The Future of Children
PRINCETON-BROOKINGS

Research-Practice Partnerships to Strengthen Early Education

VOLUME 31 NUMBER 1 SPRING 2021

A COLLABORATION OF THE PRINCETON SCHOOL OF PUBLIC AND INTERNATIONAL AFFAIRS AND THE BROOKINGS INSTITUTION

The *Future of Children* promotes effective policies and programs for children by providing timely, objective information based on the best available research.

The *Future of Children* would like to thank the Foundation for Child Development, the Spencer Foundation, the Heising-Simons Foundation, and Cynthia King Vance for their generous support.

ISSN: 1054-8289
ISBN: 978-0-9814705-2-8

The Future of Children
PRINCETON-BROOKINGS

VOLUME 31 NUMBER 1 SPRING 2021

Research-Practice Partnerships to Strengthen Early Education

Introducing the Issue

Daphna Bassok, Anna Markowitz, and Pamela Morris

High-quality early childhood education programs can have lasting impacts, particularly for children in low-income families or families that have been historically marginalized.[1] Beyond their direct benefits for children, early childhood education (ECE) programs also support families and the economy by providing safe, warm environments for young children while their parents work—an attribute that recently gained new significance when COVID-19 upended in-person instruction. But even though the promise of early education—for children, for families, for equity, and for society—is widely recognized, delivering effective ECE programming at scale remains elusive. Findings from promising research studies rarely make their way into early childhood practice; at the same time, policy and practice decisions are often made without research evidence to guide them. Many large-scale ECE programs don't consistently offer the types of high-quality experiences that research suggests support development, and thus they don't deliver on their promise to ameliorate poverty-related and other inequalities.[2] The discrepancies between what we know, what decision makers need, and the reality of children's experiences have caused many researchers, policy makers, and practitioners to wonder if there's a better or faster way to produce change.

To build a system of high-quality early care accessible to all children, some early education research must pivot to questions, methods, and timelines that align with the needs of policy makers and practitioners who are making high-stakes, systems-level decisions. At the same time, research must stay firmly rooted in the science of how children develop.[3] This issue of the *Future of Children* argues that research-practice partnerships (RPPs) are a particularly promising strategy for doing both well, and thus offer another critical tool to support high-quality early childhood education.

RPPs are long-term collaborations between researchers and policy makers and/or practitioners that are designed to improve educational outcomes through sustained collaboration and commitment.[4] These partnerships are defined by their longevity, mutual decision-making and compromise, and the commitment of both parties to larger-scale, systems-level problem-solving,

Daphna Bassok is an associate professor of education and public policy and associate director of EdPolicyWorks at the University of Virginia. Anna Markowitz is an assistant professor of education at the University of California, Los Angeles. Pamela Morris is a professor of applied psychology at New York University's Steinhardt School of Culture, Education, and Human Development and a faculty affiliate of the NYU School of Global Public Health. Bassok and Morris served as co-editors of this issue; Markowitz joined them in writing this introduction.

rather than a single project or research question.[5] When done well, RPPs can produce research that's both more relevant and useful to policy makers and more rigorous and/or innovative than the studies that researchers could undertake outside of a formal partnership.

But this work isn't easy. Researchers and policy makers are used to operating with very different timelines, goals, and incentive structures, and these differences can present challenges for effective partnership. A clear understanding of what partnership work requires, and strategies for overcoming these hurdles, are essential for creating the kind of long-term, open relationship that results in better policy *and* stronger scholarship. This issue explores these challenges as well as the potential of partnership work for helping to transform ECE systems.

Unlike a typical *Future of Children* issue, which describes what's already known about a topic that affects children's lives, this issue focuses on an *approach* to research and practice. We aim to distill lessons from existing ECE partnerships and to provide a road map for researchers and policy/practice leaders who want to collaborate. The articles examine both the challenges and the opportunities of partnership work, and show how partnerships may, in some instances and when carefully designed, pave the way for systems-level change. Each article is centered on a major challenge or opportunity, such as creating a research or funding agenda, developing tools, taking innovation to scale, navigating different timelines, finding a balance between rigor and feasibility, and building capacity. Three commentaries, written by a university dean, a practice leader, and policy organization researchers, describe how the contexts in which policy makers and researchers work—universities, state systems, and research firms, respectively—can effectively support RPP work. As a whole, the issue sets the stage for strengthening and investing in partnership work in early education.

Historical Role of RPPs in Early Education

ECE research has its roots in RPPs. The three major studies whose results have driven much of the public investment in ECE—the Perry Preschool Project in Ypsilanti, Michigan; the Abecedarian Project in North Carolina; and the Chicago Child Parent Centers—all sought to provide high-quality ECE while continually learning and refining their programs through research. These studies, two of which were randomized controlled trials, have produced crucial information about early education for nearly 50 years and have greatly influenced modern ECE policy making.[6]

Federal investments in ECE—for example, funds to expand the number of slots or improve quality—are often coupled with evaluation requirements, to learn from the implementation experience, to test whether the investments had the desired effects, and/or to use the findings for improvement. The federal government's primary early education program, Head Start, has, from its inception, been heavily influenced by research. Head Start planners drew on a nursery school movement that paired child development research with university-based child care, and sought to create a program that would not only provide a service for children and families, but would also use research to deepen our understanding of what children need to thrive, forming a feedback loop to improve the program.[7]

Head Start's commitment to using data to improve practice is visible today in the Head Start Impact Study, Head Start Program Information Reports, Head Start Graduate Student Research Grants, and the Head Start Family and Child Experiences Survey.

Other federal initiatives echo this commitment. Since 1995, the Administration for Children and Families Child Care Policy Research Partnership grant program has supported collaborations between researchers and agencies administering the federal Child Care and Development Fund (CCDF), with the goal of generating research on key issues for CCDF decision makers. More recently, the Obama administration's Race to the Top Early Learning Challenge funded the evaluation of states' efforts to improve the quality of early learning, opening the door for long-term collaborations between researchers and policy makers—though it didn't explicitly require such collaborations.

Resurgence of Partnership Research in Education

Partnerships between researchers, policy makers, and practitioners aren't new in ECE or in education research more broadly. But over the past decade, calls have intensified for partnered, context-driven research. A 2012 National Research Council publication asserted that science provides essential evidence for decision makers, and that the most useful research is designed in collaboration with policy makers and practitioners, working at the ground level.[8] The push for co-created research in education set the stage for major investments from both federal and foundation sources. The aim of these investments was to increase the number of RPPs, as well as to develop concrete methods and tools for RPP work. In short, RPPs were intended to put science to work by closing the gap between research and practice/policy.

At the federal level, from 2013 to 2019 the Institute of Education Sciences (IES) funded the Researcher-Practice Partnerships in Education Research grants program, which was designed to support research projects on issues deemed high priority by education agencies, as well as the development of a plan for future research.[9] Moreover, IES predoctoral training grants now require embedded policy partnership experiences, signaling a commitment to producing researchers who can work alongside decision makers.

Private foundations have strongly supported partnership work, both by funding partnership research itself and by working to improve the science and practice of RPPs. Since 2015, the Spencer Foundation's Research Practice Partnership grants program has funded 32 RPPs, several of which focused on early education (a new application window also opened in late 2020). The foundation has also hosted several meetings for its grant recipients to share what works best in this line of work.[10] The William T. Grant Foundation created a set of online resources for funders looking to support RPPs and for research-practice collaborators hoping to develop a partnership.[11] In 2016, the Spencer, William T. Grant, Bill and Melinda Gates, Annie E. Casey, and Wallace foundations established the National Network of Education Research-Practice Partnerships (NNERPP), aimed at developing, supporting, and connecting education RPPs.[12] NNERPP holds meetings, collects and disseminates resources, and helps RPPs collaborate with one another.

Finally, as Jacqueline Jones writes in this issue, the Foundation for Child Development recently launched an initiative to develop a set of RPPs to study the early childhood workforce in New York City.

Though investments in partnered research have grown, the evidence for its benefits for policy, practice, and scientific discovery is still underdeveloped. The best evidence to date comes from the National Center for Research in Policy and Practice's descriptive study of the IES RPP grants. This study found that researchers, practitioners, and policy makers alike valued the partnership work, felt it improved their understanding of each other's contexts, and led to an increase in the application of research.[13] About a third of partnerships reported that their collaboration had created a policy change.[14] At the same time, IES—the preeminent federal agency supporting educational research—suspended explicit funding for partnerships in 2019, justifying that decision, in part, by noting their concern that RPPs are too process-oriented and insufficiently focused on improving student outcomes.[15] This issue of the *Future of Children* aims to support the next generation of ECE RPPs by sharing lessons from a diverse set of partnerships, thus laying the foundation for future collaborations that support the systemic change we need to meaningfully shift young children's learning outcomes through ECE.

Using RPPs to Support System Building

First-generation RPPs provided compelling evidence that high-quality ECE interventions can work. But they also showed that it's difficult to bring effective programs to scale. Today, questions about the short-, medium-, and long-term impacts of ECE programs remain important, as do questions about the costs and benefits of these interventions. Yet, the question "Does ECE work?" is far less relevant now for policy makers and practitioners than are questions about how to improve ECE with constrained resources. Early learning opportunities in the United States come through a fragmented delivery system that must cope with limited funding, workforce challenges, low supply, and inconsistent quality standards.[17] In this context, policy makers and practitioners seek information about strategies to raise ECE quality (for example, through professional development or curriculum), strategies to reach specific populations (such as children with disabilities or dual language learners), and strategies to design policies that facilitate system-level change (such as licensing and accountability systems or workforce supports).

Of course, much research, including rigorous randomized trials and descriptive implementation studies, has already examined specific ECE improvement strategies. But because many of these studies are small and evaluate researcher-designed and researcher-implemented strategies, they often provide little information about implementation at scale.

Thus policy makers and practitioners are forced to make educated guesses about how to invest their limited resources. Nationwide, they're experimenting with an array of strategies to better serve young children, ranging from professional development interventions to curricular reforms to tax credits for early educators. Yet they have little research to guide them. By partnering with researchers, policy makers can evaluate whether their initiatives have the desired

effects, and can also learn real-time lessons that allow for refinement. On a larger scale, RPPs offer the chance to turn the multitude of ad hoc efforts to solve unique ECE problems into a laboratory that tests ideas for improving ECE at scale. This issue aims to learn from existing partnerships to optimize this process, and thereby enhance the next generation of RPPs. With this in mind, we present several case studies, each addressing different challenges and opportunities related to ECE partnership work. We set the stage in this introduction, offering examples from our own partnered ECE research to highlight emergent findings from the issue overall, articulating some of the unexpected benefits of partnered research, and demonstrating why now may be the right time to use partnered research to support ECE.

