Ed Tech Rapid Cycle Evaluation Coach

Guide: Think About How You Will Use the Results

We'll use this tool to think about how the answer to your research question will inform the decisions you make regarding your technology. The following questions will help you prepare you to think about the relationship between your results and the technology costs and stakes associated with your decision. To complete this tool you will have to consider the following:

What would success look like?

How much risk are you willing to take when making a decision or a recommendation?

> **CONSIDER:** When making decisions, it is important to think about the costs of the technology you're considering and the risks associated with your decision.
> For example, you might be trying to decide:
>
> • Whether to keep paying for a software tool you're already using
>
> • Whether to buy licenses for a tool you're piloting on a trial basis
>
> • Whether to change how teachers and/or students use the tool, based on a new approach tested in your evaluation

A. WHAT IS THE COST OF THE EDUCATIONAL TECHNOLOGY?

The cost of a technology will probably be a factor in your decision about how big its effect has to be in order to be a good value. Cost might also influence how certain you want to be that the technology has the desired effect.

The cost of a technology could simply be the cost per user. It could also be the savings realized by replacing one technology with another less expensive technology. Cost can also be measured in something other than dollars, such as teacher preparation time saved.

How much do you pay (or save) per student, teacher, or school to use the educational technology?

Examples

- It costs about $20 per student to use the educational technology.
- It saves about $1,000 per classroom to use the educational technology.
- We're not sure. We believe using this tool will save teachers 20 minutes per day on preparation time.

Ed Tech Rapid Cycle Evaluation Coach

B. WHAT WOULD SUCCESS LOOK LIKE?

What is the unit used to measure your outcome of interest?

Examples

- Test-score points
- Percentage points

What is the direction of intended change?

Do you hope to increase or decrease your outcome of interest?

By how many units would your outcome have to increase or decrease for you to consider the change meaningful?

To answer this question, think about what magnitude of change would or wouldn't be meaningful in your context. If you hope to improve test scores in reading, would you view a 1-point increase as essentially the same as 0 (or no change)? What about 5 points? How large would the change need to be for it to be meaningful? You should also consider factors such as the maximum possible points and current performance. A 10-point increase on a test with a maximum possible score of 100 is very different from a 10-point increase on a test with a maximum score of 500. Similarly, if students score 80 out of 100 on average, there is less room for improvement than if they score 50 out of 100 on average—and that might affect the size of the improvement that you would find meaningful.

The threshold you set here is known as the minimum meaningful effect, or MME. The Coach will use this value to calculate three probabilities for the educational technology: (1) the probability of having a positive effect at least as large as the MME, (2) the probability of having a negative effect at least as large as the MME, and (3) the probability that the effect is not meaningful (in other words, that there is a positive or negative effect smaller than the MME). See Exhibit 1.

Ed Tech Rapid Cycle Evaluation Coach

Exhibit 1. Distribution of evaluation results
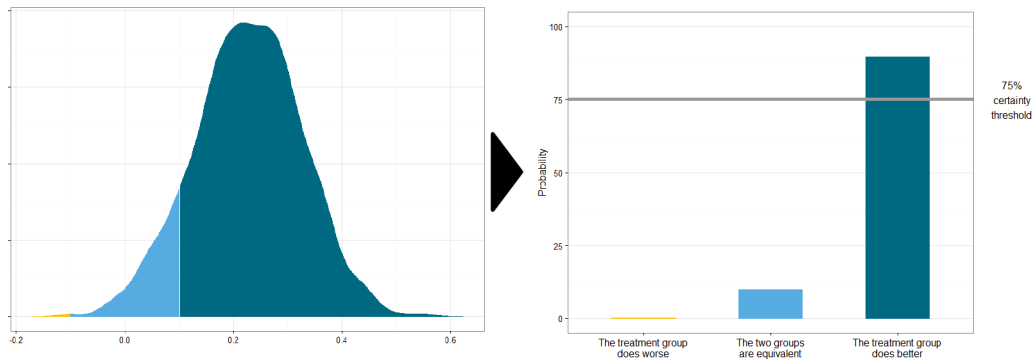
The posterior distribution of outcomes will be divided into three areas based on the minimum meaningful effect chosen, in this case 0.1 units. The area under the curve in each section represents the probability of observing that outcome. From left to right, the first area shows the probability of having a negative effect larger than -0.1 units. This probability is very small. The second area shows the probability of having an effect between -0.1 and 0.1 units, any change in this range is viewed as practically equivalent to 0 and is not meaningful. The probability of this is larger than the first area, but still small. The third and final area represents the probability of an outcome larger than 0.1 units. The probability of this outcome is very large, representing 88%. In the Coach, the area under each section of the curve will be presented to you as a bar chart, with the height of each bar representing the probability of a given category of outcomes. These bars are then compared against your certainty threshold. If any bar exceeds the threshold, the group it represents will be used to generate your conclusion. For example, here the probability for an outcome where the treatment group performs better than the comparison group, by at least 0.1 units, is 88%. That is larger than your certainty threshold of 75%, so the Coach would conclude that it is likely that the treatment group performs better than the comparison group. If no bar exceeds the threshold your results will be deemed inconclusive.

