

AFTERSCHOOL RESEARCH BRIEF

A publication of SEDL's Afterschool Research Consortium

IMPLEMENTING RANDOMIZED CONTROLLED TRIAL STUDIES IN AFTERSCHOOL SETTINGS: THE STATE OF THE FIELD

MICHAEL VADEN-KIERNAN, PhD, DEBRA HUGHES JONES, PhD, AND ZENA RUDO, PhD

Introduction

SEDL is providing analytic and technical support to three large-scale randomized controlled trials assessing the efficacy of promising literacy curriculum in afterschool settings on student academic achievement. In the field of educational research, competition among research organizations and researchers can often impede collaborative efforts in the field. One important aspect of the current work has been the creation of the SEDL Afterschool Research Consortium (ARC), which brings together SEDL researchers, key staff from each afterschool research project, and experts in the field to discuss and share accomplishments, challenges, and solutions. This unique arrangement allows a number of researchers with varied backgrounds and experience implementing rigorous research in afterschool settings to collaborate in applying best research practices in the field, to a set of ongoing studies to improve the current studies' efforts and address important challenges to the field of afterschool research. The ARC also plans to develop papers and presentations documenting the lessons learned in order to advance the effective use of rigorous experimental research approaches in applied afterschool settings. This research brief is the first in a series of papers that are intended to provide the field with the insights culled from this collaborative effort.

To date, the empirical evidence concerning the impact of afterschool programs on student academic achievement outcomes have been mixed. For instance, the national

evaluation of the 21st Century Community Learning Centers program found no statistically significant effects on academic achievement outcomes in reading or math for students participating in these programs (Dynarski, James-Burdumy, Moore, Rosenberg, Deke, & Mansfield, 2004), while more recently other reviews and meta-analyses have shown some positive impacts of afterschool programs on academic achievement outcomes (e.g., Lauer, Akiba, Wilkerson, Apthorp, Snow, & Martin-Glenn, 2004; Harvard Family Research Project, 2006). An important empirical and policy-relevant question therefore remains: Can fully developed, well-implemented afterschool programs impact student academic achievement outcomes? Many stakeholders—from the policy, practice, and research communities—increasingly need this information to inform decisions regarding program planning, implementation, and effectiveness.

To address this question, in 2006 the U.S. Department of Education (DOE) funded three randomized controlled trials (RCTs) that use rigorous methods to evaluate the benefits of promising afterschool interventions on student achievement in order to provide evidence that informs the national debate. SEDL, as part of the National Partnership for Quality Afterschool Learning (the National Partnership), coordinated the competition and is also leading a research consortium to support the afterschool RCT grantees' efforts over the 2-year funding period (2006–2008). The purpose of this research brief is to summarize the challenges that have been raised by the proposal review process and the

implementation of these studies during the first year of the project. This research brief seeks to highlight challenges to the field of afterschool research as the need for rigorous approaches to addressing questions of program impact and efficacy in applied settings increases.

Goals of the Request for Proposals

SEDL issued a request for proposals (RFP) in March 2006 that sought researchers to conduct rigorous evaluations that examine the efficacy and impact of promising afterschool interventions. The research questions addressed in the evaluations were to be well defined and clearly pertinent to the extant theoretical and empirical literature in the field. Proposed research projects were to involve efficacy trials using experimental or randomized controlled trials study designs to rigorously test a promising intervention. The purpose of the efficacy trials were to produce unbiased estimates of the degree to which a research-based, well-defined, and well-implemented intervention has a net positive impact on the student academic achievement in relation to the program or practice in which it is being compared. Promising afterschool interventions were to include reading and/or math components and target elementary-age students (grades 2–5) with a focus on improving student academic outcomes. We detail the criteria that were used to evaluate the proposals below.

Theoretical and Empirical Rationale. Research projects were to be theoretically grounded, in terms of both basic theory and program theory, relating program structures, processes, and practices to short- and long-term outcomes. Applicants were encouraged to review empirical evidence that supports the program theory and evaluate the empirical evidence to date on the effectiveness of the program in terms of student academic achievement outcomes. A key criterion for successful proposals would be their ability to demonstrate empirical support that the program or program components were promising and, therefore, merits inclusion in an efficacy trial or randomized controlled trial. A promising program or program component was defined as fully developed

and replicable and having demonstrated empirically documented gains on intended outcomes of interest. Preference was given to promising programs or program components that have one or more outcome studies (i.e., pretest/posttest design) documenting improvement in academic outcomes among participants. The theoretical and empirical rationale was intended to provide and support a coherent set of research questions that would guide the project.

