midassessment, and postassessment) for the single-group design.

Use the *Implementation Planning Worksheet* to plan when and how to use the target strategy with the experimental or comparison group. The strategy is used throughout a brief unit or time period (for example, 2 weeks). This does not mean that the strategy must be used each day; rather, the strategy should be used appropriately for the content and activities planned for the unit.



Single-Group Design

Characteristic	Description
Number of groups/classes	A single class or group of students.
Implementation	During the first half of the lesson/unit, the group gets regular instruction. This serves as the baseline phase for this group. During the second half of the lesson/unit, the group gets instruction with the strategy. This serves as the intervention phase for this group.
Content preassessment and postassessment administration	Before beginning instruction on the lesson/unit, the group takes the content assessment that covers material from the whole unit to serve as a preassessment. The group takes the same assessment in the middle of the unit (midassessment) after the time of regular instruction and before the instruction with the target strategy. The group takes the same assessment at the end of the lesson/unit after instruction with the target strategy (postassessment).
Question addressed	How well does a single group of students achieve when they receive a strategy as compared to when they did not receive the strategy?
Limitations	The learning targets from the first half of the unit may differ from the learning targets in the second half of the unit, and that difference may explain the findings rather than the strategy used.

Implementation Process for Single-Group Designs

If you are using a single-group design, complete the following steps:

Step 1. To collect baseline scores, administer the content assessment before you begin the first half of the lesson when students get regular instruction. Record each student's percent correct in the preassessment column in the Single-Group Analysis Tool.



Step 2. Administer the content assessment at the midway point of instruction of the lesson/unit just before your group starts to get instruction using the target strategy. This assessment will serve as the postassessment for regular instruction and the preassessment for target strategy instruction. Record each student's percent correct in the midassessment column in the Single-Group Analysis Tool.



Step 3. Teach the content using the target strategy during the second half of the lesson/unit (intervention).



Step 4. Administer the content assessment after the second half of the lesson/unit (intervention) is complete to collect postassessment scores from when the group received instruction using the target strategy. Record each student's percent correct in the postassessment column in the Single-Group Analysis Tool.



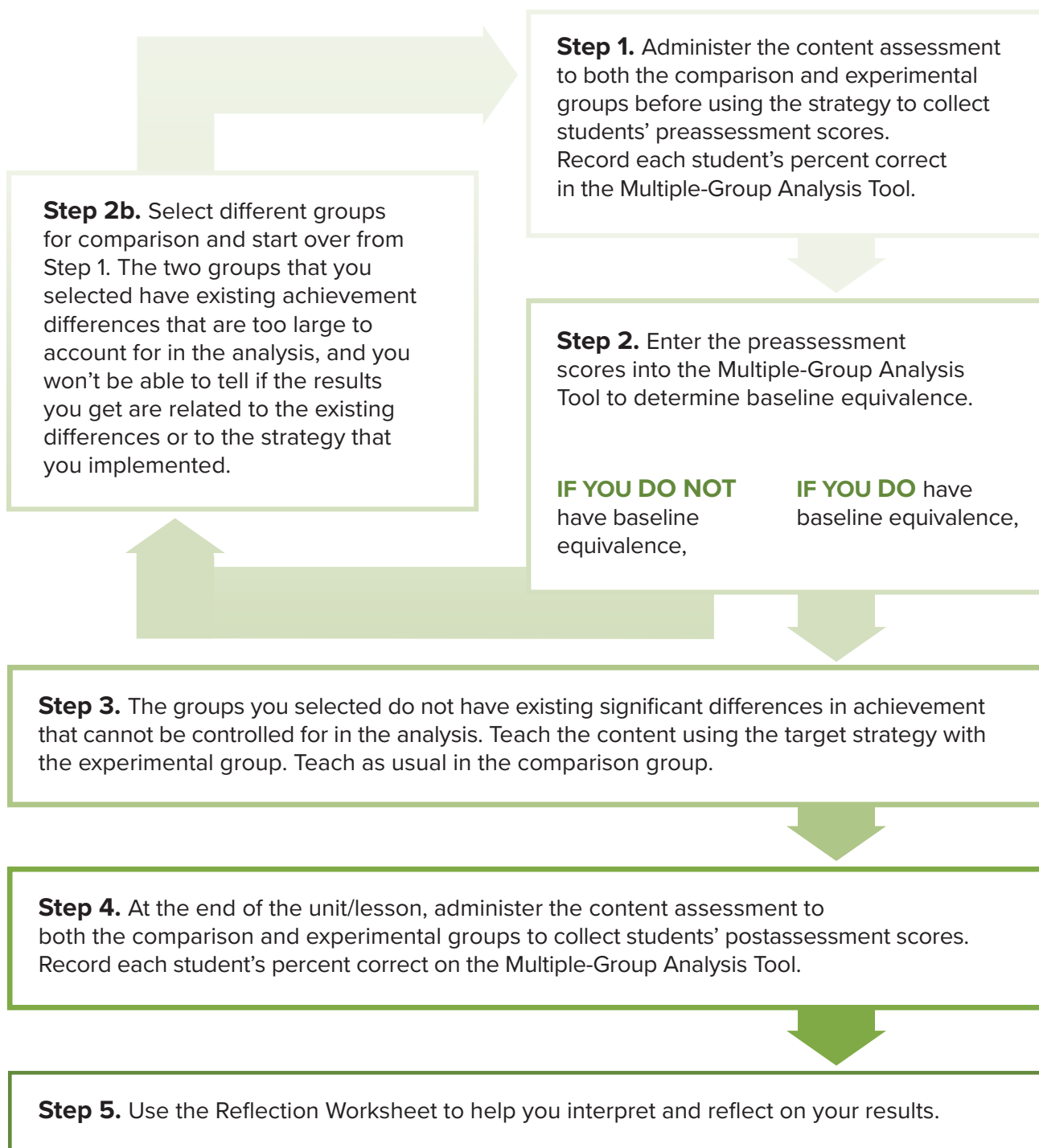
Step 5. Use the Reflection Worksheet to help you interpret and reflect on your results.

