mdrc
BUILDING KNOWLEDGE TO IMPROVE SOCIAL POLICY

REFLECTIONS ON METHODOLOGY

MAY 2018

Two-Stage, Multiyear Random Assignment at Different Levels

By Shira K. Mattera and Marie-Andrée Somers

This post is one in a series highlighting MDRC's methodological work. Contributors discuss the refinement and practical use of research methods being employed across our organization.

A two-stage, multilevel approach to random assignment is an intriguing way to test a complex set of interventions — such as interventions children experience in sequence as they move from preschool to kindergarten. In the first stage of such a design, groups (schools or centers) of participants are randomly assigned to program and control groups. In the second stage, individuals within those groups (students or children) are randomized to two different conditions. This design creates up to four groups — a group that receives both interventions, a group that receives the first intervention but not the second, a group that receives the second intervention but not the first, and a control group. By comparing the outcomes of participants in these groups, the combined and relative effects of the two interventions can be tested.

MDRC's Making Pre-K Count and High 5s studies are a recent application of a phased, two-level, two-stage design, used to rigorously test whether providing high-quality math instruction that is aligned across prekindergarten (pre-K) and kindergarten could provide children with a critical boost that would lead to long-term achievement gains across a variety of domains. The studies are the result of a partnership between Robin Hood, one of the country's leading antipoverty organizations based in New York City, and MDRC. Additional funding was provided by the Heising-Simons Foundation, the Overdeck Family Foundation, and the Richard W. Goldman Family Foundation.

MAKING PRE-K COUNT AND HIGH 5S

The original purpose of the Making Pre-K Count study was to widely test a preschool intervention that could dramatically affect low-income children's long-term outcomes. At the preschool level, Making Pre-K Count comprised a math curriculum (Building Blocks) supported by teacher training and in-classroom coaching. Building Blocks had a substantial evidence base and was a promising candidate for a large-scale test.

Findings from the field of early childhood intervention, however, suggested that it was possible that impacts from a pre-K program would fade. Therefore, the University of Michigan and MDRC developed a secondary, more exploratory kindergarten program, High 5s. High 5s math clubs, delivered by facilitators three times per week, were designed to provide a small-group math enrichment experience aligned with both the content and approach of the Building Blocks curriculum for kindergarten students who had experienced Building Blocks in pre-K. Together, the two programs were intended to improve and sustain children's math skills across pre-K and kindergarten.

By using a two-stage design, we could preserve the original test of the pre-K math curriculum while adding a test of the kindergarten component and a test of their combined effect.

STUDY DESIGN

In the Making Pre-K Count study, pre-K programs (schools) were randomly assigned either to implement Building Blocks, with associated professional development to support teacher implemen-

