



Julia Coffman

Center for Evaluation Innovation

Kristie Kauerz

University of Washington

NOVEMBER 2012

EVALUATING

PreK-3rd Grade

Reforms

PreK-3rd Grade Reform Goals

1 Resources for cross-sector work:

Mechanisms and structures exist that reflect, support, and sustain shared vision, mission, and accountabilities between early childhood and K-12.

2 Administrator and leadership quality:

Administrators--school superintendents, principals, and early childhood directors--actively and visibly create a culture that supports PreK-3rd.

3 Teacher and teaching quality:

PreK-3rd grade teachers actively and visibly provide high-quality instruction and effective experiences for children.

4 Instructional tools:

PreK-3rd grade standards, curricula, and assessments focus on both academic and social-emotional skills, and are aligned to create instructional coherence.

5 Learning environment:

Children's learning environments--campuses, buildings, schools, and classrooms--promote collaborative relationships, actively engage children in a variety of learning settings, and support the health and wellness of children and staff.

6 Data-driven improvement:

Improvements to schools, classrooms, instruction, professional development, and other systems are based on current data.

7 Engaged families:

Families are actively and systematically involved in their children's education, PreK-3rd, as a core instructional strategy.

8 Access, transitions, and pathways:

Every child, especially those most at-risk for school failure, has access to a clear pathway of high-quality education from PreK through 3rd grade.

SOURCE: Kauerz, K., & Coffman, J. (2012). *Framework for Planning, Implementing, and Evaluating PreK-3rd Grade Approaches*. Seattle, WA: College of Education, University of Washington.

Introduction

PreK-3rd grade reforms are taking hold in school districts, communities, and states across the nation.

This approach to education reform builds on research that shows that 3rd grade is a critical point in a child's education. Children who cannot read or do math at grade level by 3rd grade are unlikely to catch up later. To make sure that all children are proficient in math and reading by this point in their development, schools are instituting PreK-3rd reforms that treat children's learning as an accumulation of experiences over time.

PreK-3rd efforts connect high-quality early learning programs with high-quality elementary school experiences. They integrate and align the subject matter focus of K-3 with the child development focus of early education. To accomplish this, teachers, school administrators, and families work together across age/grade levels to ensure that all children's learning experiences build on past years and connect with those to come, striving for the kinds of goals outlined in the box at right.

While PreK-3rd reforms are spreading quickly, evaluations to assess them are not keeping pace with their growth and expansion. States, school districts, education leaders, funders, policymakers, and others want to know whether PreK-3rd grade reforms are making progress and producing results. Where they exist, evaluations tend to concentrate on isolated areas of the PreK-3rd grade continuum. For example, evaluations have examined the impact of PreK in terms of its success in achieving school readiness for children. Although high-quality PreK is an important part of PreK-3rd efforts, it is only one part of the equation. Other evaluations have focused on effective practice in the early elementary grades, but most of this work has occurred in isolation from evaluations conducted on PreK. Studies fail to consider results for children as a responsibility shared across early learning programs, grades K-3, families, and communities. In addition, few evaluations help schools and districts understand whether or not they are building approaches that will produce impacts for children by the end of 3rd grade.

Part of the challenge of PreK-3rd evaluation is that it can be hard to know where to start and what to evaluate. PreK-3rd reforms can be complicated. They cross the traditional boundaries of early care and education and K-12 education. They cross multiple grades and aim for practice and policy changes in many areas (e.g., professional development, accountability, and data) and at multiple levels (e.g., classrooms, schools, and districts). While improving children's learning and development is the ultimate objective, PreK-3rd grade reforms must first accomplish multi-faceted changes within and across multiple groups (e.g., teachers, education leaders, and families). As a result, it can be difficult to understand all of the evaluation options available and how to choose a direction, particularly if limited evaluation resources are available (and they almost always are).

