



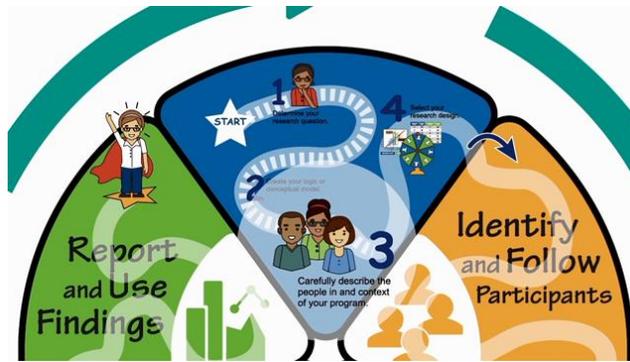
Visit Series website

Video Script: The Counterfactual Condition

	STORYBOARD IMAGE	NARRATION
1.		Welcome again to the 5 minute evaluation resource series,
2.		a series designed to help people passionate about education build evidence for educator effectiveness.
3.		This video is designed to describe the importance of Step 3 in the Plan and Design stage:

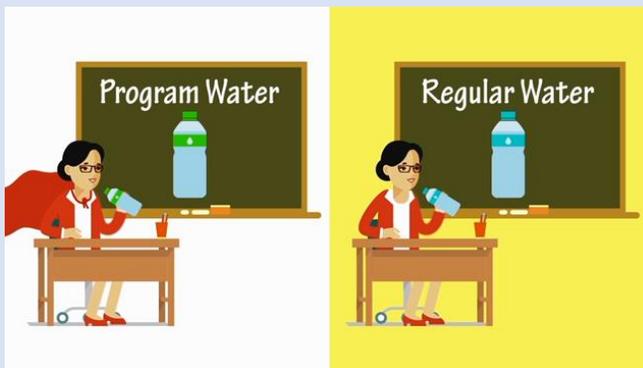


4.



“Describe participants and context to identify a valid comparison group.”

5.



To determine the impact of your program, we want to know what would have happened to the people in your intervention if they had never received your program.

6.



This is what evaluators often call ‘the counterfactual condition’ or ‘business as usual.’ This concept is really important.

7.



Without trying to figure out what would have happened without your intervention, we can’t say anything about its impact on the participants who received it.



8.



Imagine that we are time/space explorers in a superhero universe and we discover an alternate universe exactly like ours – same people, same environment, same everything.

9.



Just one thing would be different – your intervention. Let's call this other universe the counterfactual universe.

10.



To figure out if your intervention made a difference, we would compare the outcomes of the people in the counterfactual universe with the same people in our universe.

Since the only difference in the two worlds is your program, then any differences we identify must be the direct result of your intervention.

11.



But, of course, we don't have access to a counterfactual universe, so we have to do the next best thing: find a group of people who are as closely matched in characteristics, interests, and context as the people in your intervention.



12.



So, how do we do that?

This question is one of the hardest and most important questions you will have to address to ensure your study meets WWC standards.

13.



Imagine that you have this great professional development program for 9th grade teachers.

In the summer, teachers meet for 3 weeks to learn about and practice the skills addressed in your program, and check in with mentors during the school year.

14.



Your task is to find a similar group of teachers who teach a similar group of students and who work in similar schools, but who aren't involved in a similar professional development program.

15.

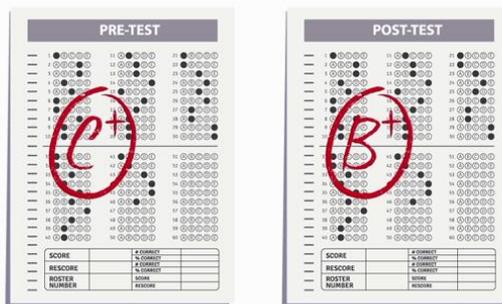
Some popular strategies
that won't help you determine impact

Because this is so difficult, people often look for creative ways to evaluate impacts of programs.

Let's look at some popular strategies that, unfortunately, won't help you determine impact.



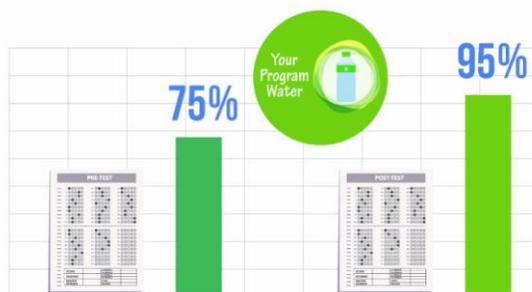
16.