Multiple-Group Design

Characteristic	Description
Number of groups/classes	Two different classes or groups of students.
Implementation	Group 1 gets regular instruction. This group serves as the comparison group. Group 2 gets instruction on similar content with the target strategy. This group serves as the experimental group.
Content preassessment and postassessment	Both groups take the same content assessment before the lesson/unit (preassessment) and the same test at the end of the lesson/unit (postassessment).
Question addressed	How well does a group of students achieve when they receive the target strategy as compared to a group of students who do not receive the strategy?
Limitations	The findings on the content assessment may be a result of preexisting differences between the two groups of students and may not be a result of using the strategy.

Process for Implementing Multiple-Group Designs

If you are using a multiple-group design, complete the following steps:



Class and Unit Selection Worksheet

Use the prompts in this worksheet to describe the classes and units you will use for a Teacher as Researcher instructional improvement cycle. If you will be using a multiple-group design, select a class that you will use as the experimental group and another class that you will use as a comparison group. When selecting classes for a multiple-group design instructional improvement cycle, it is important to ensure that the experimental and comparison classes are demographically similar and will receive the same lessons and content. If you are using a single-group design, only complete a description for Class 1 in the table below.

Class 1: Experimental group (the class in which you will use the instructional strategy)

Number of students

Grade(s)

Course name or subject

Unit topic

Demographics (e.g., percentage of students eligible for free and reduced-price lunch, SpEd, or English learners)

Unit or lesson goals (what you want students to know and be able to do by the end of the unit)

Instructional strategy to be tested (the strategy described in the Instructional Strategy Selection Worksheet)

Desired outcome (what you are addressing with the strategy)

Other strategies you plan to use during the unit

Class 2: Comparison group (the class in which you will not use the instructional strategy)

Number of students

Grade(s) (should be the same as the experimental group)

Course name or subject

Unit topic (should be identical to the experimental group)

Demographics (should be similar to the experimental group)

Unit or lesson goals (should be identical to the experimental group)

Instructional strategies you plan to use during the unit

Assessment Planning Worksheet

Use the guidance and prompts in this worksheet to identify or develop an assessment for use in your Teacher as Researcher instructional improvement cycle.