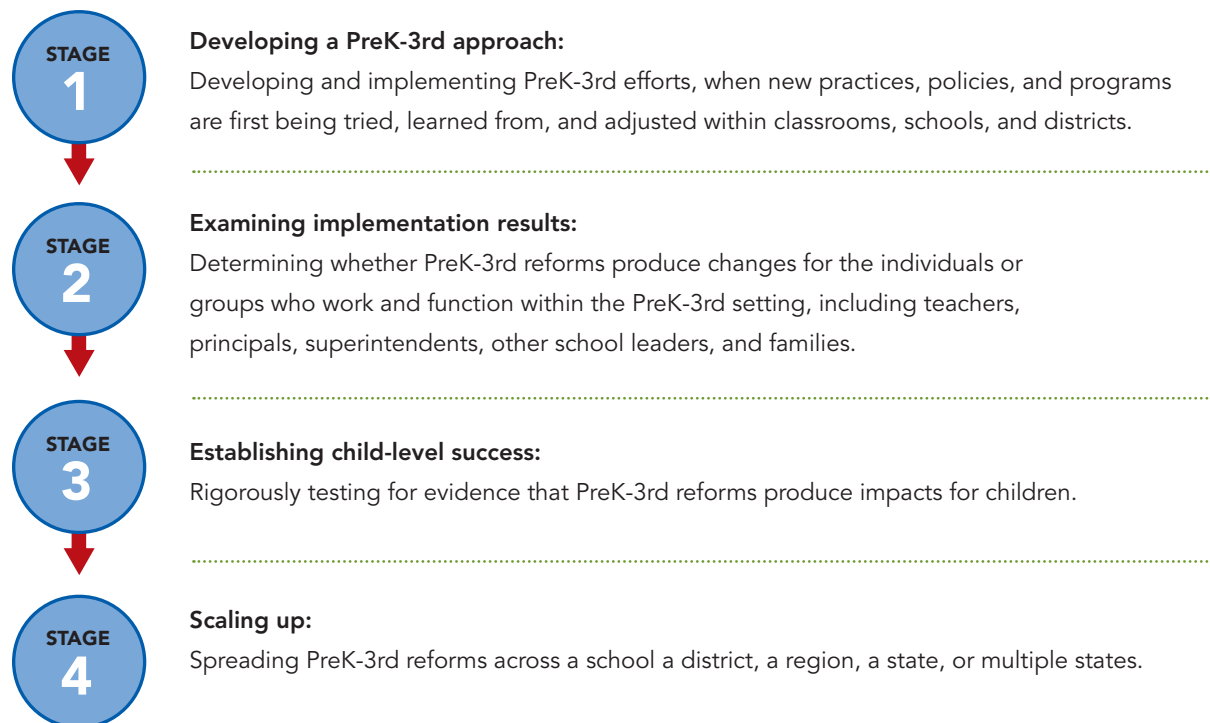
This paper is intended to help stakeholders in early childhood and elementary education think about how to use evaluation at different stages in the development of a PreK-3rd effort. It is not a how-to guide to PreK-3rd grade evaluation. Rather, it offers a conceptual model that attempts to clarify ideas, approaches, and language about evaluating PreK-3rd grade reforms. This brief spells out when and how to use evaluation, what kind of evaluation to use in order to maximize its utility, and the type and level of evidence required from evaluation.

A Conceptual Model for PreK-3rd Grade Evaluation

Evaluation is not a one-size-fits-all undertaking. Because PreK-3rd grade reforms look different in different places, the type of evaluation needed will depend on where the PreK-3rd effort is in its development and the types of reforms being implemented.

Some districts and schools are implementing comprehensive reforms that aim for improvements in all or most aspects of PreK-3rd implementation. (See the box on PreK-3rd grade reform goals at the beginning of this brief.) Others are taking an incremental approach and focusing on one or several aspects at a time while recognizing that PreK-3rd reforms should be comprehensive to be most effective. Because these efforts look different and some PreK-3rd efforts emphasize certain activities more than others, the approaches used to assess them also should look different. Evaluations should be tailored to a PreK-3rd effort's unique goals, activities, and its stage of development.

