Ed Tech Rapid Cycle Evaluation Coach

Random Assignment Technical Appendix

This document provides a technical overview of how the RCE Coach Random Assignment tool functions. For a brief introduction to randomized pilots as a research design, refer to our Random Assignment Overview.

Random assignment attempts to create two groups that are as similar as possible in both observed and unobserved characteristics. You can think of the random assignment process as flipping a coin to assign students to either pilot the technology or not. Using a fair coin, each student would have a 50 percent chance of being assigned to the group piloting the technology. Therefore, if you have a sufficiently large pool of potential users, the two groups should have the same observed and unobserved characteristics. You can think of the Coach as a way to flip a virtual coin many times, very quickly. In addition, anyone can access the code to see that the coin is truly random, and the chance of being assigned to the treatment group can differ from 50 percent to provide a smaller, or larger, fraction of the population to initially test the technology.

WHAT DOES IT MEAN TO HAVE A SUCCESSFUL RANDOM ASSIGNMENT?

A successful random assignment process will produce treatment and comparison groups that are very similar. When the groups are large, random assignment is likely to succeed at creating two similar groups. For smaller samples, it is possible that the two groups differ in important ways even though a random process was used to create them. The Coach uses the baseline equivalence standard from the What Works Clearinghouse to assess whether the two groups are similar enough. We encourage you to use some pre-intervention measure that is related to the outcome to check for baseline balance, but this is not required. The Coach will consider that random assignment worked as expected if the differences on all the variables selected to control for balance are less than 0.25 standard deviations; otherwise, the Coach will attempt the random assignment process again. The Coach will assume that random assignment worked as expected if no variables are specified to check for baseline equivalence.

HOW DOES THE TOOL WORK?

The RCE Coach uses R, a free software package for statistical computing. In particular, the Coach uses the base R function sample.int to perform the random assignment. The Coach will first perform several checks on the data to confirm that there are no data issues that would cause an error with the random assignment code. Then the Coach will randomly assign observations to the intervention or comparison group and if baseline variables were specified, will check that random assignment worked as expected. The rest of this document describes the specific data checks and random assignment algorithm.

Ed Tech Rapid Cycle Evaluation Coach

Step 1: Check for data issues

The Coach will perform the following checks to verify that the Coach has the necessary inputs to perform random assignment:

1. Data are NULL, not readable, or have 0 observations.
2. Variable specifying the randomization unit ID is not specified or does not exist.
3. Whether groups should be determined by either a set number or percentage is not indicated.
4. Percentage was selected for assignment and the number of users indicated is less than or equal to 0 or greater than or equal to 100.
5. Number was selected for assignment and the number of users indicated is less than or equal to 0 or greater than or equal to the number of records in the data.
6. Baseline variables are specified and not all exist in the data.
7. Baseline variables are specified and one or more are not numeric.
8. A variable named Treatment already exists in the data.

If any test fails, the Coach will print a message to help the user identify the problem. If the Coach detects no problems, it will proceed to Step 2.

Step 2: Randomize and check for balance

The Coach will perform the random assignment using a loop that stops either when random assignment is successful or after 10 failed attempts. Random assignment is deemed successful when all variables selected for the baseline equivalence checks are balanced across the two groups. If the process was successful, the user will be able to download the data and move to the next step in the Coach. Otherwise, the coach will randomize again with a new random assignment seed. If the Coach fails to produce a successful random assignment 10 times, it will inform the user that the process did not work as expected. The following steps describe the loop:

1. The Coach randomly selects a random assignment seed.
2. Using sample.int a number of observations are chosen to be in the treatment group. This is either a fixed proportion of the sample (rounded up to the nearest whole observation) or a fixed number of observations.
3. The Coach checks if random assignment was successful.
4. If the process is successful a file will be available to download with the original data plus one variable indicating whether the observation belongs to the treatment or comparison group.
5. If the process is not successful
    a. If the loop has run fewer than 10 times, a new seed will be chosen and the loop will run again.

b. If the loop has run 10 times, a message will explain to the user that the random assignment was not successful.

The code for the RCE Coach is open source under the General Public License Version 3 and will be available soon on our github repository.