Three Exemplars

Here we highlight three of our own ECE partnerships—with the New York City Division of Early Childhood Education in the Department of Education, the Louisiana Department of Education (LDOE), and the Virginia Department of Education (VDOE) (see also the article by Rachel Abenavoli and colleagues and the commentary by Jenna Conway). All these partnerships aim to use ECE research to support ambitious system-building. They also show the value of investing in partnerships through unusually disruptive times, such as the COVID-19 pandemic.

In New York City, leaders from the mayor's office and the Division of Early Childhood Education (NYC DECE) began building a free universal prekindergarten system known as Pre-K for all—extending free pre-K from a loosely integrated system serving 19,000 children to a fully integrated one serving 70,000 children, in just two years. To meet this challenge, they relied on a brain trust of New York University (NYU) faculty, staff, and students to provide research support for the city's decision-making, including efforts to take the pulse of the rapidly expanding system and to design lasting research architecture. The partnership led to a joint $5 million research project from the Institute of Education Sciences to rigorously examine investments in teachers' professional development to support high-quality pre-K programming and the learning and development of the city's four-year-olds (see Abenavoli and colleagues).

The Louisiana partnership formed around the passage of Act 3, a state law known as the Early Childhood Education Act. This legislation was designed to break down barriers between child care, Head Start, and state pre-K, creating a cohesive ECE system. At the heart of the reform was mandatory accountability via a Quality Rating and Improvement System (QRIS), which focused on the quality of teacher-child interactions in all publicly funded programs. Louisiana's new QRIS involved multiple in-person observations a year in every publicly funded ECE classroom in the state. Both its broad scope and its focus on process quality were unique, and policy makers were eager to partner with researchers to evaluate their quality improvement measure and its impact on ECE.

Finally, the Virginia RPP was formed to support a set of ECE quality improvement efforts led by VDOE, with support from a federal Preschool Development Grant Birth through Five (PDG). Virginia is piloting a large financial incentive program for early educators that aims to recognize their critical role in children's development,

acknowledge the substantial disparities in compensation between teachers in child care centers and those in school-based ECE, and reduce teacher turnover. Researchers at the University of Virginia, who serve as the PDG's evaluation partners, are leading a large randomized controlled trial of the financial incentive program.

All three of these RPPs aim to contribute to major ECE reform efforts in their locales, and all three rely on partnerships with policy makers who value and invest in data-driven decision-making. The three RPPs are also tackling topics of great interest for ECE, including universal access, measuring quality, and supporting early educators. The RPPs thus create opportunities to help policy partners make decisions in real time, and to provide rigorous evidence on questions of broad interest.

Driving Improvements with RPPs

RPPs can be used to refine practice and drive continued improvement as policies mature. For instance, in the first year of the NYC DECE partnership, the NYU team coordinated with a team at Westat, the research firm hired to descriptively examine the rollout of the city's Pre-K for All program by collecting fall and spring assessments of children's school readiness. Key activities included piloting the use of tablet PCs for easy-to-administer assessments of children's executive function, self-regulation, and pre-academic skills. To help the city deploy resources strategically, the NYU team also created easy-to-read visual displays combining key neighborhood characteristics with information on the location of Pre-K for All expansion sites and child assessment data.

The second year of the partnership brought a coordinated effort to use research in support of the city's commitment to quality. This involved co-developing research-based quality standards, providing research evidence to guide the selection of professional development models, and constructing a data-based decision-making process to allocate frequency of coaching support. NYU initiated feedback sessions with coaching staff who support teachers' professional learning and with teachers and leaders of preschool programs. These sessions gave the NYC DECE valuable information as it refined its professional learning model to be both integrated across training and coaching, and aligned with teachers' and leaders' needs.

Louisiana's novel approach to QRIS also relied on research for continual refinement. The QRIS required the state to regularly and accurately measure quality across all publicly funded ECE programs, ideally employing a measure that teachers and families would understand, and that could be used to improve classroom practice across all sectors. Knowing the limitations of many existing QRISs, LDOE chose to focus on a single observational measure of teacher-child interactions, one that's been consistently, though modestly, linked to children's learning gains. But the department had questions about the feasibility of accurately collecting so much observation data, and about strategies for ensuring that the quality measure was linked to children's development. Initially, the partnership examined jointly identified questions to guide design decisions and test the validity of the state's approach.[18] Over time, LDOE shifted from questions about accurately measuring quality at scale to questions about driving improvement, thus focusing on new strategies to support early educators. State and local partners have used their findings,

particularly a set of large-scale surveys of early educators, to guide real-time policy decisions and to support various grant-writing, policy, and public-facing actions.

The Virginia RPP, the most recent of the three, involves a combination of large-scale workforce surveys to better understand the experiences of the state's early educators, and a field experiment to evaluate its early childhood teacher recognition program. The survey findings have already been used to refine policy. For instance, in the second year of the state's PDG initiative, limited resources precluded giving all ECE teachers the incentive payment. Findings from teacher surveys helped the state decide to focus the program on child care lead and assistant teachers, as their turnover rates were much higher and their financial needs much greater than those of teachers in other sectors. Survey results also indicated that about a quarter of the teachers participating in the PDG were unclear about key eligibility rules for the recognition program. These results led the state to make significant changes in the way it communicated the program to early educators.

The RCT showed that the recognition program substantially reduced turnover among child care teachers.[19] During Virginia's 2020 Special Legislative Session, VDOE partners used these findings to advocate for, and ultimately receive, an additional $3 million in state funds to expand early educator incentives.

Partnership Benefits for Researchers

Giving policy makers relevant data on short timelines may help them make real-time decisions, leading to improved policies. Partnerships can also benefit applied researchers who want their research to improve on-the-ground practice. At the same time, the types of questions policy makers wish to tackle, and the timelines on which they need to tackle them, may not align with the questions typically asked in academic research, or the timelines needed for conducting rigorous evaluations. Researchers are accustomed to having enough time for thorough analyses that often focus on isolating the causal impact of particular interventions. In this way, partnership work can incur tradeoffs for researchers—between work that's relevant and timely on the ground, and work that's rigorous and contributes to scientific knowledge. We highlight these tensions throughout this issue. But our own experience, and that of the other authors, suggests that compared to independent research studies, research conducted through ECE partnerships can be more rigorous, more innovative, and more useful (that is, relevant) for furthering our collective knowledge.

Partnered research may have more mutual benefits than tradeoffs, a point too often unacknowledged. Conducting ECE research at scale is notoriously difficult. K–12 researchers have access to systemwide data collected over time on many aspects of teaching and learning, including data on student assessments, the workforce, measures of quality, and spending and finances. But the United States' fragmented ECE delivery system is characterized by a multitude of programs that are often poorly coordinated, resulting in a severe lack of systemwide data on any of these topics, and often little data even within particular ECE sectors (such as child care or Head Start).[20] For example, there are no data sets that track ECE teachers over time, either within or between centers. Data in other

key areas—such as the implementation of professional learning offered to teachers, or teacher practices that might result from professional learning experiences—is often quite limited as well.

Through partnerships, researchers and policy makers can find novel ways to overcome these challenges. For instance, policy makers may not realize that the data they collect through normal operations could provide new insights, and researchers may not realize that such data even exists. Partnerships can create opportunities to use existing administrative data to answer timely policy questions.

As an example, several years ago we learned that the data system used for LDOE's QRIS included the name of every lead teacher working in every publicly funded classroom serving toddlers or preschoolers in the state, in fall and spring, over multiple years. This information was collected during the classroom observations that are at the heart of the state's accountability system. The state simply considered this its QRIS data. But as we learned more about the state's data system, we realized it could be used to conduct the first systemwide examination of teacher turnover across a state. Because we can see which teachers are observed from one year to the next, we can provide much more systematic evidence on the prevalence of turnover, exploring differences by sector and the age of children served, and examining how turnover relates to quality.[21] Our team had long wanted to answer questions about ECE teacher turnover, but we lacked the systemwide data to do so. Through our partnership, we discovered an existing source of data that we could use to learn more about a pressing policy problem in ECE.

Besides uncovering lesser-known data sets, partnerships can also significantly improve the collection of new data. For instance, we've now collected thousands of surveys from early educators across Louisiana and Virginia, achieving much higher survey response rates (between 60 and 80 percent) than are typical for ECE workforce surveys.[22] Our partnership with LDOE , VDOE, and local ECE leaders contributed substantially to these high response rates. Our partners regularly advised us on how to make the survey language feel relevant to the local context. They also invited us to meetings with local ECE leaders (such as school principals and child care center directors) where they would introduce the project, give the leaders time to complete the surveys in person, and share hard copies for staff members who weren't comfortable taking surveys online. In addition, they sent out regular reminders to encourage survey completion, which likely resonated with local leaders more than emails from an out-of-state research team.

In the New York partnership, our RPP led to the creation of new tools to measure the implementation of a professional learning model to support children's art-based learning. And NYU researchers' experience with the Adapted Teaching Style Rating Scale—an observational measure of teacher practice that has been linked with professional learning in social-emotional learning programs—led the city to adapt and adopt this system in pre-K programs across the city, alongside other widely used measures of classroom quality.[23] The result is that new data on observed teacher behavior, which can be used for both quality improvement and future research, is being collected systematically in some pre-K sites in the NYC system.

Finally, RPPs can lead to studies that simply wouldn't have happened without regular conversations between teams. When policy makers launch new initiatives and reforms, their first thought isn't necessarily how to do so in a way that supports rigorous evaluation; in fact, they may not consider that at all. Researchers often learn about a policy too long after the fact, and lament the missed opportunity to incorporate a careful study into the rollout. Yet research can sometimes be embedded in a new initiative with relatively little added burden or cost to the policy makers, and it can provide valuable insights into the local initiative as well as broader questions in the field.