The process of developing, implementing, and scaling up PreK-3rd grade reforms can be thought of as having four stages. Ideally, PreK-3rd efforts should proceed through each stage in sequence.¹



¹ Adapted from McDonald, S. (2009). Scale-up as a framework for intervention, program, and policy evaluation research. In G. Sykes, B. Schneider, & D. N. Plank (Eds.), *Handbook of education policy research* (pp. 191–208). New York: Routledge Publishers (for the American Educational Research Association).

As a first step in thinking about the kind of evaluation needed, those implementing PreK-3rd reforms should think about where their efforts are in the just-described sequence.

Evaluation should play a role at each stage. Knowing where a PreK-3rd effort is in its development will better position implementers and evaluators to identify at each stage the appropriate evaluation questions to be asked, the data needed to address those questions, and how evaluation findings will be used and by whom. The table below summarizes key characteristics of the evaluation approaches suitable for each stage of development for PreK-3rd grade approaches.

Four Stages of PreK-3rd Reforms and Relevant Evaluation Approaches

	STAGE 1 Developing a PreK-3rd Approach	STAGE 2 Examining Implementation Results	STAGE 3 Establishing Child-Level Success	STAGE 4 Scaling Up
Why Evaluate	Support and inform PreK-3rd development, collecting data that help to continuously refine and improve implementation.	Determine the effectiveness of specific projects, policies, practices, or programs for those working or functioning within the PreK-3rd setting.	Test for evidence that PreK-3rd reforms produce impacts for children.	Examine the effectiveness of reforms when they are implemented among large numbers of individuals and across many contexts.
For Whom	Internal users (PreK-3rd implementers)	Internal and external users, such as funders and policymakers	External users	Internal and external users
Evaluation Questions	How are parts or aspects of the PreK-3rd setting changing? What can be done to improve PreK-3rd implementation?	Are PreK-3rd reforms producing changes for the groups involved in implementing those reforms?	Do PreK-3rd reforms produce impacts for children?	Are reforms being implemented with high fidelity? Do reforms produce results across places or populations?
What Data to Collect	Implementation and process data on specific, discrete PreK-3rd reforms Baseline data on child status and performance	Results measures or indicators for the group(s) of interest (e.g., teachers, superintendents, or parents)	Child-level impacts, or achievement and development data at every stage along the PreK-3rd continuum and at the end of 3rd grade	PreK-3rd implementation results and child-level impacts in all of the school districts in a certain geographic location, such as a city, county, region, or state
When to Evaluate	Early on, during the development of the reform effort	After PreK-3rd reforms have been developed and are being implemented with fidelity	When implementation data and results are sufficiently promising to warrant rigorous testing	During the scale-up process
How to Evaluate	Formative Evaluation Quantitative and qualitative methods, such as observations, satisfaction surveys or focus groups, staff interviews, and document or record reviews	Summative Evaluation Results monitoring (e.g., through surveys, interviews, or observational tools) that tracks changes in groups at different points in time (e.g., pre- and post-reform implementation)	Impact Evaluation Experimental or quasi-experimental designs that use counterfactuals to compare what actually happened with what would have happened in the absence of the PreK-3rd effort	Formative and Summative Evaluation Both implementation monitoring using quantitative/qualitative methods and impact tracking (does not require an experimental design)

1

Developing a PreK-3rd Grade Approach
with Formative Evaluation

This stage involves the initial development and implementation of PreK-3rd efforts, when new practices, policies, and programs are first being tried, learned from, and adjusted within classrooms, schools, and districts. At this stage, reforms focus on changing the way the PreK-3rd setting operates, intersects, and aligns, as well as how individuals and groups—namely teachers, administrators, and other stakeholders—work within it.

Specifically, the focus here is on making sure the implementation of specific, discrete PreK-3rd reforms is done with fidelity and, if not, making adjustments. For PreK-3rd efforts that have been identified as “best practice” or “evidence-based,” realizing the full benefits of those practices requires that they be implemented as designed or intended. For example, if a PreK-3rd effort aims to foster both horizontal and vertical teamwork among teachers, implementers will want to ensure that regular and inclusive teacher team meetings occur and that teachers’ time together is spent discussing content relevant to PreK-3rd alignment, whether it be alignment of curriculum, professional development, data, or some other aspect of their PreK-3rd grade approach.