Requirements of the Intervention and Implementation. Applicants were asked to propose an evaluation of a promising afterschool intervention that had been fully developed and implemented in an education setting, that was replicable, and for which a strong case could be made that the study of such an intervention would have important implications for practice and policy. The proposed afterschool intervention was to target programs or program components focused on reading or math with elementary-age students (grades 2–5). Projects were encouraged to focus on interventions that were designed to improve academic outcomes (e.g., achievement test scores). The intervention was to be clearly described, including the subject matter, the grade levels of the students to be targeted, the types of students to be affected, the setting in which it would be delivered, the duration, the intensity (attendance and hours or days per week), the numbers and qualifications of the teachers or other staff who would be involved, and the student outcomes that would be targeted.

The applicants were also encouraged to include a detailed plan for implementation of the intervention. This included identifying the school districts and schools or other education settings that had agreed to participate in the study and explaining, as completely as possible, how students, teachers, and/or classrooms were to be selected to participate in the proposed study. Additionally, the applicants were asked to ensure that participation of respondents would be maximized and to propose effective strategies to minimize attrition. Applicants were asked to indicate how the intervention would be maintained consistently across multiple settings

(i.e., classrooms and schools) over time. Particular attention was placed on how the proposed program implementation plan would maximize fidelity of implementation across settings and over time, such as providing ongoing technical assistance to programs or addressing any problems encountered.

Research Design. Given that the projects were asked to conduct efficacy trials, projects were encouraged to use experimental designs that randomly assign targets of the intervention (e.g., schools, teachers, or students) to treatment and control conditions to test the efficacy of promising afterschool interventions. There was a particular interest in the opportunity to conduct planned variation studies in which the control condition is not "no treatment" or "treatment as usual," but instead the design includes a contrast that isolates the impact of a particular program component (for example: an afterschool program and a reading component [treatment] versus the afterschool program only [control]; an afterschool program and a reading component versus another afterschool program; or an afterschool program and a reading component, the afterschool program and a math component versus the afterschool program control group). However, a planned variation study was not required nor was it advisable unless the contrast of program components was meaningful from the theoretical and empirical rationale and an operational and programmatic perspective of the participating schools.

The applicants were asked to clearly state the unit of randomization (e.g., student, classroom, teacher, or school) and the rationale for using that unit of randomization. Proposals should have explained the procedures for assignment of schools, classrooms, or participants to treatment and control conditions. Proposals should also have demonstrated how the researchers intend to assess the fidelity of the implementation of the intervention and strategies for avoiding contamination. A clear and complete description should have been provided for both the treatment and control conditions. Applicants were asked to demonstrate whether their research design included a sufficient number of settings and/or

participants to provide adequate statistical power for detecting meaningful effect sizes for improvements in academic achievement. Preference was given to proposals that could demonstrate a range of minimal detectable effect sizes supported in the literature and their impact on the study sample size as a justification for their decision regarding sample size.

Research Methods and Measures. Applicants were asked to demonstrate that the research methods and proposed measures were appropriate and justify the approach taken in terms of rigor. SEDL encouraged, but did not require that proposals take a mixed-method approach that integrates quantitative and qualitative strategies to address the research questions. All proposals, however, should have shown consideration of the best methods for addressing their research questions and the limits of any one method or data source. The proposals should have supplied information on the reliability, validity, and appropriateness of the proposed measures. Preference was given to proposals that included standardized measures of academic achievement in reading and math. Proposals could include observational, survey, or qualitative methodologies as a complement to experimental methodologies to assist in the identification of factors that may affect the implementation of the intervention and to provide clues as to how the intervention might be deployed more effectively and efficiently in the future. Proposals could also include measures of mediating or moderating variables for both the intervention and control conditions (e.g., student attendance or time-on-task, teacher quality). Successful proposals should have proposed a detailed data collection plan including how data collection would be managed and implemented throughout the course of the project. Applicants were also asked to include a data analysis plan that addressed how data would be analyzed to address each of the study's proposed research questions.

Personnel and Resources. Proposals were asked to demonstrate that their research teams collectively possessed the skills and experience to conduct the proposed study design. Preferred skills and experience

included design and implementation of randomized controlled trials, demonstrated substantive knowledge in afterschool and/or the subject area of the approach or intervention, expertise in statistical analysis, and experience working in education settings with districts, schools, teachers, and students. Partnerships between program personnel and research personnel were strongly encouraged. However, applicants were asked to ensure that the involvement of program personnel would not jeopardize the objectivity of the evaluation. Proposals were required to document the availability and cooperation of schools or other education settings that would be required to carry out the research proposed in the application via a letter of commitment. The letter was also to indicate an acceptance of the responsibilities associated with participating in the study, including agreement to provide a sufficient number of sites, schools, classrooms, and/or students to participate in the study and, in the case of random assignment, an agreement to random assignment of students, classrooms, or schools. The request for proposals indicated that SEDL anticipated making up to three 2-year awards for up to \$300,000 per year for each study.