For instance, as part of its PDG grant, VDOE planned to pilot a recognition program for early educators, budgeting more than three million dollars for direct payments to those teachers. Early educators, especially those working in child care settings, are typically not well compensated; this is widely seen as a problem both for the teachers and for the children they serve. Through the recognition payments—up to $1,500 per teacher—VDOE hoped to support early educators and recognize their central role in improving ECE quality. Because the PDG grant had only a 12-month turnaround, our partners at the VDOE and Virginia Early Childhood Foundation focused immediately on the complex logistics needed to get the recognition program under way. The research team saw potential to provide unique evidence about the impact of increased pay on key ECE outcomes, particularly teacher turnover. Although low earnings are among the factors driving this turnover, no studies had actually tested the impact of pay increases on teacher stability in ECE.[24]

With this in mind, we advocated for a set of randomized controlled trials that would isolate the recognition program's effects. Our partners recognized the value of this and quickly agreed, having built up trust in our perspective through years of partnership. Thus we were able to jointly seize an opportunity that wouldn't have existed without the partnership; we now have strong evidence that financial incentives substantially reduce turnover among child care teachers.

A similar opportunity emerged in the NYC DECE partnership. The city decided to offer professional learning to teachers in separate tracks, each with a different theme or primary target, such as children's math skills, social-emotional development, or creative arts. These tracks have been the focus of the most recent phase of the partnered research activity (see the article in this issue by Abenavoli and colleagues). Thanks partly to the partnered discussions, we were able to consider strategies that prioritize NYC DECE objectives in honoring program leader's choices while also offering opportunities for rigorous research. As Abenavoli and colleagues recount in this issue, partnered work helped identify several opportunities to study professional development tracks that were embedded in the city's system of allocating pre-K sites to those tracks.

In our own experience, then, the choice to do applied partnered work hasn't come at the expense of rigor. Through our partnerships, we've been able to access and collect better, more representative data, to identify policy-relevant questions in real time, and to embed rigorous research opportunities into the rollout of policies we wouldn't otherwise have known about until years later, if at all.

Partnerships during COVID-19

With COVID-19, the need for fast turnaround in data collection, as well as for research focused on local needs, has become even more pronounced. Faced with the unprecedented challenges of a global pandemic, ECE policy makers and practitioners have had to rapidly figure out how to provide ECE. They had little evidence to guide their decisions on extremely difficult questions that called for immediate action: How do we ensure the safety of early educators? How do we do this while ensuring sufficient child care for essential workers? How do we ensure that young children—particularly those in poverty or facing other challenges—still receive the support they need? Researchers didn't have these answers either, but policy makers could turn to RPPs for advice on what the existing research tells us on closely related topics, and to identify ways to gather needed evidence.

Where trusted partnerships weren't already in place, it would have been too burdensome to build them in the midst of the rapid decision-making required by COVID-19. But where partnerships existed, researchers could use their familiarity with the context to offer resources that could be useful. For example, as COVID-19 spread in New York City, researchers quickly assembled materials about remote learning, including a tool kit for teachers citywide. They also offered resources to answer policy makers' most urgent questions (such as, What tools can help teachers care for children who've experienced a death in the family or a substantial loss of income—at scale, immediately?).

RPPs could also take advantage of existing infrastructure to expedite data collection. In both Virginia and Louisiana, our research teams quickly fielded workforce surveys to give policy partners data on which rapid response policies were perceived as most helpful, as well as concrete data on the experiences of early educators during a pandemic. Work is now under way in New York City to collect and analyze data about the professional learning experiences of leaders and teachers, to identify gaps and find innovative solutions for challenges that arise during this unprecedented time, and to guide recovery efforts.

The pandemic has highlighted the critical role of child care and the inadequacy of our fragmented ECE system, prompting urgent calls for major investments and reforms. New investments may generate unique opportunities to create more-effective ECE systems, and RPPs can shape policy makers' strategies for doing that.

Overview of the Issue

This issue of the *Future of Children* offers insights from successful RPPs tackling a wide range of topics related to early learning. The issue brings together many voices, including those of researchers, practitioners, policy makers, university deans, and foundation presidents. Each article and commentary focuses on a key aspect of partnership work—such as creating a shared research agenda, balancing priorities, and adapting to partner timelines. Together, the varied perspectives and case studies present the opportunities and challenges of the work, providing a fair review of this approach to research in ECE as it exists now, and offering a vision of where RPPs in ECE can go next.

In the opening article, Michael Strambler and colleagues describe one of the first steps in developing a research partnership: forming

a joint research agenda. Every RPP has its own origin story: sometimes researchers reach out to policy makers, at other times the reverse is true. In all cases, however, researchers and partners use an iterative process to identify a set of questions that both sides consider worth the investment. The authors describe three common challenges related to the agenda-setting process: responding to partner priorities, partners' research readiness, and researchers' content expertise.

Policy makers' priorities often shift rapidly in response to changing contextual factors, especially the leadership changes endemic to education. The authors show that to strengthen the partnership and its ability to withstand such personnel changes, researchers must take a flexible approach to agenda-setting and build relationships with multiple agency personnel. Another challenge related to agenda-setting is the policy makers' readiness to engage in the research process, with respect to both their understanding and buy-in for research, and to their capacity to support the research in key ways (such as data sharing). The authors describe how researchers can build this capacity over time by showing partners the concrete ways research can be useful, both in general and in a specific context. A third challenge relates to scenarios in which researchers lack expertise in the areas their partners view as priorities. Here the authors show how research partners can link their partners to resources and expertise they may not otherwise access.

Once a research agenda is set, RPPs often face questions related to the best ways to define, measure, and track key outcomes of interest. Sometimes relevant data have already been collected, and research partners only need to gain access to existing resources. But in many cases the necessary data don't yet exist, and a major activity of the partnership involves developing the tools needed to address partner priorities. In the next article, Amanda Williford and colleagues describe how they developed measurement tools in a partnership between the University of Virginia, the Virginia Department of Education, and an advocacy organization. The partners' main priority was to gain a more comprehensive understanding of Virginia children's skills upon entering kindergarten. They sought a tool that could serve varied purposes—outlining children's school readiness, guiding teachers' instructional practice, and helping the state target support.

The authors, including developmental psychologists and policy makers, describe the iterative development and rapid statewide implementation of the Virginia Kindergarten Readiness Program. They point to the benefits of partnering in this context, particularly the ability to create measurement tools that are methodologically sound and also address the needs and perspectives of early educators, school leaders, and policy makers. At the same time, the refinement and scale-up of the assessment—which moved from measurement development to full statewide implementation in just a few years—was fast-paced. The data collected through VKRP provide a population-level look at readiness across multiple domains, and therefore are uniquely suited for research. But at times, the immediate demands of the statewide implementation compromised the university partners' ability to pursue their research aims, at least in the short term—an outcome that highlights a recurring theme of the issue, the challenges of balancing researchers' and partners' needs.

Laurie Brotman and colleagues dive more deeply into the issue of systemwide scale-up. They describe a long-standing partnership between researchers at the NYU School of Medicine and NYC DECE around ParentCorps, an intervention aimed at supporting the parents and teachers of pre-K children. In 2014, DECE launched the city's Pre-K for All program, rapidly tripling the number of children in free, full-day pre-K. Two years later they rolled out ThriveNYC, a citywide mental health initiative with a focus on early childhood.

At the time these two initiatives started, the NYU School of Medicine team had already partnered with the state for nearly two decades, during which they developed ParentCorps, tested its efficacy, and refined it. They were thus well positioned to take on the citywide scale-up of their proven intervention. Nonetheless, they were challenged by the scope of the scale-up and the need to modify the intervention, which had been developed and tested in schools, for use in the community-based organizations that house many of the city's pre-K programs. The article describes how their long-established partnership principles (for example, ParentCorps's commitments to racial equity, centering parents' voices, and continuous learning) helped guide their actions, their strategy development, and ultimately their plan for scaling ParentCorps locally and nationally.

The next three articles tackle three common challenges arising in RPPs: competing timelines between researchers and their partners, tensions between analytic rigor and policy urgency, and insufficient research capacity. Christina Weiland and colleagues discuss the mismatch between the tight timelines of educational decision makers and the typically longer timelines of researchers who are pursuing rigorous analyses. The authors tell how, during a 12-year partnership with the Boston Public Schools Department of Early Childhood, they've balanced these competing demands by pursuing both fast-turnaround descriptive analyses to satisfy partners' immediate needs, and longer-term causal studies that ensure rigorous analyses on timely issues of interest both to the partners and to researchers more broadly. They also emphasize their commitment to communicating the strengths and limitations of fast-turnaround work in ways that can be easily understood by their partners and other key stakeholders.

Next, Rachel Abenavoli and colleagues argue that while some might see RPPs as being inherently at odds with scientific research, these partnerships actually facilitate the rigorous study of relevant policy questions in ways that would otherwise be infeasible. Growing numbers of randomized trials show us that early childhood intervention can yield sizable benefits. But when we move from relatively small, tightly controlled studies to scaled-up initiatives, the results are often disappointing. The authors describe how their partnership with New York City's Department of Education, as the city rapidly rolled out its universal pre-K initiative, gave them opportunities to collect experimental and quasi-experimental evidence while placing a minimal burden on educators. They argue that this type of research can answer the most pressing ECE questions, which are less about whether ECE can make a difference and more about the conditions under which early interventions are effective at scale.

Although many of the articles highlight the research that partnerships can and do

facilitate, effective partnered research requires capacity that many researchers and practitioners lack. In the next article, Maia Connors and colleagues discuss the issue of capacity constraints, using examples from their own "embedded partnership" between the program implementation and research teams within a single nonprofit organization (Start Early, formerly the Ounce of Prevention Fund). Even within a single organization, they write, effective partnership between practitioners and researchers can be challenging. What's needed is an organizational culture that values research evidence, sound measurement, and continuous learning; interdisciplinary human capital—people who bring diverse perspectives and a commitment to collaborative work; and sustainable infrastructure, including administrative support, technology, and financial resources.

These capacity-building challenges are compounded when creating a partnership across two or more agencies. University researchers, for instance, may find themselves in schools or departments where the incentives and culture favor more individualistic, single-discipline research projects. Practitioners and policy makers may not see the value of investing in research and evaluation, or they may hesitate to prioritize this work above more pressing operational and strategic aims. As described in this issue, some researchers and practitioners do manage to foster successful, longstanding, mutually beneficial partnerships, but those may be the exception rather than the rule.