NEW YORK
16 East 34th Street
New York, NY 10016
Tel: 212 532 3200

OAKLAND
475 14th Street
Suite 750
Oakland, CA 94612
Tel: 510 663 6372

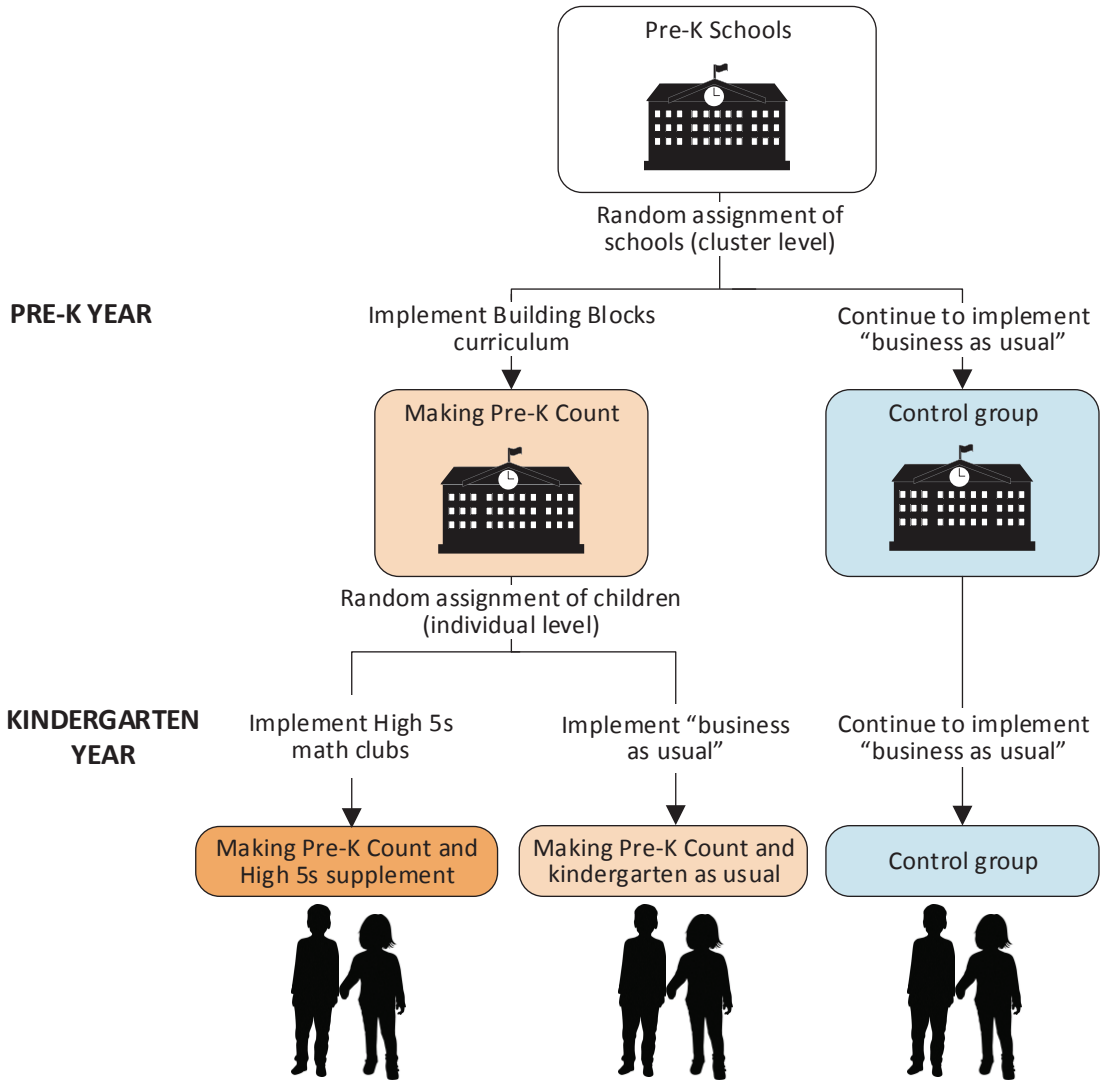
WASHINGTON, DC
1990 M Street, NW
Suite 340
Washington, DC 20036

LOS ANGELES
11965 Venice Boulevard
Suite 402
Los Angeles, CA 90066

www.mdrc.org

tation of the curriculum, or to a pre-K-as-usual control condition. In the High 5s study, kindergarten students who had been in Making Pre-K Count program classrooms in pre-K were then individually randomly assigned within schools to High 5s small-group supplemental math clubs or to a business-as-usual kindergarten experience. A notable feature of this two-stage design, illustrated in the figure below, is that the level of random assignment differs across stages: In the first year, a cluster-level design was used to randomly assign *the schools*; in the second year, individual-level random assignment was used to assign *students* from the first-stage treatment group.

Multiyear, Multilevel Random Assignment Design of Making Pre-K Count and High 5s



This design created three research groups at the end of kindergarten: (1) children with two years of math enrichment (Making Pre-K Count in pre-K and High 5s in kindergarten), (2) children with one year of math enrichment (Making Pre-K Count in pre-K only), and (3) children with no math enrichment (a control group with pre-K and kindergarten as usual).

THE BENEFITS OF A TWO-STAGE MULTI-ARM DESIGN

The two-stage, multi-arm random assignment research design made it possible not only to learn whether the specific interventions worked, but also to build evidence about the combined effect of two well-aligned educational programs over time.

The study design is also resource-efficient. We used the cluster-level design in testing the pre-K curriculum to minimize program spillover from one classroom to another. But randomizing enough clusters (schools) to create three arms for testing the pre-K and kindergarten programs would have been challenging in its operational demands and use of resources. Randomly assigning children to kindergarten math clubs within the program arm allowed for reasonable statistical power while balancing the reality of implementing the clubs in so many schools.

Moreover, the design minimized the loss of power from attrition of children across years. As children moved from pre-K to kindergarten, randomizing again in the kindergarten year meant that only children who stayed in the sample for kindergarten were randomly assigned, minimizing attrition and maximizing power.

In sum, the two-stage, multiyear, multilevel study design provided a test of two interventions — impact findings were published in March 2018 — as well as rigorous and policy-relevant evidence about ways to sustain the effects of preschool interventions over time, while still making efficient use of evaluation resources.