Proposal Reviews: Challenges to the Field

SEDL received 13 proposals, which were reviewed by a five-person review committee. The review process was supported and facilitated by SEDL and also included representatives from the U.S. Department of Education. The following three studies received funding through the review process:

- *A Randomized Evaluation of the Adventure Island Afterschool Reading Curriculum With English Language Learners*, conducted by the Success for All Foundation. Adventure Island, an adaptation of the Success for All curricula, is currently being evaluated in a large-scale randomized experiment; however, that experiment includes few English language learners (ELLs). Because ELLs are a population of great interest, the new experiment will be conducted in 14 majority-Hispanic schools in Alabama, Texas, and Utah, with approximately 1,260 children in grades 2–4. The study will assess impacts on students in Adventure Island compared to students in the typical afterschool program offered in their school.
- *What Works in Afterschool Programs: The Impact of a Reading Intervention on Student Achievement in the Brockton Public Schools*, undertaken by MPR Associates, Inc., in partnership with Scholastic Publications and Brockton (Massachusetts) Public Schools. This study is designed to compare Scholastic's READ 180 curriculum, which has many characteristics that have been associated with positive academic outcomes, with Brockton's standard afterschool services. The study will include approximately 1,100 students in grades 4–6, to provide methodologically rigorous information about READ 180 and to capture potential gains in reading skills and other outcomes.
- *Afterschool Randomized Controlled Trials: The Voyager Passport Curriculum in Kentucky 21st Century Community Learning Centers*, designed by the Center for Evaluation and Education Policy at Indiana University, in collaboration with the Kentucky Department of Education. This study aims to compare the impact of previously established 21st Century Community Learning Center (21st CCLC) programs with similar 21st CCLC programs that include the Voyager Passport Reading curriculum (Passport). The study tests the hypothesis that economically disadvantaged youth in grades 2–5 who participate in a high-quality afterschool program with Passport achieve significantly greater learning gains over a 2-year period than do disadvantaged youth who participate in the same program without Passport.

Challenges Across Proposals: What It Says About the Field

All the proposals reviewed faced multiple challenges. By reviewing these challenges, others in the field may be able to address them up-front in their proposals. The most notable challenges were the following:

- **Auspice of the Principal Investigator (PI).** PIs for the proposed studies came from a mix of

academic settings, both in centers and university departments, and included senior researchers in private nonprofit and for-profit research organizations. The diversity of sources where the proposed work would be housed showed strength in terms of expertise from various fields but also signaled a need to create other settings, such as the Afterschool Research Consortium SEDL formed to allow diverse researchers to come together and collaborate on such work.

- **Past Performance of PI in conducting RCTs.** One of the key strengths of successful proposals was the PIs' ability to demonstrate that they have conducted RCTs successfully in the past. Most of the proposals reviewed did not have a PI with a strong past performance in conducting RCTs successfully.
- **Focus of the Intervention.** Most, if not all, of the interventions proposed focused on literacy interventions. This indicates there is a need to continue to focus on identifying and assessing promising interventions in other academic content areas, such as math and science.
- **Scope of the Intervention.** The scope of the proposed interventions varied from focusing on a set of schools in one district (districtwide), to including multiple sites across a state (statewide), to including sites across several states (multistate). The diversity of the scope of the interventions was a strength, particularly in being able to generalize findings to multiple policy-relevant contexts, but raised challenges in terms of conducting large-scale research projects effectively with limited resources.
- **Strength of Evidence to Support the Intervention's Promise.** Most of the proposals did not present adequate evidence to support the intervention as being promising, which became a clear criterion for distinguishing successful proposals from those that were not funded. However, even the interventions that were funded often had limitations to their support that made them, more often than not, only potentially promising. For instance, a common case was one in which the curriculum had undergone

some experimental trials but not in an afterschool setting.