In the final article, Jacqueline Jones examines how philanthropic organizations and other funding agencies can play a powerful role in facilitating partnerships, circumventing capacity challenges, and in turn supporting research that is more timely and more likely to impact policy decisions. She describes one unique partnership, the Early Childhood Research Network (ECRN), which brought together multiple New York City agencies and eight teams of researchers to explore a set of timely questions about the implementation of the city's universal early childhood program. Jones highlights how foundations can grease the wheels, so to speak, in the process of developing partnerships, making this form of research more feasible and accessible to a broader set of researchers, including early career scholars.

We close the issue with three commentaries that examine the roles that universities, research firms, and educational practice organizations (such as departments of education) can play in reducing these capacity constraints and creating institutional cultures that value evidence and partnerships.

Conclusions

Taken together, the articles and commentaries in this issue offer varied perspectives on the opportunities and challenges that partnership work affords. They also reflect on the next steps for universities, foundations, researchers, and policy makers who hope to promote this kind of scholarship. Calls for partnered research aren't new. Partnership work has a long history in education research, and it has anchored ECE research in particular. Our knowledge about the benefits of ECE for children, families, and society stems from close collaborations between researchers, policy makers, and practitioners. This may

be an especially important time for RPPs in ECE: we now have an extensive body of work documenting that ECE "works" alongside several decades of randomized experiments about ways to improve instruction and quality. In this context, RPPs offer a uniquely promising approach to the next phase of ECE research, supporting policy makers and practitioners as they wrestle with the complexities inherent in providing high-quality early learning in varied contexts and at scale. These topics are even more relevant in light of the devastation wrought by COVID-19 on children, families, early educators, and our society more broadly. The pandemic has created large gaps in the services provided to our youngest learners, and opened the door for new collaborations as policy systems race to meet children's needs. RPPs can support efforts to rebuild and reimagine ECE systems that can help all of our nation's children acquire strong foundations for kindergarten and beyond.

Endnotes

1. W. Steven Barnett, "Effectiveness of Early Educational Intervention," *Science* 333 (2011): 975–8, https://doi.org/10.1126/science.1204534; Frances A. Campbell et al., "Early Childhood Education: Young Adult Outcomes from the Abecedarian Project," *Applied Developmental Science* 6 (2002): 42–57, https://doi.org/10.1207/S1532480XADS0601_05; David Deming, "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start," *American Economic Journal: Applied Economics* 1, no. 3 (2009): 111–34, https://doi.org/10.1257/app.1.3.111; Deborah A. Phillips et al., *Puzzling It Out: The Current State of Scientific Knowledge on Pre-Kindergarten Effects* (Washington, DC: Brookings Institution, 2017).

2. W. Steven Barnett and Allison H. Friedman-Krauss, *State(s) of Head Start* (New Brunswick, NJ: National Institute for Early Education Research, 2016), https://nieer.org/research-report/states-head-start; Margaret Burchinal, "Measuring Early Care and Education Quality," *Child Development Perspectives* 12 (2018) 3–9, https://doi.org/10.1111/cdep.12260; Diane M. Early et al., "Teachers' Education, Classroom Quality, and Young Children's Academic Skills: Results from Seven Studies of Preschool Programs," *Child Development* 78 (2007): 558–80, https://doi.org/10.1111/j.1467-8624.2007.01014.x; Allison Friedman-Krauss et al., *The State of Preschool 2017: State of Preschool Yearbook* (New Brunswick, NJ: National Institute for Early Education Research, 2018), https://nieer.org/state-preschool-yearbooks/yearbook2017.

3. Christina Weiland, "Pivoting to the 'How': Moving Preschool Policy, Practice, and Research Forward," *Early Childhood Research Quarterly* 45 (2018): 188–92, https://doi.org/10.1016/j.ecresq.2018.02.017.

4. Cynthia E. Coburn and William R. Penuel, "Research–Practice Partnerships in Education: Outcomes, Dynamics, and Open Questions," *Educational Researcher* 45, no. 1 (January 1, 2016): 48–54, https://doi.org/10.3102/0013189X16631750.

5. Cynthia E. Coburn, William R. Penuel, and Kimberly E. Geil, *Research-Practice Partnerships: A Strategy for Leveraging Research for Educational Improvement in School Districts* (New York: William T. Grant Foundation, 2013), https://eric.ed.gov/?id=ED568396.

6. Clive R. Belfield et al., "The High/Scope Perry Preschool Program: Cost-Benefit Analysis Using Data from the Age-40 Followup," *Journal of Human Resources* 41 (2006): 162–90, https://doi.org/10.3368/jhr.XLI.1.162; Campbell et al., "Early Childhood Education"; Stevens H. Clarke and Frances A. Campbell, "Can Intervention Early Prevent Crime Later? The Abecedarian Project Compared with Other Programs," *Early Childhood Research Quarterly* 13 (1998): 319–43, https://doi.org/10.1016/S0885-2006(99)80042-8; James J. Heckman et al., "The Rate of Return to the HighScope Perry Preschool Program," *Journal of Public Economics* 94 (2010): 114–28, https://doi.org/10.1016/j.jpubeco.2009.11.001; Peter Muennig et al., "The Effect of an Early Education Program on Adult Health: The Carolina Abecedarian Project Randomized Controlled Trial," *American Journal of Public Health* 101 (2011): 512–16, https://doi.org/10.2105/AJPH.2010.200063; Milagros Nores et al., "Updating the Economic Impacts of the High/Scope Perry Preschool Program," *Educational Evaluation and Policy Analysis* 27 (2005): 245–61, https://doi.org/10.3102/01623737027003245; Arthur J. Reynolds et al., "Long-Term Effects of an Early Childhood Intervention on Educational Achievement and Juvenile Arrest: A 15-Year Follow-Up of Low-Income Children in Public Schools," *JAMA* 285 (2001): 2339–46, https://doi.org/10.1001/jama.285.18.2339; Arthur J. Reynolds et al., "Effects of a School-Based, Early Childhood Intervention on Adult Health and Well-Being: A 19-Year Follow-Up of Low-Income Families," *Archives of Pediatrics & Adolescent Medicine* 161 (2007): 730–39, https://doi.org/10.1001/archpedi.161.8.730; Arthur J. Reynolds, Suh-Ruu Ou, and Judy A. Temple, "A Multicomponent, Preschool to Third Grade Preventive Intervention and Educational Attainment at 35 Years of Age," *JAMA Pediatrics* 172 (2018): 247–56, https://doi.org/10.1001/jamapediatrics.2017.4673.

7. Edward Zigler and Sally J. Styfco, *The Hidden History of Head Start* (New York: Oxford University Press, 2010).

8. Kenneth Prewitt, Thomas A. Schwandt, and Miron L. Straf, eds., *Using Science as Evidence in Public Policy* (Washington, DC: National Academies Press, 2012), https://doi.org/10.17226/13460.
9. Institute of Education Sciences, "Research Practice Partnership Grants Program," 2020, https://ies.ed.gov/ncer/projects/program.asp?ProgID=81.
10. Spencer Foundation, "Research-Practice Partnerships: Collaborative Research for Educational Change," 2020, https://www.spencer.org/grant_types/research-practice-partnerships.
11. William T. Grant Foundation, "Research Practice Partnerships," 2020, https://rpp.wtgrantfoundation.org/.
12. National Network of Education Research-Practice Partnerships website, 2020, https://nnerpp.rice.edu/.
13. Caitlin C. Farrell et al., *A Descriptive Study of the IES Researcher–Practitioner Partnerships in Education Research Program: Final Report* (Boulder, CO: National Center for Research in Policy and Practice, 2018).
14. Ibid.
15. Mark Schneider, "Research-Practice Partnerships, Redux," Institute of Education Sciences, February 4, 2020, https://ies.ed.gov/director/remarks/2-4-2020.asp.
16. Michael Puma et al., *Head Start Impact Study: Final Report* (Washington, DC: Administration for Children & Families, 2010), https://eric.ed.gov/?id=ED507845; Phillips et al., *Puzzling It Out*.
17. Daphna Bassok, Katherine A. Magnuson, and Christina Weiland, "Memo: Building a Cohesive, High-Quality Early Childhood System," blog post, *Brown Center Chalkboard*, Brookings Institution, December 19, 2016, https://www.brookings.edu/blog/brown-center-chalkboard/2016/12/19/memo-building-a-cohesive-high-quality-early-childhood-system/; Institute of Medicine and National Research Council, *Transforming the Workforce for Children Birth through Age 8: A Unifying Foundation* (Washington, DC: National Academies Press, 2015).
18. Daphna Bassok et al., "Are Parents' Ratings and Satisfaction with Preschools Related to Program Features?," *AERA Open* 4, no. 1 (2018), https://doi.org/10.1177/2332858418759954; Anna J. Markowitz, Daphna Bassok, and Daniel Player, "Simplifying Quality Rating Systems in Early Childhood Education," *Children and Youth Services Review* 112 (2020): 104947, https://doi.org/10.1016/j.childyouth.2020.104947; Virginia E. Vitiello et al., "Measuring the Quality of Teacher–Child Interactions at Scale: Comparing Research-Based and State Observation Approaches," *Early Childhood Research Quarterly* 44 (2018): 161–9, https://doi.org/10.1016/j.ecresq.2018.03.003.
19. Daphna Bassok et al., "The Effects of Financial Incentives on Teacher Retention in Early Childhood Education: New Evidence from Virginia," paper presented at the annual fall research conference of the Association of Public Policy Analysis and Management, virtual, November 11–13, 2020.
20. Marcy Whitebook, Caitlin McLean, and Lea J. E. Austin, *The Workforce Data Deficit: Who It Harms and How It Can Be Overcome* (Berkeley, CA: Center for the Study of Child Care Employment, University of California, Berkeley, 2018).
21. Daphna Bassok et al., "New Evidence on Teacher Turnover in Early Childhood," *Educational Evaluation and Policy Analysis* 43, published online ahead of print, January 18, 2021, https://doi.org/10.3102/0162373720985340.
22. Daphna Bassok et al., *The Early Childhood Education Workforce in Louisiana: Findings from the 2018 Early Childhood Workforce Survey in Jefferson and Rapides Parishes* (Charlottesville, VA: University of Virginia, 2019); Daphna Bassok et al., *Who Teaches Virginia's Youngest Children? Sector Differences in the Racial/Ethnic Composition of Early Educators* (Charlottesville, VA: University of Virginia, 2020); Anna J. Markowitz et al., *Child Care Teachers' Experiences with COVID-19: Findings from the Study of Early Education in Louisiana* (Charlottesville, VA: University of Virginia, 2020).

