The Rapid Cycle Evaluation Coach: Building capacity and informing decisions

Presentation at the 2018 OPRE Methods Meeting
Washington, DC

October 25, 2018

Kate Place
Agenda

• Background
• RCE Coach overview
• Research questions
• Methods
• Case study
• Lessons learned
Background
Background: The initial motivation

• Billions of dollars invested in educational technology
  – Little evidence of what works

• Schools and districts need timely and useful information to make decisions
  – Traditional evaluations: too long and too expensive
Rapid Cycle Evaluation (RCE)

Rapid
Rapid identification of results

Cycle
Continuous improvement model

Evaluation
Rigorous experimental techniques
RCE Coach overview
EdtechRCE.org

• Free, open-sourced, online toolkit
• Facilitates the use of rigorous evaluation to inform decisions
• Designed for users with little time, limited knowledge of statistics
• Funded by the Office of Educational Technology of the US Department of Education (2015-2018)
The RCE Coach steps

1. GETTING STARTED
   The Coach will recommend an approach to evaluate your technology.

2. PLANNING YOUR RESEARCH
   The Coach will help you design an evaluation based on the outcomes you are interested in and your unique context.

3. PREPARING YOUR DATA
   The Coach will use your data to create two groups—a technology user group and a similar comparison group.

4. ANALYZING YOUR DATA
   The Coach will automatically conduct the analysis and give you the results.

5. SUMMARIZING YOUR FINDINGS
   The Coach will compile your results and all of the information you have entered into one succinct document or presentation.
Research questions
Research questions:
Broad questions that districts and schools ask

• Is this program or product working?
• Is it worth expanding this pilot program to all students?
• What is the best way to implement this intervention?
• Which of two programs or products works better?
Research questions:
Specific questions that guide their RCEs

- Narrow, specific research questions

- **Craft Your Research Question** is a tool in the Coach
  - The blueprint: Does A do B among C compared to D?
  - **Example:** Does eZumi Learning increase student achievement on the benchmark English Language Arts assessment among 5th-grade English learners compared to similar students with no access to eZumi Learning?
The methods
The methods: Behind the scenes statistics

• Supports random assignment and matched comparison designs
• Applies Bayesian methods to determine the effect of the intervention
• Uses R to do statistical analysis
The methods: Randomized pilots

Beginning of the evaluation

Intervention is introduced

Only one group has access to the intervention

Treatment group Receives the intervention

Comparison group Continues with business as usual

Compare outcomes
The methods: Matched comparison designs

Potential users

- Some receive intervention, but others do not

Beginning of the evaluation

Intervention is introduced

Unmatched
- Not included in our final sample

We match similar students

Unmatched
- Not included in our final sample

Treatment group
Receives the intervention

Comparison group
Continues with business as usual

Compare outcomes
The methods: Benefits of Bayesian

• Requires decision-making prior to evaluation & facilitates discussion
  – What does success look like? (minimal meaningful effect)
  – How confident do we need to be in the results? (certainty threshold)

• Provides more intuitive results for non-researchers
  – Probability rather than p-values (e.g. there is an 80% chance the intervention has a positive effect on student achievement)
  – Less likely to be misinterpreted (Chandler et al., 2017)

• Is more aligned with questions districts and schools want to answer
The methods: Presentation of Bayesian results

The treatment group does worse
The two groups are equivalent
The treatment group does better

Probability (percent)

75% certainty threshold
### The methods: Frequentist vs. Bayesian results

<table>
<thead>
<tr>
<th>Grade</th>
<th>Outcome</th>
<th>Impact estimate (standard error)</th>
<th>Effect size</th>
<th>p-value</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>MAP reading comprehension test</td>
<td>1.72 (1.33)</td>
<td>0.11</td>
<td>0.2</td>
<td>[-0.91, 4.35]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade</th>
<th>Meets goal?</th>
<th>Possible change in outcomes (treatment vs. comparison)</th>
<th>Probability of change</th>
<th>Estimated size of effect</th>
<th>Comparison group average</th>
<th>Number of students in treatment group</th>
<th>Number of students in comparison group</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>Not sure</td>
<td>↑ Increase &gt;=1</td>
<td>70%</td>
<td>1.7</td>
<td>217</td>
<td>58</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Between -1 and 1</td>
<td>28%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>↓ Decrease &gt;=1</td>
<td>2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Case study: How does a district use the Coach?
Uplift Education Charter Schools

**Uplift’s schools**

- Charter network of 37 schools
- Pre-K to 12th grade
- Data team of 11 full-time staff support schools
- Dallas/Forth Worth, Texas

**Uplift’s students**

- About 19,000 students
- 74% economically disadvantaged
- Mostly minorities

![Pie chart showing the racial breakdown of Uplift's students: 72% Hispanic or Latino, 22% Black, 5% Asian, and 1% White.](chart.png)
“When you have a 2nd grader who can’t read, you don’t have the luxury of time.”
–Kim Lammers, Uplift Education
Evaluation details

• **Research question:**
  – Did the reading intervention increase reading achievement among 2nd grade students compared to similar 2nd grade students with no access to the intervention?

• **Sample:**
  – 54 treatment, 54 comparison students

• **Evaluation method**
  – Backward-looking matched comparison design

• **Results**
  – 0% probability that reading scores increased by at least 7 points
  – 12% probability of any impact on reading scores

THE READING PROGRAM

Students at least 6 months behind in reading assigned a volunteer tutor

Students are pulled out of class and work with tutor for 45 minutes, twice per week, following a structured curriculum

Reading scores improve by at least seven points on MAP reading assessment

Only a 12 percent probability of any positive impact on MAP reading scores

INTENDED EFFECT

ACTUAL EFFECT
Using evidence for decision-making

- Data team and Uplift leaders reviewed the evaluation results

- Results given to schools
  - 2 schools discontinued the program and hired reading specialists
  - 1 school kept the program to look at another year of data

- Uplift conducted several other evaluations

**EVIDENCE-BASED DECISION-MAKING PROCESS**

1. Examine reading program evaluation results
   - Not having the intended impact
   - Discontinue the reading program

2. Reallocate $40,000 per year, per school to the hiring of additional reading specialists to help struggling students
Lessons learned
Lessons learned: Use of the Coach

• Toolkit developed to evaluate educational technology, but served broader need

• Districts and schools want evidence specific to their context

• Need to meet schools and districts where there are
  – Few are able/willing to conduct randomized controlled trials
  – Coach can build capacity, change mindsets about program evaluation
  – Process is as important as the product
Lessons learned: Sustainability

• Less sustainable in:
  – Very large districts—internal evaluation capacity, external research partners
  – Very small districts or schools—limited staff time and capacity for independent research, conflicting priorities

• More sustainable in:
  – Mid-sized districts with data teams, champions
References

• Chandler, J., Finucane, M., Martinez, I., Resch, A., & Terziev, J. (2017). Speaking on the data's behalf: Different presentations of the same data lead to different decisions. Paper presented at Association for Public Policy Analysis & Management Conference, Chicago, IL.
For more information

• Kate Place
  – KPlace@mathematica-mpr.com