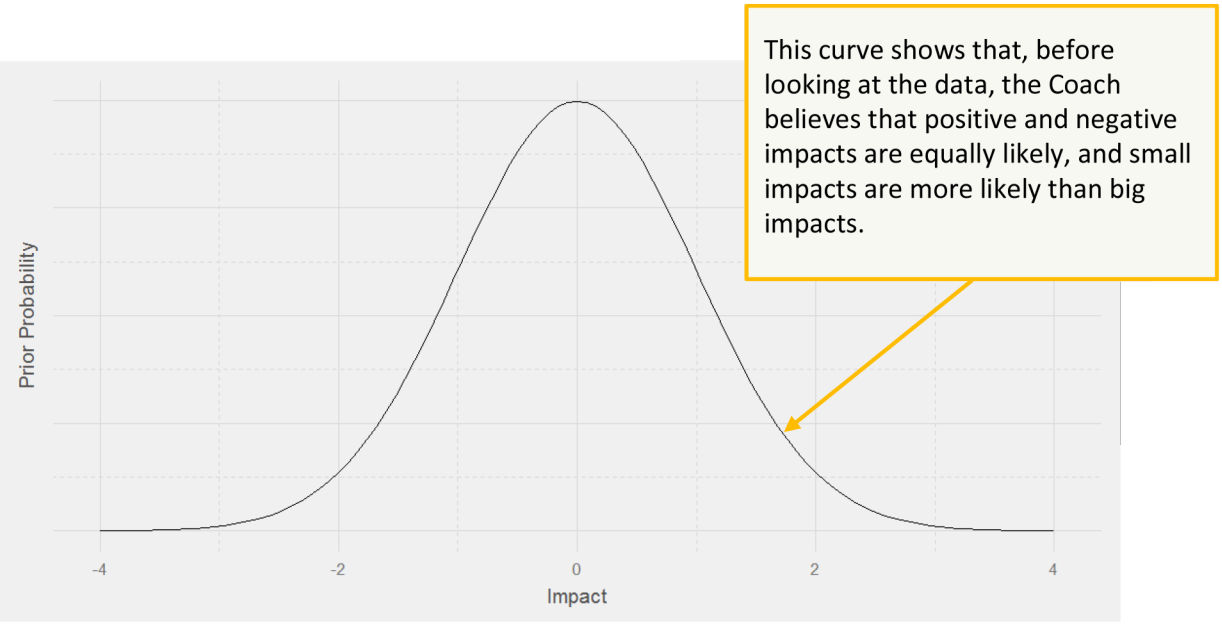
Ed Tech Rapid Cycle Evaluation Coach

Impact Estimation Technical Appendix

OVERVIEW

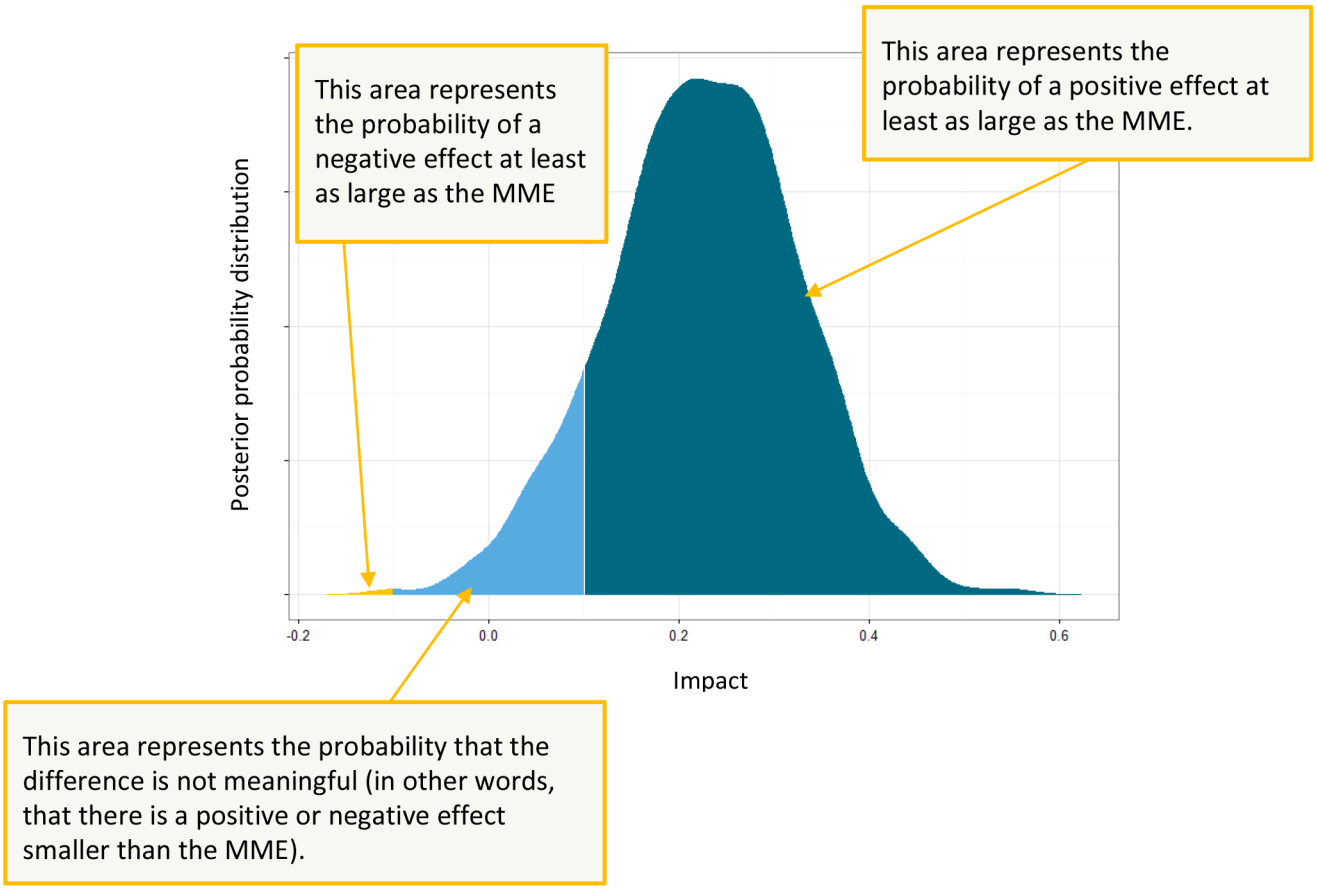
The RCE Coach will help you determine if an educational technology is moving the needle. Before you load your data, the Coach believes that negative and positive effects are equally likely, and that small impacts are more likely than big impacts. This is what statisticians call the prior distribution and it is depicted in Exhibit 1. A prior distribution describes the analyst's beliefs about a measurable outcome before any data are taken into account.