23. Natalia M. Rojas et al., "Measuring Preschool Teachers' Social-Emotional Practices: A Comparison of Two Measures," *Early Education and Development* 32, published online ahead of print, January 11, 2021, https://doi.org/10.1080/10409289.2020.1864838; Pamela Morris et al., *Impact Findings from the Head Start CARES Demonstration: National Evaluation of Three Approaches to Improving Preschoolers' Social and Emotional Competence: Executive Summary* (Washington, DC: MDRC, 2014), https://eric.ed.gov/?id=ED546649.

24. Claudia Hale-Jinks and Herman Knopf, "Tackling Teacher Turnover in Child Care: Understanding Causes and Consequences, Identifying Solutions," *Childhood Education* 82 (2006): 219–26, https://doi.org/10.1080/00094056.2006.10522826; Marcy Whitebook, Deborah Phillips, and Carollee Howes, *Worthy Work, STILL Unlivable Wages: The Early Childhood Workforce 25 Years after the National Child Care Staffing Study* (Berkeley, CA: Center for the Study of Child Care Employment, University of California, Berkeley, 2014); Marcy Whitebook and Laura Sakai, "Turnover Begets Turnover: An Examination of Job and Occupational Instability among Child Care Center Staff," *Early Childhood Research Quarterly* 18 (2003): 273–93, https://doi.org/10.1016/S0885-2006(03)00040-1.

Seeking Questions from the Field: Connecticut Partnership for Early Education Research

Michael J. Strambler, Joanna L. Meyer, Clare Waterman Irwin, and George A. Coleman

Summary

One of the first steps in developing a research-practice partnership is forming a joint research agenda. In this article, Michael Strambler, Joanna L. Meyer, Clare Waterman Irwin, and George A. Coleman describe the collaborative process that Connecticut's Partnership for Early Education Research (PEER) used to develop an agenda driven by practitioners' interests and concerns.

The authors describe three challenges that often arise during the agenda-setting process: responding to partners' priorities, partners' research readiness, and the researchers' content expertise. Policy makers' priorities often shift rapidly in response to changing contextual factors, especially the leadership changes endemic to education. The authors show that to strengthen the partnership and its ability to withstand such personnel changes, researchers must take a flexible approach to agenda-setting and build relationships with multiple agency personnel. Another challenge is the policy makers' readiness to engage in the research process, with respect both to their understanding and buy-in for research, and to their capacity to support the research in key ways (such as data sharing). The authors describe how researchers can build this capacity over time by showing partners the concrete ways research can be useful, both in general and in a specific context. A third challenge is that researchers may lack expertise in the areas their partners view as priorities. The authors show how research partners can link their partners to resources and expertise they may not otherwise be able to access.

www.futureofchildren.org

Michael J. Strambler is an associate professor of psychiatry at Yale School of Medicine and the director of Child Wellbeing and Education Research at the Consultation Center at Yale. Joanna L. Meyer is the assistant director of Child Wellbeing and Education Research at the Consultation Center at Yale. Clare Waterman Irwin is a research scientist at Education Development Center. George A. Coleman is an early childhood program development specialist at Cooperative Educational Services. Together, Strambler, Meyer, Waterman Irwin, and Coleman make up the founding management team of PEER.

A main goal of research-practice partnerships is to conduct research, driven by researchers, educators, and policymakers, to seek out new information to improve students' learning. As this issue of the *Future of Children* illustrates, research-practice partnerships (RPPs) take various approaches to define their mission and establish an agenda to accomplish their goal. This article describes the decision points, processes, benefits, and challenges for establishing a collaborative research agenda, with an emphasis on agendas that are driven by practitioners' interests and concerns. Three main challenges are inherent in this collaborative approach: responding to partners' priorities, partners' readiness for research, and the researchers' content expertise. We explore these topics in the context of a Connecticut-based early childhood partnership, the Partnership for Early Education Research (PEER). Using examples from PEER's work, we discuss strategies for addressing the above issues, and we examine the benefits of using a collaborative approach to establish an RPP's research agenda.

Background

The Connecticut partnership PEER conducts research to inform early childhood education policy and practice. Specifically, PEER strives to produce research evidence that can be used to improve access to high-quality early childhood education and to reduce educational disparities among young children, both locally and statewide. A core aspect of PEER's approach is to pursue questions developed in collaboration with its members, thus ensuring that the research is relevant to early childhood teachers, administrators, policy makers, and advocates.

To foreground both researchers' and practitioners' voices, PEER is led by the Yale School of Medicine, Education Development Center (EDC), and Cooperative Educational Services (CES). Both the Yale School of Medicine and EDC (a nonprofit research firm focused on education and health) contribute research expertise in education, psychology, statistical methodology, and interdisciplinary fields. CES, one of Connecticut's six legislatively authorized Regional Education Service Centers (RESCs), contributes its longstanding connections to the region's public schools and early childhood education programs as a trusted provider of professional development, as well as direct experience from its own early childhood and elementary programs.

PEER was formed in 2014 in response to the lead organizations' shared interest in collaborative research to help improve early childhood education for Connecticut. PEER's founders were also interested in increasing coordination between early childhood education programs and public school systems. Many Connecticut school districts operate a preschool program, but many preschoolers and the vast majority of infants and toddlers are served by a diverse array of community-based providers, including early childhood education centers and family childcare providers. School districts and community-based providers generally operate independently, with no structures for coordinating learning goals, instructional practices, professional development, information sharing, or

children's transition to the public school system. PEER's founders believed that an RPP focused on the largest urban centers in the CES region (Bridgeport, Norwalk, and Stamford) would offer a promising approach for using research and evidence to enhance regional and state-level early childhood education.

The Institute of Education Sciences' (IES) Researcher-Practitioner Partnership grant program provided the initial funding to launch PEER. The IES grant supported three specific research projects: on early childhood assessment, kindergarten readiness, and the association between teacher and classroom factors and kindergarten performance. The grant also helped build the partnership and establish a long-term research agenda. PEER has since relied on funding from the Spencer Foundation and a variety of contracts with its partners to pursue that long-term agenda.

Partnership Structure and Theory

Like many RPPs, PEER was designed as a long-term, place-based partnership focusing on problems of policy and practice, and committed to mutualism. According to a common classification of RPPs, it best fits the description of a research alliance.[1] Rather than aiming to solve specific problems by a repeated cycle of evaluation and design, or to focus on a single problem of practice, PEER seeks to support its partner organizations by studying a range of policy- and practice-related questions.

Figure 1. PEER's Theory of Change

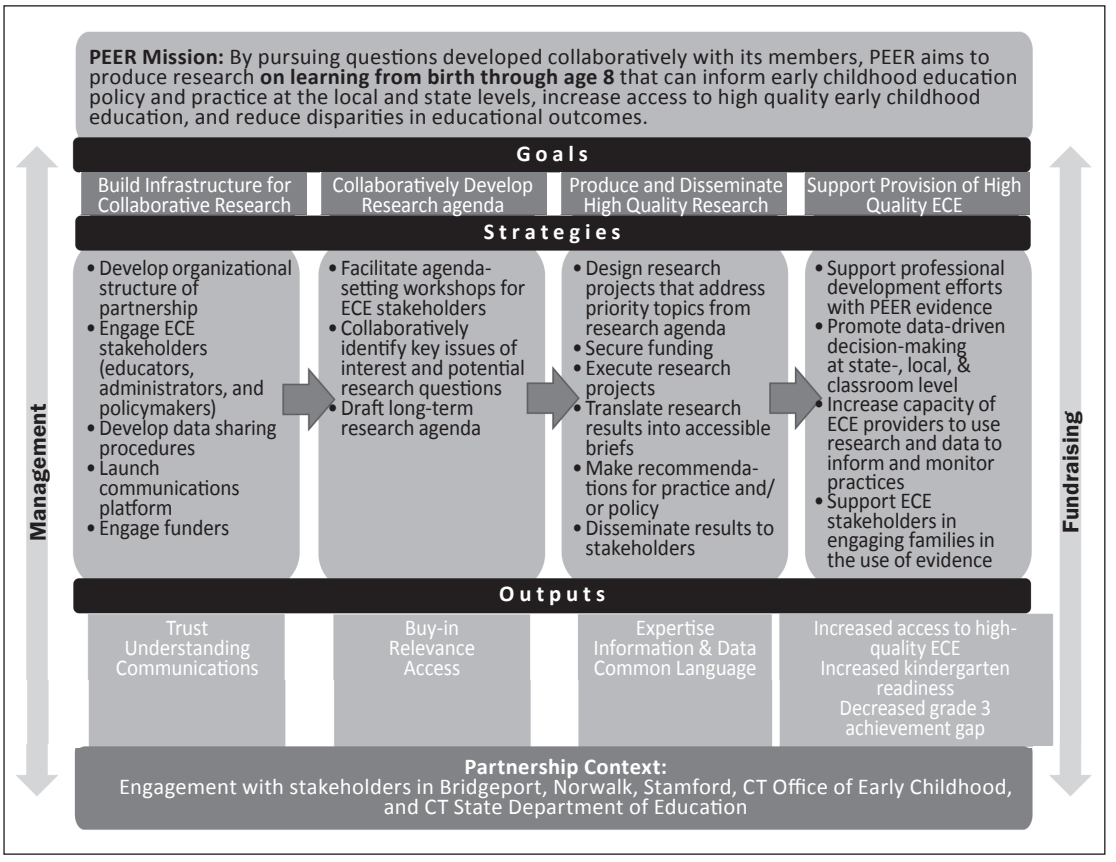
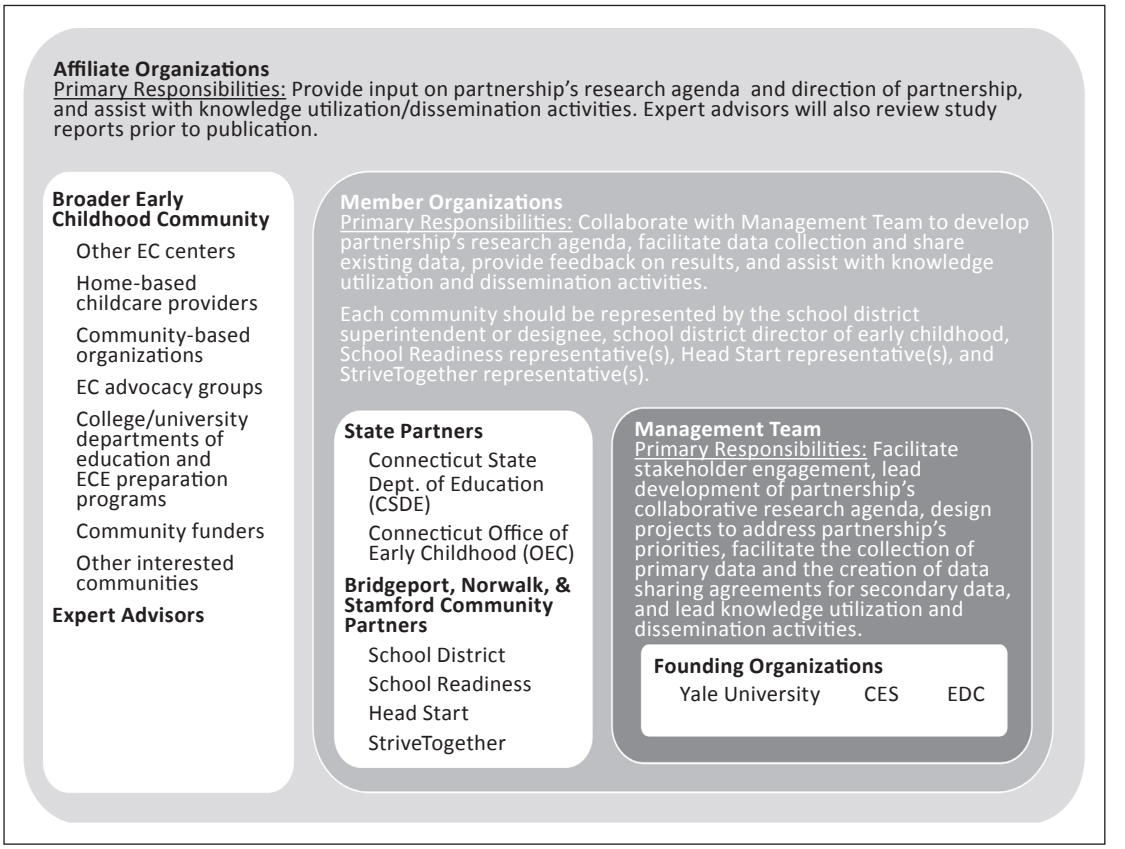