This stage is important, as implementation must be functioning properly if impacts for children are to be realized. PreK-3rd objectives must be clear and agreed upon across stakeholders; resources should be in place; and PreK-3rd implementers should be reasonably confident that activities are designed to have an impact and are being implemented properly. PreK-3rd efforts that are in the “early” stages of development might stay in this initial stage for some time, especially if the implementers are taking an incremental approach to reform.

Stage 1 evaluation should support the development of the PreK-3rd effort, collecting data that can help to continuously refine and improve it. At this point, PreK-3rd reforms are not ready to be examined to determine if they are producing impacts for children. Rather, evaluation should focus on implementation. This is, however, a crucial time for gathering baseline data on children’s status or performance to allow for impact evaluation later in Stage 3.

Specifically, *formative evaluation* is appropriate at this stage. It concentrates on examining and changing implementation as it occurs and provides timely feedback about activities that allows implementers to make adjustments as they go. Formative evaluation is usually conducted for an internal audience of those who are developing or implementing reforms.

Formative evaluation in Stage 1 typically uses some form of program monitoring or process evaluation. These approaches address questions about the extent to which reforms are reaching their target populations, whether implementation matches design expectations, and whether adequate resources are being used to deliver the reform.² Monitoring and process evaluation focus on the internal operations of the PreK-3rd setting in order to understand the strengths and weaknesses, and examine *how* the *implementation* of reforms is trying to produce results rather than examining the results themselves. A wide array of both quantitative and qualitative methods can be used in a formative evaluation, such as observations, participant surveys or focus groups, staff member interviews, and document or record reviews.

² Rossi, P., & Freeman, H. (1993). *Evaluation: A Systematic Approach*. Newbury Park, CA: Sage Publications.

2 Examining Implementation Results with Summative Evaluation

PreK-3rd efforts move to the second stage of development when there is indication that they can produce changes for the individuals or groups who work and function within the PreK-3rd setting, including teachers, principals, superintendents, other school leaders, and families. PreK-3rd reforms require people to change how they act. For example, if PreK-3rd reforms are instituting new professional development approaches for teachers, teachers' practices and performance in the classroom should be measurably different after participating in professional development. During Stage 2, this may include examining how teachers organize their classrooms, choose their learning activities, and use data to assess and discuss children's progress. Similarly, teacher teamwork should be changing the way teachers use data to make instructional decisions and how they differentiate instruction for children in their classrooms. Teacher teamwork may also influence how teachers share and receive feedback from each other and from their principal.

Summative evaluation is useful for determining whether changes are occurring with individuals and groups functioning within the PreK-3rd setting. It determines the effectiveness of distinct PreK-3rd efforts, policies, practices, or programs. Summative evaluation is conducted after a reform effort has been developed and implemented with confidence and fidelity. Typically, summative evaluations are conducted for an external audience of decision makers, and often for the purpose of deciding whether to continue the effort or reconfigure it if results are not being achieved. Like formative evaluations, summative evaluations should include the collection of some baseline data early on.

Most evaluation at this stage is a form of results monitoring. This approach typically involves capturing results measures or indicators for the group(s) of interest (e.g., teachers, superintendents, parents) at various points and tracking whether changes are occurring over time. For example, this may involve administering a survey or questionnaire to teachers or parents at the start and end of the school year to capture changes in knowledge, attitudes, beliefs, skills, or behaviors. Unless a control or comparison group is used, results monitoring does not infer causality between the PreK-3rd effort and its results unlike Stage 3 impact evaluations discussed below. Observed changes in results may be attributable to multiple factors, not just PreK-3rd reforms.