If you are implementing a multiple-group design instructional improvement cycle, you will administer the same assessment at the beginning and end of the unit that you will teach during the cycle in both your experimental and comparison class.

If you are implementing a single-group design instructional improvement cycle, you will administer the same assessment three times during the unit you will be teaching during your cycle. You will teach the unit without using the strategy you plan to test in between administration of the assessment at the beginning and middle of the unit, then implement the strategy in between administration of the assessment at the middle and end of the unit.

Assessment Selection Checklist

- ☐ Assessment covers content from the entire unit.
- ☐ The assessment questions are clearly aligned with the unit learning objectives.
- ☐ The assessment has enough questions to measure a progression of learning over time.
- ☐ The assessment is at a level of difficulty that allows for accurate measurement of learning in between each administration (i.e., it is not too hard to measure growth in between a pre- and midassessment and not so easy that students are unable to demonstrate growth).
- ☐ The assessment is designed for multiple administrations (i.e., two times in a multiple-group design and three times in a single-group design).

Assessment Design Template

Use the table below to identify or develop an assessment that addresses the criteria described above and that covers all the unit learning objectives.

Learning objective	Number of questions in the assessment
1.	
2.	
3.	
4.	
5	

Implementation Planning Worksheet

Use the prompts in this worksheet to help plan for implementation of your instructional improvement cycle. Before completing the planning worksheet, consider reviewing research and evidence regarding your instructional strategy and/or reflecting on your previous experience using the strategy to determine how often to use the strategy and how much time to allot for each use of the strategy.

Planning questions

Implementation plans

Are you implementing a single-group or multiple-group design instructional improvement cycle?

Single-Group or Multiple-Group

How long is your instructional improvement cycle?

On what dates will you teach the lesson or unit?

Describe the assessment you plan to administer (What content is assessed?
What is the format of the assessment?)

On what date will you administer the preassessment?

On what date will you administer the midassessment? (applicable to single-group designs)

On what date will you administer the postassessment?

How often will you use the strategy (e.g., daily, after each reading passage)?

How much time will you allocate for implementation of the strategy each time you use it (e.g., 10 minutes in groups, 15 minutes during problem solving, 5 minutes during direct instruction)?

Step 3: Collect data on strategy implementation

Use the *Analysis Tools* to analyze student assessment data and see the results on how well the strategy worked. Teachers assess students before and after implementing the instructional strategy and enter the scores into the Single or Multiple Group Analysis Tool. The multiple-group analysis produces three results: baseline equivalence, confidence level, and effect size. The single-group analysis produces three results: baseline effect size, intervention effect size, and confidence level. Formulas for the analyses are provided in the technical documentation. The *Implementation Note Catcher* should be used during implementation to document instructional strategy implementation.



Implementation Note Catcher

How is implementation going overall?

What successes have you experienced?

What challenges have you experienced?

What adaptations have you made to the instructional strategy?

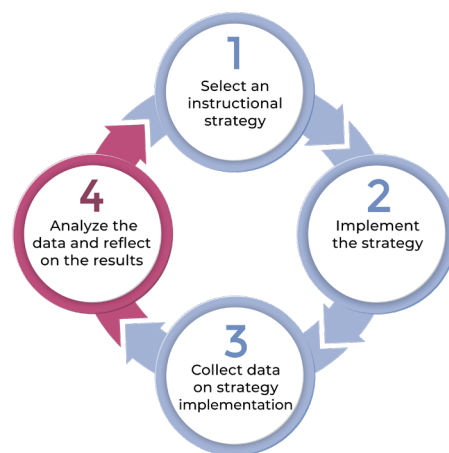
What else is happening that would be important to keep track of?

If you're doing a multiple-group design, what are you noticing about the students who aren't part of the experimental group?

If you're doing a single-group design, what did you notice during the baseline phase in which you were not implementing the strategy?

Step 4: Analyze the data and reflect on the results

After entering your data into the Analysis Tool, look at the effect size(s) and the confidence in the effect size. The data in these cells provide information about how the implementation of a strategy may relate to student achievement in a given situation. Review the guidance below and use the *Reflection Worksheet* to interpret your results.