Figure 2. Structure of PEER



Although the member organizations are deeply involved in establishing PEER's research priorities and questions, as well as in using and disseminating the study results, PEER researchers work independently to conduct research before they share findings with member organizations. This structure promotes both research integrity and objectivity.

As PEER's theory of change (figure 1) shows, building processes for engaging a diverse set of partners allows the partnership to collaborate on developing a research agenda. The partnership can then produce and disseminate rigorous, actionable research, and support the application of research to policy and practice in its partner communities. Rather than establishing a linear process, PEER envisions a cycle of collaboration, research, dissemination, and application. Sustaining the cycle requires strategic and logistical management, along with grant-writing and other forms of fundraising.

Figure 2 illustrates PEER's structure. The management team (Yale, EDC, and CES) oversees all aspects of partnership-building, research, and dissemination. Member organizations include two state agencies and each community's school district, Head Start program, and StriveTogether initiative (described below), as well as other community-based providers of early childhood education. These organizations collaborate with the management team to define and refine the

partnership's research agenda, guide the development of research projects, collect data and share existing data, give feedback on reports, and design activities to use and disseminate knowledge.

Affiliate organizations also play a role. These are smaller community-based providers of early childhood education, home-based childcare providers, community-based organizations that serve children and families, early childhood advocates, college and university departments of education and early childhood education teacher preparation programs, and representatives of other Connecticut communities that want to learn more about the partnership. Affiliate members provide input on the general direction of the partnership; they also use and help disseminate findings. PEER consults expert advisers on specific projects as well. Including such a range of voices has many benefits, which we describe later in this article.

State and Local Context

Before they reach kindergarten, Connecticut children experience a variety of early care and education settings. These are supported by various federal, state, municipal, and private funding sources, which sometimes overlap. Early childhood education programs funded by multiple sources are subject to the requirements of each funder; these may set parameters for staffing, the physical condition of the site, learning experiences, children's eligibility, and other program elements. Such diversity presents a significant management challenge for early childhood education, and can also make it difficult for K–12 systems to coordinate with these programs.

Because of the challenges faced by Connecticut cities and CES's central role in launching the partnership, PEER covers the three largest urban communities in southwestern Connecticut, the region served by CES, which includes three of the state's six largest cities: Bridgeport (pop. 145,000), Norwalk (89,000), and Stamford (130,000). Children of color make up a substantially higher percentage of these cities' school district populations than of the state's public school population as a whole (between 71 percent and 87 percent, compared with the state average of 43 percent). All three districts enroll a much higher proportion of English language learner students than the state average (between 13 percent and 16 percent, compared to the state average of 7 percent).[2]

Leadership transitions have affected each community's level of participation in PEER: each school district has welcomed at least one new superintendent since PEER was launched, and Norwalk has seen three Head Start providers. Finally, each PEER community hosts a collective impact initiative that's affiliated with the StriveTogether network. These initiatives aim to organize communities to work collaboratively on complex social problems that limit opportunities for youth and adults alike. The StriveTogether model seeks to do this by helping communities bring local partners together to assess needs, define goals and action plans, secure resources, and evaluate impacts.[3]

Establishing a Research Agenda

Typically, an RPP's research agenda outlines goals for the next three to five years or more. Research agendas center the work on a specific set of topics or focus areas,

and describe a cohesive set of key questions regarding practice or policy for which practitioner partners need answers. The research questions should originate with and have significance for the practitioner partners, with input from researchers.

Without a representative set of partners, RPPs risk focusing on topics and questions that aren't important to practice or policy, thereby undermining their core mission.

RPPs have found a variety of ways to establish a successful research agenda. Because RPPs often develop out of unique circumstances, the partnership will be most effective when partners consider the context while developing their agenda. For example, the researchers who initiated the ParentCorps partnership (see the article in this issue by Laurie Brotman and colleagues) wanted to study a researcher-designed intervention, and they set their research agenda around that specific need. By comparison, the Boston Public Schools' early childhood research partnership (see the article in this issue by Christine Weiland and colleagues) began with broad research questions identified by the school district; the researchers recommended study designs that could be used to answer those questions. Alternatively, as with PEER, researchers and practitioners may connect before doing any research and decide on the RPP's parameters. After agreeing to work together around a broad topic (for example, closing achievement gaps via early childhood education), the partners then collaborate to define their research agenda.

Creating a research agenda together at the outset works well for partnerships that take the form of a research alliance, because it allows many voices to participate. In this process, partners think deeply about their challenges and needs, contribute ideas about what to focus on, prioritize those ideas, and, finally, create a set of researchable questions, targeting the ideas that were generated. Though this process is led by researchers, it ensures that the topics and questions are relevant for policy makers and, especially, practitioners.

For its agenda-setting process, PEER adapted workshop materials from the Regional Educational Laboratory (REL) Northeast and Islands that were created to help RPPs develop their research agendas by drawing out diverse perspectives.[3] These materials guide partnerships through a series of steps: ensuring that all partners understand the possible approaches to the research; generating research topics; prioritizing, categorizing, and voting on the most relevant topics; and generating and refining researchable questions.

Who Should Be Involved?

An essential part of establishing a research agenda is to ensure that the appropriate people are involved. Without a representative set of partners, RPPs risk focusing on topics and questions that aren't important to practice or policy, thereby undermining their core mission. Perhaps worse, failing to involve people who can represent their organization's needs and priorities may produce an agenda that's not important to those with the authority to commit to the research.

The first step in determining who should be involved is to define the structure of the RPP. Will it contain one practitioner entity, such as a single school district, or will there be multiple practitioner partners? There can be tradeoffs either way. For example, PEER decided to partner with a variety of member organizations in part because of the nature of early childhood education—a fragmented system of school- and community-based providers. Also, we aimed to conduct research that would be valuable not only to the partners but also to the state and the field of education as a whole. It would have been hard to accomplish our goals by working with one relatively small community. However, while we benefited from the increased representation and the greater generalizability of our findings, we faced challenges in managing the partnership, which we'll discuss below.

When the partners in an RPP represent a variety of organizations and, in particular, when those organizations approach the work from different perspectives (for example, district- and community-based early childhood education providers with different philosophies of early learning), it's important to build relationships and identify common interests between the researchers and each of the partner organizations, even before convening the entire group to set the agenda. It's essential to keep in mind the adage "Progress moves at the speed of trust." Though working across many organizations can be hard, there are benefits as well; for example, partners get the opportunity to think critically with other professionals who are engaged in similar work. In this way, participating in an RPP has some of the same benefits for practitioners as a professional learning community.

Defining Shared Objectives

The process of setting a research agenda is a critical time to ensure that practitioner partners understand different kinds of research (for example, reviews of research studies, meta-analyses, primary data collection, secondary data analysis) and research goals (to add to existing knowledge, to determine whether an intervention is effective, or to guide practice or policy?). If the partnership intends to create a cohesive agenda with questions to study over the course of several years, it also helps to discuss what constitutes a "researchable" question. However, it's generally up to the researchers to refine the questions generated by practitioner partners to ensure they're specific, measurable, attainable, relevant, and time-bound (condensed into the acronym SMART). For example, here's one SMART question that was generated by PEER's research agenda–setting process: "What are the characteristics of successful coordinated systems among public schools, families, and community-based programs that have served young children over the past 10 years?" This question is *specific* because it targets a specific population. Though the question doesn't define "successful," the concept of success is *measurable* and the data for doing so are *attainable*, either by collecting them or by acquiring them from existing sources. Finally, the question is *relevant* to a large portion of the partner communities and *time-bound* because it describes a specific timeframe. The research agenda–setting process is also a good time to establish partnership norms, including how often to revisit and revise the agenda.

