# Estimating Long-term Program Impacts When the Control Group Joins Treatment in the Short-term

Andrew Jaciw, Xiaohui Zheng, Boya Ma, Qingfeng Zhao
Empirical Education Inc.

## Background.

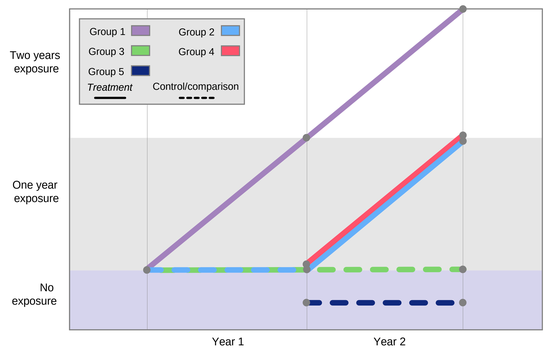
Randomized trials of educational interventions often face logistical hurdles (Cook, 2002). For example, sometimes, as a condition for participation, a control is allowed to receive the treatment program prior to the full-term required to observe a mature treatment effect. Other methods are then required to infer the full-term impact. We compare two such approaches:

(1) Quasi-experimental: (Shadish, Cook and Campbell, 2002) We use a matched comparison group to infer how the control would have performed had they not joined treatment.

(2) Extra-experimental: (Bell and Bradley, 2008) We contrast performance between the initially randomized groups, before the control joins treatment and at the full-term, and use both contrasts to construct an estimate of how the controls would have performed had they not joined the program.

We analyze data using both methods to see if we can obtain convergent validity.

## Research Design.




| Student Group | Year 1 Treatment | Year 1 Control | Year 2 Treatment | Year 2 Control | Year 2 Comparison |
|---|---|---|---|---|---|
| **Group 1** | X | | X | | |
| **Group 2** | | X | X | | |
| **Group 3** | | X | | X | |
| **Group 4** | - | - | X | | |
| **Group 5** | - | - | | | X |

*Note*. "-" indicates that the student group was not involved in Year 1.

## Student Groups.

**Five student groups are involved in the study:**

Group 1 — Randomized to treatment in Year 1; stays in treatment in Year 2

Group 2 — Randomized to control in Year 1; phases into treatment in Year 2

Group 3 — Randomized to control in Year 1; stays in control in Year 2 (Logistical issues prevent uptake into treatment.)

Group 4 — Enters study in Year 2 and receives treatment.

Group 5 — Enters the study in Year 2 and does not receive treatment.

## Empirical Study.

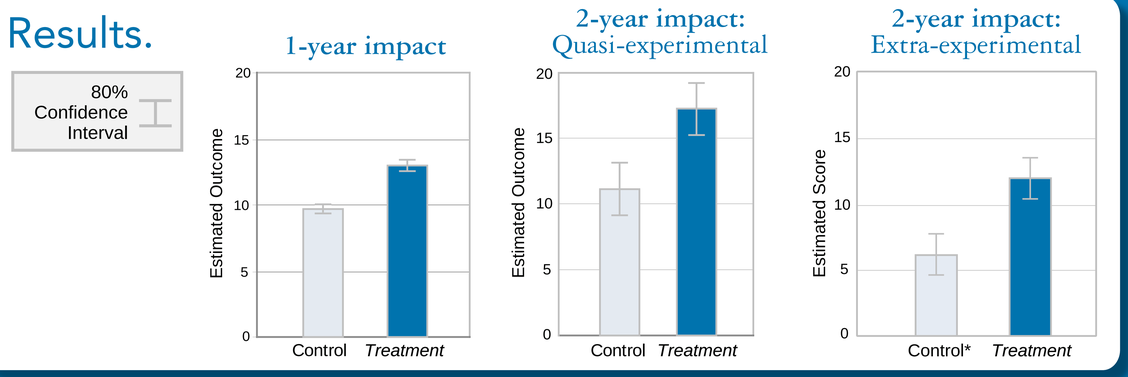
**Data collection**

- Data are collected from a longitudinal study of a Special Education reading program implemented in two Florida school districts.
- The program is designed for students with severe learning and developmental disabilities from grades 3 through 8.
- The primary research question is whether students who receive the program, have stronger sight word recognition skills than those who don't, after one and two years.

## Methods.

1. **Experimental (1-year impact).** We compare performance of randomized groups after 1 year.
2. **Quasi-experimental (2-year impact).** Students in Group 1 are matched to similar students in groups 3 and 5 and we compare the difference between them in performance.
   **Extra-experimental (2-year impact).** We estimate the difference in performance between the randomized groups (groups 1 and 2) after the first and second years, and combine the differences to construct an estimate of how the controls would have performed had they not joined the program.

## Results.




## Conclusion/Discussion.

| Impact | Method | Estimate | Effect size | *p* value | Percentile standing | Bias due to selection | Bias due *to* a changing treatment | Bias due to attrition[c] |
|---|---|---|---|---|---|---|---|---|
| **1-year** | Experimental | 3.17 | 0.55 | <.01 | 21% | Ruled out | Ruled out | Possible |
| **2-year** | Quasi-experimental | 6.12 | 0.89 | .06 | 31% | Possible[a] | Ruled out | Possible |
| **2-year** | Extra-experimental | 5.81 | 0.98 | .02 | 34% | Ruled out | Unlikely[b] | Possible |

[a] Balance is achieved on a series of covariates.
[b] No evidence that the treatment changed between the first and second year (for a description of the assumptions underlying the extra-experimental method see Bell and Bradley (2008))
[c] Remaining groups pass balance checks.

- We compared two methods for obtaining two-year impact estimates in the situation where the experimental control joined treatment after one year. Each method has strengths and limitations that the other does not.
- In this study, the two-year impact estimates from both methods are consistent, which gives us convergent validity and greater confidence in the result.

References:
Bell, S. H., & Bradley, M. C. (2008, March). *Calculating long-run impacts in RCTs that release the control group into the intervention prior to the end of follow-up*. Paper presented in a paper discussion at the Annual Research Conference of the Society for Research on Educational Effectiveness, Crystal City, VA.
Cook, T. D., (2002). Randomized experiments in educational policy research: A critical examination of the reasons the education evaluation community has offered for not doing them, *Educational Evaluation and Policy Analysis*, 24, 175-199.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.