Exhibit 1. Prior distribution of the effect of the educational technology on the outcome



When the data are loaded and the other inputs are specified, the RCE Coach calculates the probability that the effect of an educational technology exceeds the minimum meaningful effect (MME) threshold, in either direction, or whether it falls into the region of practical equivalence (the area between –|MME| and |MME|). Exhibit 2 illustrates this.

Exhibit 2. Probability of an educational technology's effects, given the data uploaded



The Findings Brief for each evaluation includes this information.

In this example, the user told the RCE Coach that any increase in the outcome larger than 0.1 units would be considered meaningful. The Coach calculated that there is an 88 percent probability that the educational technology increased outcomes by at least that amount.

Note that the answer to whether the educational technology is moving the needle will depend on the threshold chosen by the user (0.1 units in this example) and how much uncertainty he or she is willing to tolerate. For example, some users might conclude that a technology is moving the needle if there is at least a 75 percent probability that the technology meets the threshold they chose, whereas others might want to be more confident, leading them to select a higher probability. In the previous example, if the user sets a certainty level of 95 percent, then this technology would not be found to be moving the needle because 95 is larger than 88.

THE RCE COACH

The RCE Coach uses a program called Stan (see Carpenter et al., 2016) to estimate a linear model using Bayesian statistics. In particular, the Coach uses the R package RStan. The Coach acts as a simple graphic interface to help the user construct a formula. The user has to select the outcome, treatment indicator, and any additional covariates that the user might want to include. Before estimating the effect of the intervention, the Coach temporarily standardizes the data using the R function "scale." Then, the Coach estimates the following model:

$$y_i = \alpha + \eta T + \beta x_i + \varepsilon_i$$
$$\varepsilon_i \sim N(0, \sigma)$$
$$\sigma \sim N(0,1)$$
$$\beta_k \sim N(0,1)$$
$$\eta \sim N(0,1),$$

where y_i is the outcome of interest for individual i, *T* is the treatment indicator, and x_i is a matrix with the other covariates.[1] The term η tells you the effect of the educational technology—that is, how much higher the outcome was for the average student using the educational technology, *T* = 1, compared with the average student not using it, *T* = 0. The Coach computes the posterior distribution of η, which describes our understanding of the educational technology's effect after observing the data. Using this distribution, the Coach can calculate the probability that the treatment effect (1) positively exceeds the MME, (2) negatively exceeds the MME, and (3) does not exceed the MME (falls into the region of practical equivalence), as shown in Exhibit 2.

WHICH VARIABLES SHOULD I INCLUDE AS CONTROL VARIABLES?

You should include all the variables, measured before you implement the technology, that you think can affect the outcome of interest and for which you have good data. In the case of student achievement, it's common to include previous achievement and other characteristics, such as indicators for English as a second language, socioeconomic status, and so on.

In addition, sometimes you might have to include an element to the model to account for clusters in the data, such as classrooms or schools. This applies when students are assigned to use or not use the technology as groups, rather than individually. To correctly estimate the effect of the intervention, you will have to tell the Coach that you used clusters during Step 5. By giving the Coach this information, you enable it to account for the possibility that students belonging to a cluster can do better or worse just because they belong to that cluster and not because they are using the

[1] In the clustered case, we add cluster-specific random effects to the regression model that's given.

technology. For example, when entire classrooms are assigned to conditions, the quality of teachers in each classroom can affect outcomes, in addition to any effect of the technology. In this case, you would want to indicate the variable that identifies classroom clusters so that the Coach can account for classroom-based differences in outcomes.

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

CITATIONS:

Carpenter, Bob, Andrew Gelman, Matt Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Michael A. Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software*, in press.

Stan Development Team. "RStan: the R Interface to Stan, Version 2.10.1." Available at http://mc-stan.org.