Pursuing the Collaborative Research Agenda

The research agenda defines the partnership's direction, and knowing the key questions of practice and topic areas that the partners want to address allows the partnership to seek funding that supports their shared interests and needs. By gathering input on immediate challenges and by including partners in the proposal process, the RPP can ensure that the proposed work meets important needs. After developing PEER's research agenda, the management team met with a number of practitioner partners to get feedback about which topics to work on first. The partners named supporting dual language learners as a top priority, because Connecticut's growing diversity had created high need. But resources were inadequate, so the management team worked with the partners to craft a proposal that would provide actionable information about this topic.

Beyond developing the research topics and questions, practitioner partners must be included in the process of designing and conducting research. This helps to avoid losing momentum, and also keeps partners feeling they're part of the work. It also keeps the research agenda grounded in the practitioner partners' interests. Such concerns are often addressed by placing partnership members in advisory committees for the purpose of advising the entire RPP, specific projects, or both. PEER opted to create project-specific committees that advise on such activities as data collection (primary and secondary), identifying interest groups that should be apprised of the work, troubleshooting, and recommending resources to make the project a success. During the course of a project, the advisory committee also meets with the researchers to help them interpret findings from the analyses and identify the implications of those findings.

Including practitioner partners in decision-making at each step of the research process maximizes a project's relevance to the partners and increases the likelihood that they'll use the findings to guide their practice and policy decisions. That's particularly important in early childhood education, which encompasses many different systems that provide services for children. These systems often function independently, making it hard to conduct research that's relevant to all of them. PEER includes a range of early childhood partners across different systems and engages them at each step of the research process to make sure the research is relevant to them—and is used to improve the services provided to children and, in turn, to improve children's educational outcomes.

Challenges and Responses

It's almost certain that RPPs driven by practitioner input will encounter challenges that threaten their effectiveness. Some of these challenges are unique to practitioner-driven partnerships, while others are common across most RPPs but manifest differently in practitioner-driven partnerships. In this section we highlight some of these challenges, show how they've emerged in the context of PEER, and describe possible responses. We focus on three broad types of challenges: responding to partners' priorities, partners' research readiness, and content expertise.

Responding to Partners' Priorities

Even when RPPs invest substantial time to identify the priorities of their partner

organizations, it can be difficult to translate those priorities into research questions that are of equal interest to all the partners. Partners' priorities may differ and can also shift quite quickly, particularly when there's a leadership change in an organization. These differences can make it hard to keep partners interested in work that they may have previously identified as being valuable to them—and can also make it hard to design future studies. Here we describe the challenges in terms of identifying *shared priorities* and responding to *shifting priorities*.

Shared priorities. When a partnership works with organizations across multiple communities, demographic differences can affect each community's priorities and the types of questions they want to ask. Socioeconomic and racial/ethnic diversity can be especially influential. In the case of PEER, these differences are substantial, especially between Bridgeport and the other two communities. Though Bridgeport, Norwalk, and Stamford are all in one of the richest counties in the United Sates, their populations are quite different, and they face very different challenges. In Bridgeport, for example, 32.9 percent of children under 18 live in poverty, compared with 11 percent in Norwalk and 10 percent in Stamford (and 14 percent in Connecticut overall). Although children of color make up 70 percent or more of the school-age populations in all three cities, more than half the overall populations of Norwalk and Stamford identify as white and not Hispanic or Latino, compared with only 23 percent in Bridgeport and 78 percent in Connecticut overall.[4] These differences shape the questions partners ask. Consider achievement gaps: in a city with less socioeconomic and racial diversity, like Bridgeport, partners tend to frame achievement gaps in terms of differences relative to other communities, to the region, or to the state as a whole. But communities with greater diversity, like Norwalk and Stamford, tend to frame achievement gaps as group differences within their communities. It's important to be aware that when such topics arise, practitioner partners may think about them in different ways.

Another factor that affects the agenda-setting process is the differences in education landscapes among communities. This is especially important for RPPs that focus on early childhood, because early childhood education is highly varied and complex. States vary, for example, in the type and availability of funding for preschool, often having multiple and overlapping forms of funding. And with so many distinct early childhood education providers in each community, it's hard to get them all together to define shared priorities. For this reason, PEER typically focuses on the largest providers in each community.

As time passes, RPPs will inevitably see priorities change—and change quickly, in some cases.

The level of participation in early childhood education is another difference that can arise. For example, in the past few years Bridgeport has struggled to fill all available preschool spots because a large number of families have chosen not to participate. So for Bridgeport, a key priority might be encouraging parents to enroll their children in preschool, whereas Norwalk and Stamford might be more

concerned with improving their existing programs.

As the partners gained an understanding of the types of research PEER might conduct to address their needs, differences in research priorities emerged. In 2017, partners began asking whether PEER could conduct research on practices or policies being implemented in their own organizations or communities. Though these inquiries were aligned with the topic areas described in PEER's research agenda, the specific policies and practices weren't shared across the three PEER communities. This challenged our management team to think beyond research questions common to all partners and to better respond to partner-specific concerns.

Shifting Priorities. Collaborating to address the shared priorities of partner organizations is the first step toward establishing a practitioner-driven research agenda. But as time passes, RPPs will inevitably see those priorities change—and change quite quickly, in some cases. Although practitioners' interests are influenced by their own views on problems of practice, they can't ignore the demands placed on them by parents, municipalities, states, the federal government, a natural disaster, or a public health crisis. New state policies may be implemented, new school or district leaders may arrive with their own strategic goals, and parent and community needs may evolve. These changes can ultimately alter what practitioners want or need to know. Such shifts in research priorities can be challenging for RPPs that have invested time and resources in developing long-term research agendas.

RPPs should anticipate leadership changes, which are common in urban districts.[5] PEER's communities have been no exception. As noted above, the three school districts in PEER communities all saw at least one leadership transition during the RPP's first three years. The timing of these transitions was particularly difficult for PEER, because the superintendents who were in place when the research agenda was developed had left their positions by the time PEER launched the first project based on that agenda. In two of the districts, the superintendent changed between the submission of a major grant proposal and its acceptance by the funder. The PEER management team worked to build relationships with the new superintendents, meeting with them to share PEER's mission, history, and research agenda; to tell them about the newly funded project; and to ask them to identify which areas might benefit from research support. Some new superintendents were more receptive than others to research projects that had been launched before their tenure began.

Addressing challenges related to priorities. As the differences in partners' priorities became apparent, and as partners' priorities shifted, the management team explored ways to adjust their approach. The version of the research agenda released by PEER in 2016 included specific research questions developed during the second agenda-setting workshop. But it became apparent that the depth of detail made the research agenda less responsive to the partners' evolving needs and priorities, especially given the differences among the communities. And many of the research questions were directed toward reviews of past studies, which was limiting when it came to seeking funding for PEER's research.

For these reasons, in 2018 the management team decided to reframe the agenda so it focused more generally on the four topic areas and less on specific questions. In this sense, the research agenda became more of a frame than a set of specific research questions, while staying true to the input provided by the partners. Other RPPs that aim to respond to practitioners' interests may find it best to develop a research agenda that's specific enough to focus the partnership, but also flexible enough to adapt to changes in priorities and interests.

The research agenda may also need flexibility to evolve alongside the practitioner partners' knowledge of collaborative research. As we've said, PEER received multiple partner-specific requests once our partners began to understand how PEER might support their organizations with research and evaluation. So the management team started exploring how PEER could respond to such requests, both immediately and in the long term. This led to two changes in our approach. First, rather than assuming that all of PEER's projects would engage early childhood partners in all three communities, we accepted that some research topics might be more relevant to one partner than another. Thus, we agreed to consider research projects that focused on the needs of specific partners. Second, the management team expanded its focus to include evaluation and technical assistance, which allowed PEER to address partners' immediate and unique needs more effectively. The support included helping partners develop a theory of change, developing formative tools to guide decision-making, and evaluating programs' effectiveness. By necessity, these services are currently limited to partners with the funding to support such work. But we hope to acquire core funding so we can provide this support at no cost or reduced cost in the future.

One example of the projects taken on by PEER since this shift occurred is our work with a recently launched Science, Technology, Engineering, and Mathematics (STEM) preschool to develop its theory of change; to decide on appropriate teacher, student, and implementation measures; and to create an evaluation plan. Another project involved helping a community understand the usefulness of a new communication tool intended to share student information across the preschool-to-kindergarten transition. For this project, PEER worked with a group of community partners to collect data on how effectively the information gathered with this tool by preschool teachers was communicated to kindergarten teachers and elementary administrators, and how useful these partners found the information for guiding practice. More recently, when most Connecticut early care and education programs closed at the onset of the COVID-19 pandemic, one of our partners asked for support concerning the new challenges faced by families and teachers. We worked with this partner to develop a needs assessment survey and to collect data that could guide the organization as it pivoted to new ways of supporting its families and teachers while its facility was closed to children.

When it comes to shifting priorities, leadership transitions have convinced us that we must engage district leaders beyond the superintendent—for example, mid-level leaders such as directors of research and directors of early childhood. We've found that new superintendents are more

enthusiastic about remaining engaged with PEER when the remaining district leaders havebeen involved with the process of setting our research agenda. These leaders could tell the new superintendent what the district had already done, explain the benefits of collaborative research, and advocate for participation to continue. In districts where mid-level leaders were less involved, or where there were leadership transitions across many levels, it was harder to engage the new superintendents.

Research Readiness

Some partner organizations may be more ready than others to participate in research, in terms of both mindset and resources. *Mindset readiness* refers to attitudes and beliefs that are conducive to engaging in research. Here, we highlight attitudes regarding the rigor and timeliness of research. *Resource readiness* means the partners' capacity to get involved in the process of research. The most important resource is personnel, but technological systems for managing, reporting, and sharing data are also relevant.