- **Attendance/Participation Rates.** Determining accurate attendance/participation rates for students in afterschool settings in order to ensure that a majority of students would receive the treatment at the intended intensity level or "dose" was perhaps the most important factor in many research designs. Projects that could use estimates from previous years or other similar interventions were able to articulate much more clearly the strengths and limitations of their design in terms of sample size, power (the sample size needed to detect a significant effect), and analyses, particularly the feasibility of conducting sub-group analyses.
- **Monitoring the Fidelity of Random Assignment.** A critical challenge to most proposals was describing a set of procedures for monitoring and ensuring the fidelity of random assignment procedures. Although this was often addressed at the beginning of the study period, successful proposals included monitoring procedures that continued over the course of the study.
- **Student Academic Achievement Outcome Measures.** Most of the successful proposals collected their own norm-referenced student academic achievement outcomes. Although state achievement tests were appealing given their ease of access and the associated reduced costs to the study, their limitations in terms of measuring appropriate outcomes and producing grade-level (i.e., not vertically equated) criterion-referenced data made them much less desirable.

The proposals submitted and the proposal review process used indicate that the afterschool research field is in a dynamic stage of development. The field is evolving based on strong research efforts assessing intervention outcomes to also include more rigorous research efforts to determine the degree and size of the impact on students. The proposals also indicate that this is a recent evolution and that studies involving quasi- and experimental designs need focused attention and supports, such

as those being provided by the Afterschool Research Consortium. Although there is some evidence to support addressing more rigorous studies in the areas of literacy and reading, there appear to be continuing challenges in the field to conducting rigorous research in other areas, such as math and science. There is a need to continue to focus on identifying and assessing promising interventions in these areas while also moving forward toward more rigorous studies to determine these interventions' impacts.

The proposals also indicate that there is increasing capacity and sophistication in designing rigorous studies with high degrees of internal validity in the afterschool field. Past performance and experience conducting RCTs was a key criteria distinguishing successful proposals as well as associated expertise in addressing key challenges to the internal validity of such studies. These capacities have become even more important with the afterschool RCT grantees' work as implementation of the current study designs have begun. Supporting such efforts in a way that promotes collaboration and learning across projects is critical to maintaining momentum in the afterschool research field.

Implementation of Randomized Controlled Trials: Challenges to the Field

SEDL is providing analytic and technical support to all three of the funded projects through the ARC. The ARC brings together SEDL staff, key staff from each research project, technical working group (TWG) members,¹ and representatives from the U.S. Department of Education on a regular basis to facilitate cross-fertilization of ideas and provide strategies to ensure full implementation of RCT designs in afterschool settings. The ARC focuses on substantive content issues in the afterschool field, technical and analytic issues in conducting rigorous RCTs, programmatic and research-to-practice issues, and policy issues. Project teams discuss ongoing implementation, data collection, and analytic challenges presented by the work, explore

opportunities to collaborate and disseminate the findings to the field.

Challenges to Current Implementation: What it Says About the Field

All three research projects faced challenges as they began to implement their study designs in afterschool settings. Many challenges were identified and discussed through the ARC. The most notable challenges were as follows:

- **Recruitment.** Recruitment of sites to participate is a particular challenge for experimental studies. The issue of denial of services, inherent in a design with a "no treatment" or "treatment as usual" control condition, may arise as schools and districts attempt to justify participation. However, successful strategies that the RCTs have used to ensure and gain solid commitment from districts and schools have largely involved making additional efforts, such as holding schoolwide meetings or parent meetings, to discuss the study's benefits and address their concerns. Perhaps the best predictor of successful recruitment has been the development of a good open relationship with the schools and districts (i.e., principals and afterschool program coordinators) early and consistently working to maintain that relationship.
- **Retention.** Retention has been an important issue related to recruitment. Some of the same principles regarding good practice apply here also, including extra efforts to maintain interest and support for the study, such as special holiday events or parties, as well as constant monitoring and inquiry into why students may be dropping out of the program in order to avoid any non-random or systematic problems that may be at play. These efforts are increasingly important as the study is implemented in terms of impacting sample size for analyses and power to detect effects in the current year of the study. It also has implications for the ability of the study to continue into year 2 under the current randomization or may signal the need to resample and re-randomize.

¹ The TWG members currently include Harris Cooper (Duke University), Fred Doolittle (MDRC), Rebecca Maynard (University of Pennsylvania), Elizabeth Warner (U.S. Department of Education), and Peter Witt (Texas A&M University).