One is using a pre-test and post-test on the teachers' students.

Let's play this out a bit.

17.



You give the students a pretest at the beginning of the school year and they score an average of 75%; at the end of the school year, you test the students again, and they score an average of 95%.

The students improved!
Did your intervention work?

18.



We don't know!

We expect the students will improve naturally as they mature and learn from the world around them.

We don't know how much, or if any of that improvement was because of your intervention.

It's impossible to tell.

19.

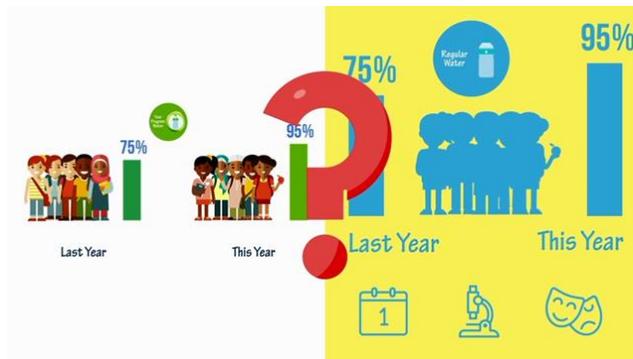


Another creative option is to look at the same teachers' 9th grade students' scores the year before. Then, they scored 75% at the end of the school year on a similar test, and this year a new cohort of students scored 95%.

So doesn't that give some indication your intervention worked?



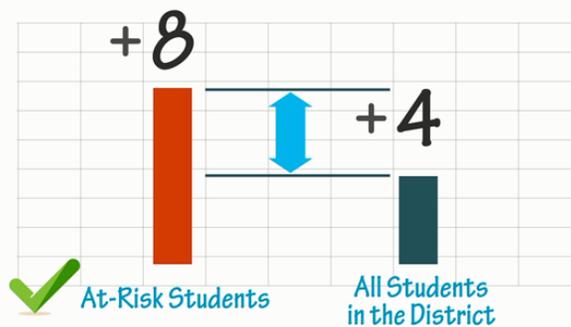
20.



Well, unfortunately, not really.

Both the teachers and students have been exposed to different events, different school initiatives, or other activities that may have influenced the current students' scores.

21.

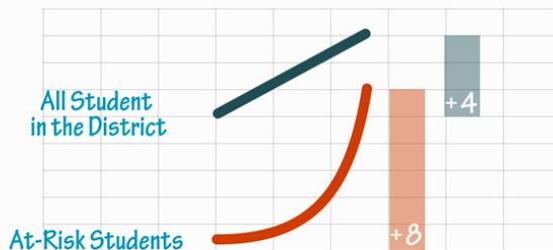


Let's do one more.

Your intervention is targeted towards teachers who teach the most at-risk students in the district. Someone suggests that you compare their student outcomes to the rest of the district.

If your students have greater test score gains than students in the rest of the district, that should indicate that the program worked, right?

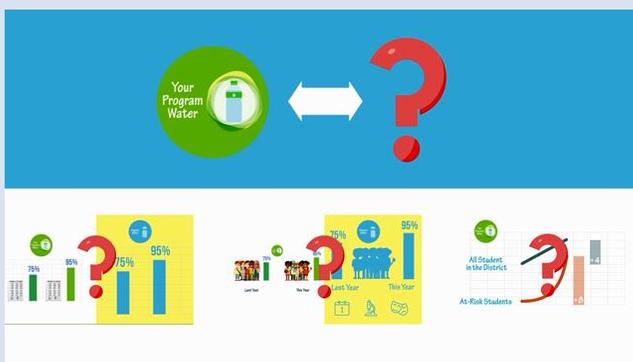
22.



Not necessarily.

The at-risk students might appear to be making greater gains because they had more room for improvement.

23.



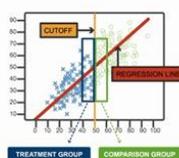
Like our other examples, we still don't know if their gains are due to the program or for other reasons.



24.

WWC

What Works Clearinghouse

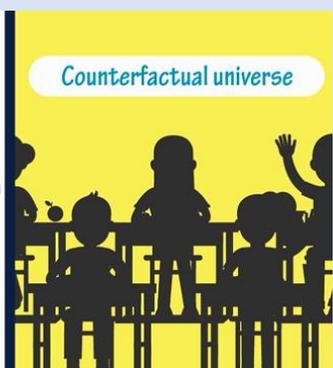
TREATMENT SCHOOL	PRETEST SCORE	COMPARISON SCHOOL	PRETEST SCORE
1	70	1	70
2	95	2	92
3	77	3	81
4	84	4	87
5	74	5	70
AVERAGE	80	AVERAGE	80

So, what will work?