3 Examining Child-Level Success with Impact Evaluation

PreK-3rd grade efforts move to the third stage of development when implementation results are promising and reforms are ready to be more rigorously tested for evidence that they produce impacts for children. Child-level impacts refer to strong achievement and healthy development results at every stage along the PreK-3rd continuum and by the end of 3rd grade. The goal of Stage 3 is to assess the magnitude of the cause-and-effect relationship between PreK-3rd reforms and their impacts on children to determine if those reforms are worthy of being scaled up to other classrooms and schools.

The definition of when a PreK-3rd grade effort is “ready” to be rigorously evaluated for its impacts for children is a blurry one. Timing is important, as evaluation at this stage can carry high costs. Unfortunately, no readiness checklists or hard-and-fast rules exist. It is, however, important to be sure that multiple elements are in place. At the least:

- Reforms--the details of which will vary across locations--should include core aspects from each of the eight key aspects of PreK-3rd implementation (see the box on the first page).
- The actual implementation of PreK-3rd reforms should match its intended implementation in the classroom, at the school level, and at the district level (determined in Stage 1).
- The behavior of teachers, administrators, and families should be measurably changed from what it was before PreK-3rd reforms were instituted (determined in Stage 2).

This stage is critical if the intent is to scale up PreK-3rd reforms. It makes little sense to invest in disseminating reforms without being confident that they work. Replicating PreK-3rd grade reforms in multiple sites before there is substantial certainty that they are effective is both politically risky and an inefficient use of resources.

Impact evaluation is needed at this stage. This is a rigorous type of summative evaluation that answers the overall question of whether “net effects” have occurred and whether changes in child achievement and development would have occurred if the PreK-3rd effort had not been undertaken. Impact evaluations use rigorous research designs to answer cause-and-effect questions, establishing whether impacts can be directly attributable to PreK-3rd efforts.

For PreK-3rd grade efforts, impact evaluations must be longitudinal endeavors, gathering baseline data early on and then tracking children as they move from PreK through at least 3rd grade. Many studies will extend beyond that timeframe and examine child impacts into adolescence and adulthood.

Evaluations at this stage often use experimental designs, also known as randomized controlled trials

(RCTs). These designs have a defining characteristic: the random assignment of individuals (or classrooms or schools) to intervention and control groups (also called the counterfactual) to compare what actually happened with what would have happened in the absence of the PreK-3rd effort. The intervention group participates in the program or intervention while the control group does not. Experiments or RCTs are considered the strongest design option when an evaluation is exploring the cause-and-effect relationship between PreK-3rd reforms and their impacts.

Quasi-experimental designs are another impact evaluation option. They are like experimental designs in that they aim to make cause-and-effect statements about impacts. Such designs are often used when random assignment is not possible. Because they cannot use randomization, most quasi-experimental designs construct comparison groups or other types of counterfactuals to examine an intervention’s impacts for those who do and do not participate.

4

Scaling Up with Formative and Summative Evaluation

PreK-3rd grade efforts should move to this fourth stage when reforms have been tested and established as effective, and the decision has been made to scale them up across a school district, a community, a region, a state, or multiple states. Scaling up involves disseminating PreK-3rd reforms so that more schools and children can experience their benefits. The challenge at this stage is scaling up effectively and efficiently. For that reason, evaluation can add tremendous value during the scale-up process.

This stage aims to examine the effectiveness of reforms when they are implemented among large numbers of individuals and across many contexts. This may mean, for example, examining PreK-3rd grade implementation results and child-level impacts in all of the school districts in a certain geographic location, such as a city, county, region, or state.

Data collected at this stage can provide feedback to help refine the scale-up process or to develop guidelines that ensure PreK-3rd reforms operate as intended across real-life settings with all of their complications. If child-level impacts are not present across settings, this stage can help to determine why not, and what it might take to ensure they do, such as revisiting the fidelity of implementation across sites.

During scale-up, evaluation should be both formative and summative. It should both support the scale-up process and continue to measure outcomes at scale. For example, formative evaluation should examine the contextual factors that are influencing PreK-3rd implementation in different settings. These data can help to refine dissemination efforts. Summative evaluation can ensure that expected impacts are sustained at scale. At this stage, evaluation should be used to answer questions such as: Are impacts being maintained as reforms go to scale? What implementation problems or challenges are arising during scale-up and how can they be addressed? What kinds of capacities and resources are needed to support the scale up of reforms?