Interpreting Single-Group Design Results

The **baseline effect size** tells you how well your students performed without the strategy. A positive effect suggests that they performed well. Remember that the effect is smaller as the effect size is closer to zero, so an effect size of 0.01 indicates a smaller effect than an effect size of 1.00.

The **intervention effect size** tells you how well your students performed with the strategy. A positive effect suggests that they performed well. Remember that the effect is smaller as the effect size is closer to zero, so an effect size of 0.01 indicates a smaller effect than an effect size of 1.00.

The **confidence in the effect size** tells you the level of certainty in the effect size by estimating what the range of the effect size would be if repeated samples were taken (confidence intervals). For example, an effect size of 0.69 might have a range from 0.05 to 1.00. If the range crosses zero (goes from negative to positive)—for example -0.1 to 0.5—then you cannot have certainty that the intervention would have a positive or negative effect if implemented in different samples.

While the findings from this one study cannot be used to conclusively state whether or not a strategy is effective, they do give some information about how that strategy worked in your setting. Formulas for the analyses are provided in the Technical Documentation.

What Do Your Results Mean?

Intervention		Baseline		Suggests that students...
--------------	--	----------	--	---------------------------

Positive
Intervention
effect size



Baseline
effect size



Performed better when the
strategy was implemented*

If the intervention effect size is positive and larger than the baseline effect size, this suggests that students performed better when the strategy was implemented.

**If the analysis tool indicated that there is not confidence in the effect size, then you may not see the same results in another sample.*

Intervention
effect size



Positive
baseline
effect size



Did not perform better when the
strategy was implemented*

If the intervention effect size is smaller than the baseline effect size and the baseline effect size was positive, this suggests that students did not perform better when the strategy was implemented.

**If the analysis tool indicated that there is not confidence in the effect size, then you may not see the same results in another sample.*

Negative
Intervention
effect size



Negative
baseline
effect size



Did not perform well with
or without the strategy*

If both the intervention effect size and the baseline effect size were negative, this suggests that students did not perform well with or without the strategy.

**If the analysis tool indicated that there is not confidence in the effect size, then you may not see the same results in another sample.*

Interpreting Multiple-Group Design Results

The **effect size** “tells you how much larger (or smaller) you might expect the average score to be in a class where students use a particular strategy as compared to a class where the strategy is not used” (Marzano, 2007, pp. 10–11). Remember that the effect is smaller as the effect size is closer to zero, so an effect size of 0.01 indicates a smaller effect than an effect size of 1.00.

The **confidence in the effect size** tells you the level of certainty in the effect size by estimating what the range of the effect size would be if repeated samples were taken (confidence intervals). For example, an effect size of 0.69 might have a range from 0.05 to 1.00. If the range crosses zero (goes from negative to positive)—for example, -0.1 to 0.5—then you cannot have certainty that the intervention would have a positive or negative effect if implemented in different samples.

While the findings from this one study cannot be used to conclusively state whether or not a strategy is effective, they do give some information about how that strategy worked in your setting. Formulas for the analyses are provided in the Technical Documentation.

What Do Your Results Mean?

Effect size	Confidence	Suggests that the strategy was...
+ Positive	and Yes	RELATED to increased achievement in your sample and a POSITIVE EFFECT may be seen in other samples
+ Positive	and No	RELATED to increased achievement in your sample, but we are UNCERTAIN OF EFFECTS on other samples
– Negative	and Yes	NOT RELATED to increased achievement in your sample and a NEGATIVE EFFECT may be seen in other samples
– Negative	and No	NOT RELATED to increased achievement in your sample and we are UNCERTAIN OF EFFECTS on other samples

Reflection Worksheet

Regardless of whether the effect size was positive or negative, teachers should also reflect on other information when interpreting the findings.

Reflection Responses

If you had **POSITIVE** results . . .

Why do you think the strategy worked well?

What else was happening, other than the strategy, that may have caused the result? (For example, were you using another strategy that may have also helped?)

If you had **NEGATIVE** results . . .

Why do you think the strategy did not work well?

What else was happening, other than the strategy, that may have caused the result?
(For example, were several students absent during the week that you implemented the strategy?)

Consider how the strategy was implemented

Do you feel like you had a good understanding of how to implement the strategy?