WWC allows for three evaluation designs—we mentioned these in our earlier videos:

25.

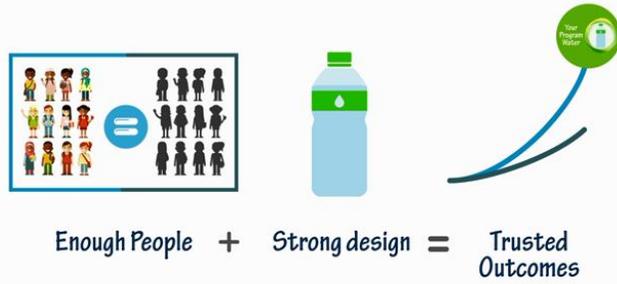
- ✓ Randomized Controlled Trial
- ✓ Regression Discontinuity Design
- ✓ Matched Comparison Design



- randomized controlled trials,
- regression discontinuity designs,
- or matched comparison designs.

Each design has a different way to help you identify your counterfactual universe.

26.

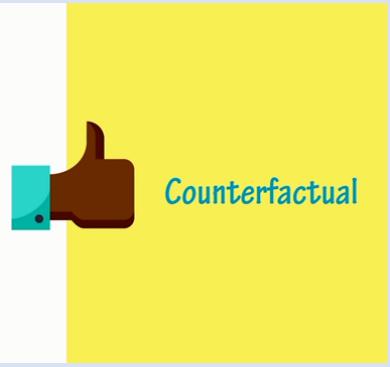


Enough People + Strong design = Trusted Outcomes

The underlying theory at work here is that if you have enough people with similar characteristics as the people in your program, then you can create two equivalent groups.

Differences in outcomes that may emerge between these two equal groups can be attributed to your program.

27.



For these designs to work, though, your evaluator needs to get the counterfactual right.



28.



This can be a little tricky because education interventions often occur at the teacher, school, or district level, but the purpose of the intervention is to make an impact at the student level.

Evaluators call these different levels 'nesting.'

29.



You need to help your evaluator figure out which nesting levels could make a difference in how effective your program is.

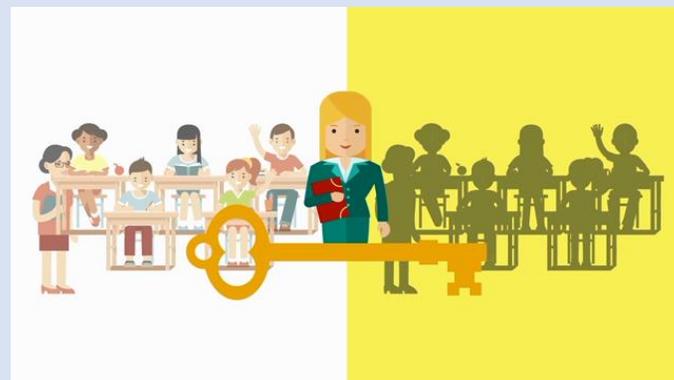
Characteristics of students, teachers, schools, districts, and even states are all potentially important to whether your program succeeds or not.

30.



Thankfully, your evaluator has access to some pretty cool statistical techniques to take these nesting issues into account.

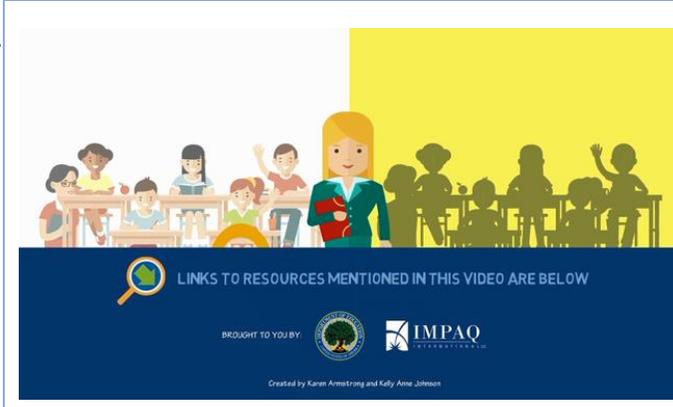
31.



However, you are the key for making sure the counterfactual group is truly comparable to the people in your intervention.



32.



For more information on this and other evaluation issues, please check out the resources below and in the evaluation resource database.

Evaluation Resource Database: <https://innovation.ed.gov/what-we-do/teacher-quality/supporting-effective-educator-development-grant-program/evaluation-resources/>