Conclusion

PreK-3rd grade reforms are complex endeavors. Understanding how to evaluate them—including when to do evaluation, what type of evaluation to use, and how to use it—can seem equally complex. This brief attempts to bring some overall clarity to the question of how PreK-3rd efforts can and should be evaluated. It offers a conceptual model to help PreK-3rd funders, implementers, and evaluators identify where a PreK-3rd effort is in its development, as well as the type of evaluation needed given its stage of development. Several overall conclusions also emerged during the development of this brief. They are noted below.

Evaluation is useful and necessary at all stages of PreK-3rd development. From the very earliest phase of an effort's development and implementation, evaluation can play a useful role. Early on it is most useful for identifying implementation challenges and suggesting ways of addressing them. Later it is useful for measuring whether implementation results have occurred or whether a cause-and-effect relationship exists between PreK-3rd reforms and child-level impacts. For PreK-3rd efforts that are scaling up, evaluation can help to ensure that outcomes are sustained and that implementation is of sufficient fidelity to continue to produce impacts.

Some PreK-3rd grade evaluations are in progress, but an overall lack of evaluation is occurring among PreK-3rd grade efforts. Many PreK-3rd efforts across the nation are just getting started. Unfortunately, many have not integrated evaluation into their growth and development. For some, this is a resource issue. However, not all evaluation has to be expensive. Formative evaluation and summative evaluation in the form of results monitoring can be fairly resource-efficient, especially if the evaluation is focused on priority areas among PreK-3rd reforms, such as capturing results in which decision makers are most interested. In addition, some formative and summative evaluation may be managed internally, without requiring an external evaluator. While longitudinal impact evaluation does require outside expertise and can be costly, only certain PreK-3rd efforts are at a point where they are ready for that level of evaluation.

Finally, the *Framework for Planning, Implementing, and Evaluating PreK-3rd Grade Approaches* is a useful companion for those identifying what to evaluate and how. Organized around the eight main goals or “buckets” of PreK-3rd reforms listed at the beginning of this brief, the Framework offers several layers of detail that can be useful for the evaluation of PreK-3rd reforms:

- For those in Stage 1 of their development, the Framework identifies important PreK-3rd activities and self-assessment questions that signal an effort’s depth of implementation and alignment.
- For those in Stage 2, it offers individual and group-level indicators that relate to each reform area, along with example methods and tools that can be used to capture those indicators.
- For those in Stage 3, the Framework can be used to help determine if a PreK-3rd effort is ready for rigorous testing of its effects on child-level impacts (e.g., whether there are substantial aspects underway across all eight goals and whether implementation is deep enough).
- For those in Stage 4, the Framework can be a guide for selecting the indicators that should continue to be monitored after a PreK-3rd effort goes to scale.



Julia Coffman (jcoffman@evaluationinnovation.org) is founder and director of the Center for Evaluation Innovation in Washington D.C.

Kristie Kauerz (kauerzk@uw.edu) is a Research Assistant Professor at the University of Washington (UW) College of Education in Seattle, WA. She also played a lead role in the PreK-3rd Grade Initiative at Harvard Graduate School of Education (HGSE) in connection with a series of professional education institutes co-chaired by HGSE faculty members Hiro Yoshikawa, Rick Weissbourd, Nonie Lesaux, and Stephanie Jones.

*Suggested Citation: Coffman, J. & Kauerz, K. (2012). *Evaluating PreK-3rd Grade Reforms*. Seattle, WA: College of Education, University of Washington.*

*The development of this paper and the *Framework for Planning, Implementing, and Evaluating PreK-3rd Grade Approaches* was generously funded by the Foundation for Child Development, the W.K. Kellogg Foundation, the W. Clement and Jessie V. Stone Foundation, and the Bill & Melinda Gates Foundation.*