Do you feel like you implemented the strategy the way it was intended to be implemented?

Do you feel like the strategy you chose was appropriate for the class and/or lesson?

Consider the characteristics of the assessment used

Did the content assessment have questions that were clearly aligned to the learning objectives identified in the lesson/unit?

Did the content assessment have enough questions to fully capture student learning?

Was there enough time between the preassessment and postassessment for students to learn the material?

Were the questions at an appropriate difficulty level to allow for increases in learning to be captured from preassessment to postassessment? (For example, were some questions too difficult, so that even on the postassessment most students couldn't answer correctly, therefore making the gains look smaller than they may have been? Or were some questions too easy, so students got them right both at preassessment and postassessment, making the gains look smaller than they were?)

Next Steps

Based on your responses to the reflection questions, what steps do you plan to take? (For example, consider adjusting the strategy and trying it again, implementing the strategy for other lessons, adjusting your assessments, implementing the strategy for a longer period of time, or using other strategies to test.)

Technical Documentation

Effect size significance

To determine the significance of effect sizes, the confidence intervals were reviewed. If the confidence intervals crossed zero, the effect size was not significant.

Single-group analysis

To compute effect sizes, we used Becker's (1988) formula for a standardized mean gain as presented in Lipsey and Wilson (2001). A standardized mean gain is computed for the comparison and experimental condition. The standardized mean gain is the difference between the average postassessment and preassessment score divided by the pooled standard deviation. In this formula, χ_{T2} is the mean postassessment score, χ_{T1} is the mean preassessment score, s_g is the standard deviation of the gain score, and r is the correlation between the preassessment and postassessment scores.

$$d_g = \frac{\chi_{T2} - \chi_{T1}}{s_g / \sqrt{2(1-r)}}$$

The formula from Lipsey and Wilson (2001) was used to calculate the standard error for each standardized gain score. In the formula below, r is the correlation between the preassessment and postassessment scores, η is the number of students, and d_g is the standardized gain score. Using the standard error, the confidence interval was then computed using $d \pm (1.96 \times SE_d)$.

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

Multiple-group analysis

To compute effect sizes, the same formula for effect sizes used by What Works Clearinghouse was applied. This formula calculates the difference between the adjusted experimental and comparison group postassessment means divided by the pooled standard deviation (derived from the unadjusted postassessment standard deviations) (What Works Clearinghouse, 2013). In the formula below, χ_E and χ_c are the adjusted postassessment means for students in the experimental and comparison groups, η_E and η_c are the student sample sizes, and s_E and s_c are the student-level standard deviations of the postassessment.

$$g = \frac{\chi_E - \chi_c}{\sqrt{\frac{(\eta_E - 1)s_E^2 + (\eta_c - 1)s_c^2}{\eta_E + \eta_c - 2}}}$$

The formula from Lipsey and Wilson (2001) was 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 was then computed using $g \pm (1.645 \times SE_g)$.

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

Baseline equivalence: multiple-group design

To compute baseline equivalence, an effect size for preassessment differences between comparison and experimental groups was calculated using the What Works Clearinghouse formula. This formula calculates the difference between the experimental and comparison group preassessment means divided by the pooled standard deviation (What Works Clearinghouse, 2013). In the formula below, χ_E and χ_C are the adjusted preassessment means for students in the experimental and comparison groups, η_E and η_C are the student sample sizes, and s_E and s_C are the student-level standard deviations of the preassessment. To determine if baseline equivalence was acceptable, the What Works Clearinghouse guidelines were used: If the effect size differences between experimental and comparison groups were between 0.00 and 0.25, then adjustments were made and baseline equivalence was met; if the effect size difference was greater than 0.25, then baseline equivalence was not met (What Works Clearinghouse, 2013).

$$g = \frac{\chi_E - \chi_C}{\sqrt{\frac{(\eta_E - 1)s_E^2 + (\eta_C - 1)s_C^2}{\eta_E + \eta_C - 2}}}$$

References

Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, 41(2), 257–278. <https://doi.org/10.1111/j.2044-8317.1988.tb00901.x>

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE.

Marzano, R. (2007). *The art and science of teaching*. Solution Tree.

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>