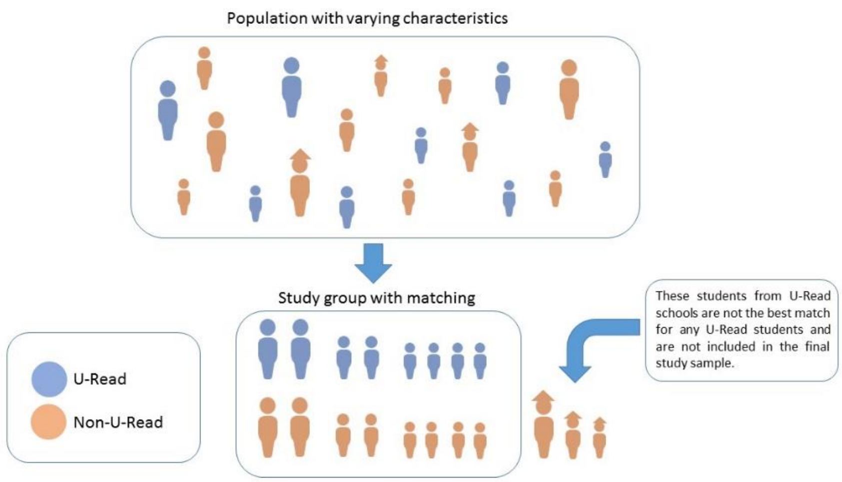
Matching Technical Appendix

This document provides a technical overview of how the RCE Coach Matching Dashboard functions. For a brief introduction to matching as a research design you can refer to our Matching Overview.

Matching attempts to group students (or the individuals you are trying to impact) with similar characteristics. The ideal match would be two students identical on every observable and unobservable characteristic. In practice, it is very unlikely that two students will be identical on all observed characteristics, so you must determine which observable characteristics you will match on based on how likely those characteristics are to impact the outcome of interest.

Matching on only one variable is conceptually simple. Suppose that you have a set of students using i-Ready, an educational technology, (blue) and another not using it (orange). You are trying to create a matched data set using only a pre-test score. A simple algorithm would look like this:



1. Define a distance measure. For example, the absolute distance between pretest scores of 240 and 265 is: |240-265|= 25
2. Take a student using U-Read and calculate the distance between his pre-test and the pre-test of each of the students not using U-Read.
3. Match this student to the student with the minimum distance.
4. Repeat steps 2-4 for all students using U-Read.

Notice that with this algorithm the order in which you do the matching matters. Moreover, defining distance when you are trying to match students using more than 1 characteristic is not as simple. Once you are using multiple matching variables, you have to summarize this information into a single distance number. For example, imagine that you want to include an indicator for English as a second language, and one for free or reduced price lunch eligibility. In this case, we can create a measure that summarizes all this information and match based on the summary. We can combine all of this information, even though the scales differ, by calculating the probability that a student with a given set of observable characteristics uses the educational technology. The goal is that after doing matching, students that look the same should have the same probability of using the educational technology.

HOW DOES THE MATCHING DASHBOARD WORK?

To create a matched data set, the RCE Coach uses the R package MatchIt. Specifically, we use one of the most common, and easy to understand methods, 1-to-1 nearest neighbor matching (Rubin, 1973). For each individual using the educational technology, the Coach will select a control individual with the smallest distance.

Once the Coach has selected matches for each individual using the educational technology, it will assess whether the two groups – the intervention group and the matched comparison group – are similar on these key characteristics. The Coach uses the baseline equivalence standard from the What Works Clearinghouse to make this assessment. If the baseline differences on all matching variables are less than 0.25 standard deviations, we consider the matched group to be a valid comparison group.

If the Coach fails to generate a valid comparison group, it will define a maximum distance known as a caliper and repeat the matching algorithm. If the distance between the individual using the technology and the nearest individual not using it is larger than this caliper, the Coach will drop this observation. If the Coach cannot generate a valid comparison group with this caliper, it will use a smaller one. The RCE Coach will start by using no caliper, and then move from 1 standard deviation of the distance measure to 0.25 standard deviations in 0.25 intervals. Using a caliper can mean that some users of the educational technology will not receive a match. If this is the case, one has to be careful when interpreting the results because you are only measuring the effect for the individuals who have a good match.

WHICH VARIABLES SHOULD BE INCLUDED WHEN USING THE MATCHING DASHBOARD?

You should always include some baseline, or pre-intervention, measurement that is related to your outcome of interest. In the case of U-Read, you would include pre-test scores as one of your matching variables. However, there are cases in which you may want to use additional characteristics for matching in order to improve precision.

In our example, if you have substantially more students not using U-Read than using U-Read, the literature recommends including:

- ☐ Variables that affect the probability of using U-Read
- ☐ Variables that affect the outcome you want to study (e.g., achievement), controlling for use of the educational technology

(Rubin and Thomas 1996; Heckman et al. 1998; Ho et al. 2007b)

If you do not have substantially more students not using U-Read than the number of students using U-Read, you **should not** include all available characteristics (Ho et. al. 2007b). The intuition is that, although it would increase your predictive power for the probability that a student is using the technology, as you increase the number of characteristics used for matching you will decrease the likelihood of finding a good match.

SHOULD MATCHING VARIABLES ALSO BE INCLUDED AS COVARIATES IN THE ANALYSIS?

YES! Matching is seldom perfect – some imbalance often remains on at least some variables. Regression adjustment for the matching variables can mitigate any remaining imbalance.

R-CODE

The Coach uses the following R code:

```
match_result <- matchit(
            as.formula(strformula),
            data = match_data[[grade]],
            method = 'optimal',
            ratio = 1)
```

Where `` `strformula` `` is created based on the inputs you selected and `` `match_data` `` is the data you uploaded. All the code for the RCE Coach is open-source under the GPL-V3 license, and available on our github repository: https://github.com/mathematica-mpr/MPRDashboards

CITATIONS:

Daniel Ho; Kosuke Imai; Gary King; and Elizabeth Stuart (2007), 'Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference,' Political Analysis 15(3): 199-236, http://gking.harvard.edu/files/abs/matchp-abs.shtml

Daniel Ho; Kosuke Imai; Gary King; and Elizabeth Stuart (2007b) 'Matchit: Nonparametric Preprocessing for Parametric Causal Inference,' Journal of Statistical Software, http://gking.harvard.edu/matchit

Hansen, B.B. and Klopfer, S.O. (2006) Optimal full matching and related designs via network flows, JCGS 15 609-627, http://dept.stat.lsa.umich.edu/~bbh/hansenKlopfer2006.pdf

Heckman, J., Ichimura, H., Smith, J., & Todd, P. (1998). *Characterizing selection bias using experimental data* (No. w6699). National bureau of economic research.

Rubin, D. B., & Thomas, N. (1996). Matching using estimated propensity scores: relating theory to practice. *Biometrics*, 249-264.