- **Monitoring Contamination.** A continuing challenge to the implementation of RCT studies is setting up a regular (weekly) monitoring procedure to protect against contamination. Contamination can occur when a teacher or student is exposed accidentally or systematically to another condition or treatment than they were originally assigned to receive in the study. The most common type of contamination is when students or teachers in the control condition are exposed to the treatment or intervention condition. Typical best approaches have included sharing a weekly database report indicating students' assignment in terms of the condition of the study and involve the research staff talking with the school staff or program coordinators through any potential or realized problems with contamination that have come up and providing clear guidance and rules for addressing such problems.
 - **Collaborating With Program Developers.** The context of an efficacy trial involves trying to implement an intervention under ideal circumstances in order to estimate what sort of effects may be expected from an ideal implementation. Afterschool RCT grantees have worked hard to bring this to the program developers' attention in order to maximize resources brought to the program implementation. This is a constant aspect of the early phases of the project but also continues throughout the course of the study. Researchers, while building firewalls between themselves and the developers to ensure objectivity of the research endeavor, must also monitor implementation and encourage corrective actions by the developer if there are any problems. The nature of the relationship with the developer before the study begins is important and maintaining that relationship is critical over the course of the study. Developers have a lot at stake in an efficacy trial so nurturing and maintaining a good relationship under this situation is often difficult but obviously critical for successful implementation of the trial.
 - **Measuring Implementation.** Afterschool RCT grantees have struggled with how to measure program implementation, how many resources to invest in it, and how to measure it in a way that facilitates better understanding of the role of implementation in the overall analytic plan. Given these concerns, afterschool RCT grantees have focused on key aspects of implementation in the first year of the study, including capturing implementation fidelity, as well as the duration, intensity, and integration of services as part of the participants' experience in the intervention. Afterschool RCT grantees have focused on trying to develop implementation measures that adequately capture these elements to address concerns, if any, about whether the intervention has been fully implemented while also attempting to, if possible, develop quantitative measures of implementation that might be helpful in assessing whether program implementation may be an important predictor in the experimental outcome analyses. One challenge raised during this project has been the need to measure afterschool program quality across all the studies more generally as an additional way to understand program implementation and how it may relate to outcomes. However, while there are many program quality measures, ARC discussions have not focused on any one particular measurement tool as adequate for this work. There are expectations that through other related efforts of the National Partnership there may be a measurement tool available to pilot in the afterschool RCT grantees' sites in year 2 of the studies.
- To date, our conversations and collaborations as part of the ARC have left the impression that, like most studies, the driving concerns raised by implementing RCTs in applied afterschool settings distill down to addressing common threats to internal and external validity—a common goal in any study. In addition, we increasingly see the need for and value of developing a manual of procedures (MOP) approach in afterschool RCTs to maintain procedural integrity in the project and sustain constant and fair application

of rules and procedures over the course of the studies as well as across all the projects. This is a common approach used in clinical trials in medical and bio-medical areas, and we see the need for this to be part of good common practice in social science efficacy trials. From recruitment and monitoring fidelity of random assignment to data collection, management procedures, and analytic plans, such an approach would move the field forward and maximize learning from such trials in a way that would sustain and improve the quality of future trials.

The State of the Field

This set of three efficacy trials has been designed and developed to increase the quantity and quality of evidence to address critical questions in the afterschool research field—whether and under which conditions well-implemented, academically focused curricula have an impact on students' academic achievement in afterschool programs. Given the mixed evidence in the field regarding program efficacy, these are important questions to pursue, and our experience to date is that the resources to do so have been well placed. The efficacy trials of the three reading interventions will provide valuable evidence regarding their potential promise.

However, critical questions in this field remain. These trials focus on academic interventions in afterschool programs, which reduces these studies to addressing what is traditionally a small window of the after-

school day (approximately 45 minutes of a 2- or 3-hour program, 4 days a week at most). Although these trials address an important and recent focus in overall afterschool programming efforts, academic content areas and student achievement outcomes have been only some of the components of intended programming and outcomes for afterschool programs. There remains a need to address overall program impacts on a range of intended outcomes from academic to more traditional youth development and other related outcomes.

Many of the challenges of implementing rigorous RCTs in afterschool settings have been addressed by individual afterschool RCT grantees and shared collectively with the ARC. The ARC holds the potential to capture some of the successes in addressing those challenges through dissemination vehicles such as research and policy briefs, presentations of findings at professional conferences, and peer-reviewed publications. We remain convinced that significant implementation challenges addressed and lessons learned need to be shared with others to address the steep learning curve that exists around implementing such designs in applied settings. However, the value of the ARC in serving in this role is still an open question and one we hope to assess and improve through formative feedback in order to support more high-quality research to inform better practices and policies.

Copyright © 2008 by SEDL

This research brief was produced in whole or in part with funds from the U.S. Department of Education under contract number ED-03-CO-0048. You are welcome to reproduce it and distribute copies at no cost to recipients; please credit SEDL as publisher.

AFTERSCHOOL RESEARCH CONSORTIUM

Advancing Research, Improving Education



Available online at www.sedl.org/afterschool/resources/