U.S. Department of Education
http://www.ed.gov/

The Power of Sample Size

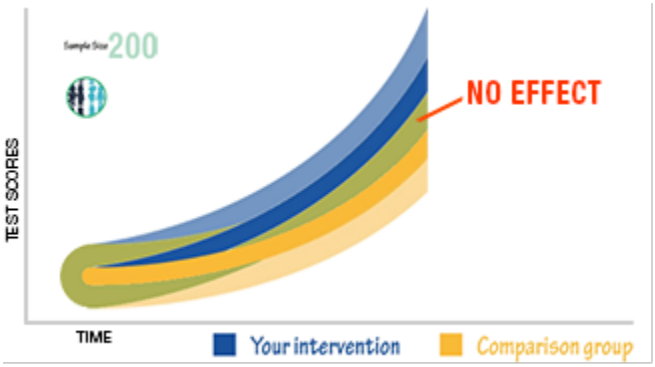
When you conduct an evaluation, you need to have enough people in your study to ensure that your results are meaningful. The number of people in your study – people who will be participating in your program and people who will be in your comparison group – is called the study's sample size.

Sample size is important. If you have too few people in your study your evaluation will not be able to detect program effects even when your program has actually had a meaningful impact.

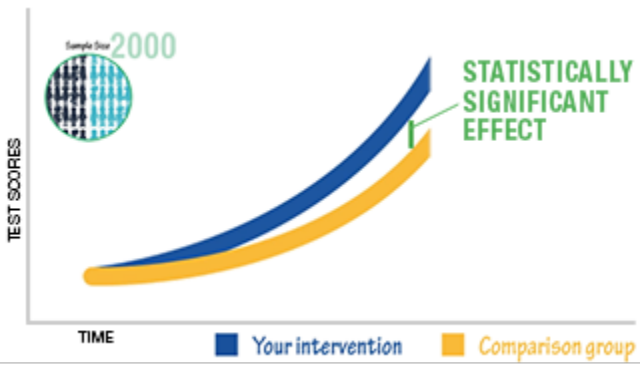
To report an effect, evaluators like to be at least 95% confident that the results they see aren't the result of random variation.

Think of larger sample sizes as a way to sharpen your vision. Look at the two line graphs below – A and B. Assume you conducted a rigorous study on your new program designed to improve teacher effectiveness. One important outcome measure you are looking at is how students' test scores improve because of the program.

Graph A: Study with sample size of 200



Graph B: Same study, but with sample size 2000



The dark lines in both graphs indicate the average results of students in your program and students in the comparison group. The only difference between the two graphs is the 'shadow' or 'fuzziness' around the two lines. In Graph A, the fuzziness overlaps – so you can't say with confidence that there actually is a difference. In Graph B, even though some fuzziness exists, you see that there is space between the two lines, indicating a statistically significant difference.

The only difference between the two graphs is the sample size. If you put a lot of time and energy in conducting an evaluation but used too small a sample size for your study, then your results will report "No effect." If you, however, started with a larger sample size, your results would show a "statistically significant" effect. Evaluators call the fuzziness around the results the confidence interval: evaluators are 95% confident that the true effect lies within the area covered by the confidence interval.

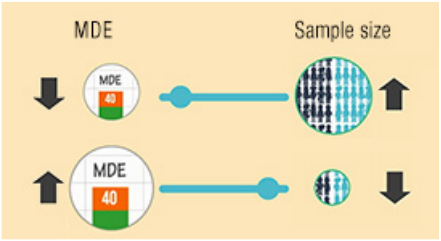
So, how can you determine how large your sample size needs to be to detect a difference?

This publication was produced by **IMPAQ International, LLC.**, under U.S. Department of Education Contract No. ED-ESE-15-A-0009/0001/0002.

U.S. Department of Education
http://www.ed.gov/

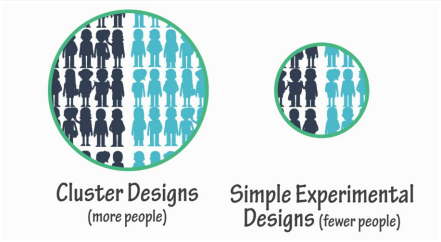
Evaluation design decisions that affect the sample size you need

To determine the sample size that you need to detect an effect, evaluators conduct a "power analysis." The power analysis considers several different variables. See below how decisions about some variables will increase or decrease the sample size you need.



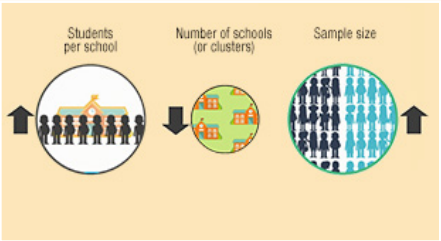
Minimum Detectable Effect (MDE)

The MDE is the smallest impact on your outcome of interest that your study is designed to find. As the MDE gets smaller, the sample size requirement is larger.



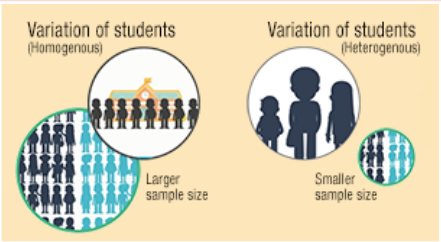
Simple or cluster evaluation design

Simple evaluation designs occur when students are assigned directly to a program or comparison group. Cluster designs are when students' classrooms or schools are assigned to the program or comparison group. Simple experimental designs require smaller sample sizes than cluster evaluation designs.



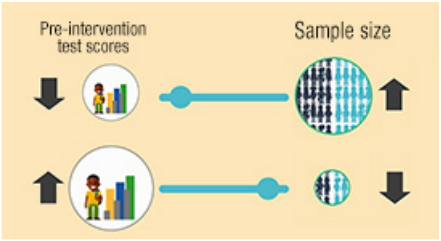
Number of individuals within clusters

The number of clusters and the number of students within clusters influences sample size. The more students that are in a cluster, the larger the overall sample size, but the fewer clusters that are needed.



Variation of students within cluster

If students within a cluster are heterogenous and represent many different kinds of students, then you can get by with a smaller sample size. If the students within clusters are fairly homogenous, then you will need a larger sample size.



Baseline data on outcome of interest

Having data from before the program begins that explains variation in the outcome of interest (for example, test scores) will reduce the sample size required. The more correlated the pre-intervention data is with the outcome, the smaller the sample size.

This publication was produced by **IMPAQ International, LLC.**, under U.S. Department of Education Contract No. ED-ESE-15-A-0009/0001/0002.