Consider: The MME could be 0 or even negative. A 0 means that you would consider any change in the outcome to be meaningful. If you select 0, the Coach will only calculate two probabilities: (1) the probability of having a positive effect and (2) the probability of having a negative effect. If the technology saves you a lot of money, you might be willing to consider any change that is greater than a small negative number to be a success.

C. HOW MUCH CERTAINTY DO YOU NEED?

Rarely do we have enough evidence to be certain that an educational technology produces the desired results. Most of the time, we have to make choices with incomplete evidence. You have to decide how much uncertainty you are willing to tolerate when making a decision based on this analysis.

We make these types of decisions all the time. For example, you make a choice under uncertainty when you decide whether to take an umbrella after looking at the weather forecast. Some people will take an umbrella if the probability of rain is 10 percent, whereas others might do so if the probability of rain is 60 percent. Different people have different thresholds for how much risk they are willing to take. And you may have a different threshold for different kinds of decisions.

As you think about what the probability threshold would be for you to be comfortable making a given decision, consider the stakes involved. You might want a higher level of certainty if your results will influence a high-stakes decision than if your results will be used for a lower-stakes decision. For example:

- High stakes—if you are deciding whether to implement a district-wide curriculum
- Low stakes—if you are deciding how to use a small amount of class time

Though it might be tempting to set a high probability threshold for any kind of evaluation, it is important to note that the higher the threshold, the harder it will be to meet. This might not necessarily be because the technology isn't working, but it could be driven by the number of users you are testing the technology with or your unit of assignment. With a very small number of users, you will be unlikely to achieve high confidence even if the product is very effective.
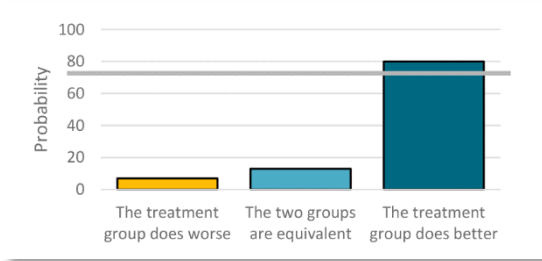
The RCE Coach will compare each of the three probabilities calculated based on your MME to the certainty threshold you set here. If the probability of a given effect exceeds the threshold, the Coach will conclude that the given effect is the likely outcome. If none of the probabilities exceed the threshold, the Coach will say that the results are inconclusive.
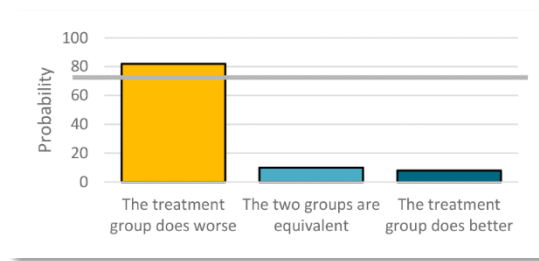
Ed Tech Rapid Cycle Evaluation Coach

D. WHAT WILL YOU DO IF . . . ?

This RCE will provide you with one of the following four answers to your research question. Think about what you will do if . . .
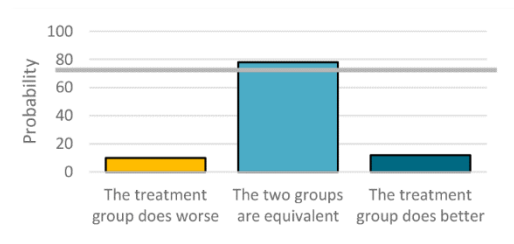
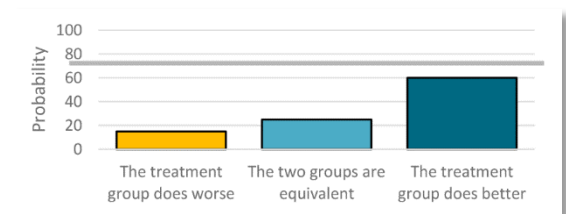**1.** The data suggest that the treatment group does better than the comparison group

2. The data suggest that the treatment group does worse than the comparison group

3. The data suggest that the two groups are equivalent

4. The results are inconclusive

Examples

If the treatment group will likely do better than the comparison group . . .

- *We will roll out eZumi to all students in grades 3 through 5.*

If the treatment group will likely do worse than the comparison group . . .

- *We will not renew our eZumi license for the next school year.*

If the two groups are likely to be equivalent . . .

- *We will collect more data to see if one technology works better for students with specific characteristics.*

If the results are likely to be inconclusive . . .

- *We will collect more data to get a higher level of certainty about our results.*