Mindset Readiness. Although researchers tend to believe deeply in the value of research, practitioners may be less confident about its applicability and relevance. In the early stages of an RPP, it's common to encounter a range of mindsets about the utility of research and evidence for educational practice. These views directly affect partners' preparedness to work with the research partnership. Practitioners on one end of the spectrum believe that personal or professional observation or experience provides the best evidence for the presence of an effect or the effectiveness of a practice. From their point of view, personally witnessing or directly experiencing a phenomenon carries more weight than data produced by research tools like assessments and surveys. (Researchers, on the other hand, tend to view personal observation as useful for generating hypotheses but weak for causal inference.) For an individual with this mindset, producing evidence via rigorous or systematic methods is valuable only insofar as it can "prove" what the practitioner already believes. Such practitioners consider that the real value of evidence lies in convincing funders, administrators, and others to whom they're accountable that a specific approach has a beneficial effect.

The belief that practice—the instruction of children—is the real work of educators is perhaps one of the most challenging views encountered when working with practitioners to develop and execute a collaborative research agenda. This belief implies that while research may be useful, it should take a back seat to instruction; in other words, research is a separate and expendable aspect of serving students. This perspective is especially challenging because it's partly true. In the under-resourced field of early childhood education, when leaders perceive a need to choose between educating children or conducting research, it makes sense to choose educating children. But this either/or framing is mostly driven by the belief that research is resource- and time-intensive for practitioners. Though that may be true in some cases, research often uses existing data to produce actionable findings, which minimizes the burden on practitioners. PEER partners have often told us they have plenty of data but simply lack the research capacity to use the information. In these instances, research can occur without disrupting classroom practices, though coordination at the leadership level

is likely required. For projects that do require collecting new data, RPPs can work with practitioners to design feasible studies that use available resources and minimize disruption to teachers and students.

Another challenge occurs when practitioners lack incentives to use research to guide their work. They may adopt a "good enough" mindset, assuming that quasi-experimental and experimental research is unnecessary. This perspective also implies that scientific rigor has its place in such contexts as academia or medicine but is superfluous for early childhood education.

Researchers sometimes see the mindsets we've just described as being opposed to rigorous research. In our view, practitioners who think this way aren't rejecting such research altogether; rather, they're not aware or convinced of its importance and value. And researchers themselves often undervalue descriptive analyses in their work. In many cases, the research most useful to educators includes in-depth descriptive analyses that can help them understand who they're serving, who's working in their programs, or the characteristics of those programs.

Where rigorous research is not the norm, education researchers can easily underestimate the effort required to get practitioners to buy in. This challenge can be magnified when working with early childhood providers, for at least two reasons. First, early childhood settings, especially private programs, often function independently of one another—more so than schools, which are part of a large and organized system. Accordingly, researchers working across multiple early childhood settings will rarely find a common set of leaders with a shared mindset regarding research. Instead, RPPs that focus on early education are likely to include multiple entities with varying understanding of, appreciation for, and capacity to engage in research. This diversity of viewpoints makes it hard to resolve decisions about research directions.

Second, whereas public school systems often have common practices that specify which assessments are administered and when, many states and communities lack such a system for early childhood. For example, Connecticut has a history of common measures for formative assessment in preschools, but no statewide summative assessment. Without common assessments for measuring preschool children's academic and social-emotional outcomes, it's hard to conduct collaborative research across multiple communities without collecting primary data. If, say, a partnership wants to know how classroom quality and instructional supports are associated with the development of preschool children's skills, researchers would need to either administer an assessment across the participating sites or figure out how to compare results from the various summative assessments implemented by the preschools. The former requires a great deal of resources and may burden preschool staff and children. The latter makes it harder to statistically answer research questions or to reach conclusions that apply to all settings.

Resource Readiness. Even when practitioners want to participate in research, they may lack the organizational resources to do so effectively. Although most RPPs aim to increase research capacity, collaborative research requires significant support from schools or districts. For example, if researchers and practitioners are to share data with each other, the researchers

need collaborators at the school, district, or organization level so they can develop data-sharing agreements, understand the nature and availability of data elements, and determine the means and schedule for transferring data. The availability of such collaborators can vary widely. Some school districts and state agencies have well-staffed research departments; in others, the research or data teams lack adequate resources. And community-based organizations may have no research teams at all. Moreover, some organizations are better prepared to partner because they've developed structures that can facilitate collaboration. For example, one PEER school district has an administrator responsible for facilitating and managing the district's partnerships. Some school districts establish similar structures by creating professor-in-residence positions that explicitly link universities and school districts. No matter how motivated a practitioner organization may be, participating in research is challenging without the necessary organizational supports. The variation in resources is a barrier for researchers who are trying to collaborate with diverse partners to perform timely, rigorous, and useful research.

Fostering a culture that values data and research is especially important while developing a research agenda, because it encourages buy-in and creates momentum for the work ahead.

Responses to readiness challenges. Because practitioner partners hold different views about research, PEER has worked to show partners how research can help them serve children better. Being committed to fostering a culture that values data and research is especially important while developing a research agenda, because it encourages buy-in and creates momentum for the work ahead. PEER has built such a culture through an "informing and doing" approach. For the "informing" part, we communicate what it means to conduct research and explain its fundamental value. This includes the instructional elements of the agenda-setting workshops previously described, as well as the production and dissemination of bite-size, practitioner-friendly articles related to collaborative research.

For example, PEER's newsletter and website have been important media for informing early childhood partners. Every month we publish a brief article on a topic relevant to our partners. Some articles feature PEER's research or other research related to early childhood; others cover more general topics in education research, such as why evidence matters in helping children succeed in school, the value of logic models in education, and using assessments wisely in early education settings. We can then refer practitioners to these articles as research primers.

We also hold various partnership meetings throughout the year to foster common views around the value of research and its role in practice. The most frequent meetings gather the partners with whom we're actively working; we may meet as often as four times per year to seek feedback and share progress and results. We also hold annual meetings with a wider cross-section of partners, where we discuss broader issues such as the direction of the partnership and hear from

a keynote speaker with expertise in RPPs. In 2016, PEER hosted a conference for the Connecticut early childhood community, with workshops where community and state partners shared their research and its implications for practice and policy.

PEER has also tackled the issue of research readiness by "doing." As we began to create research products, it became easier for the partners to appreciate the value of our shared work. In a few instances, we've offered free research and evaluation advice to partners when we saw that doing so would help build capacity relative to the time commitment. For several years PEER has advised StriveTogether groups that are leading community efforts focused on early childhood. Since the collective impact model emphasizes data, evidence, and assessing change and improvement, partnering with each PEER community's StriveTogether group fits with our mission of using research to inform practice. By advising cross-sector groups that work with children from birth to age eight, PEER helps strengthen data collection, data sharing, and evaluation in a way that helps organizations with limited capacity to conduct their own research. Though not every community served by an RPP will have such robust community-wide efforts, many communities have organizations with similar goals, if not a similar scope of work. Allocating advisory time for RPP team members does have a real cost, but the investment can be worthwhile.

Content Expertise

In practitioner-driven partnerships, the partners' research priorities can lead researchers into content domains outside their main area of expertise. When a partner has already developed a theory of change and done some of the deep thinking around conceptualizing constructs of interest, a research team may be able to use that partner's methodological, data analytic, and evaluation skills to perform research in a new area. But in the early stages of research, it's more common for practitioners to request help in performing such tasks as identifying what works best in a specific content area or for a specific population. In these cases, the research team is more effective when it has in-house expertise in the relevant content. Yet it's unlikely that the research team will always have expertise that's relevant to every research question identified by the partnership.

Responses to the content expertise challenge. RPPs that aim to base their research agendas on the needs and interests of practitioner partners must consider any potential gaps in expertise when designing the structure of the partnership, and then start developing resources to address the issue. For example, PEER's management and organizational structure allows us to tap into expertise within our organizations and across various levels of the partnership—practitioners, policymakers, and researchers. Having CES as a member of PEER's management team connects us to Connecticut's network of regional educational service providers; these support the school districts and offer content expertise on a variety of topics relevant to professional development. Because CES's role involves supporting educational agencies in various ways, its input is valuable for a range of topics relevant to practitioners, including curriculum development, professional development, assessment, and the formative use of data. Not every RPP can tap into such a state-sponsored entity, but most have access to some type of local technical assistance or a professional

development provider. For any RPP facing an expertise gap, it's worth identifying what resources exist and giving the organizations that provide those services a prominent role in the partnership.

RPPs may also use network-based and adviser approaches to address gaps in content knowledge. Another strength of the PEER management structure is its connection with the Regional Educational Laboratory (REL) Northeast and Islands, which is directed by one of PEER's founding organizations, the Education Development Center. Since a key mission of RELs is to support, conduct, and enact applied research with educational agencies, they can be valuable resources for expert research consultation. PEER has consulted experts associated with REL Northeast and Islands and EDC both for their own knowledge and to find others with the expertise we seek. Our connection with REL Northeast and Islands also helped us stay informed about other research being conducted in the region that may complement our own. The IES supports 10 RELs around the United States; other RPPs may find them valuable as well.

It's also worthwhile to engage community and state leaders who oversee early childhood education policies. For example, PEER has developed relationships with directors of school districts' early childhood programs, directors of center-based early learning programs, and directors of community-based organizations that serve children. At the state level, we've built connections with representatives of the Office of Early Childhood, the Department of Education, and early childhood advocacy groups. One-on-one communication with such leaders is helpful, but it's often more valuable to convene them in advisory groups and partnership meetings because of the interactions and feedback made possible in these settings. Furthermore, to ensure that advisors' feedback is specific enough to the topics being studied, RPPs might consider developing advisory panels attached to specific partnership projects, rather than (or in addition to) a panel to advise the partnership as a whole.

Conclusions

Developing an RPP centered on a collaborative research agenda can produce research that's relevant, useful, and actionable for those most likely to put the research into practice. Such an approach can also improve education research in general by developing knowledge about contextual factors that enhance or diminish the outcomes of children, their families, and educators. When a collaborative research agenda is established by integrating input from on-the-ground partners who are close to the practices and systems being studied, it's likely that the RPP will identify and define research questions that can shed light on such contextual factors. As a function of their roles, practitioners will bring insights that prompt new and interesting hypotheses, offer more nuanced interpretations of research findings, and help identify effective ways to put evidence into practice.

As with most efforts intended to bridge gaps in related but distinct areas of work, the partnership approach also faces challenges. We discussed three of these in this article: responding to partner priorities, research readiness, and gaps in content expertise. As we've shown, there are many strategies that can mitigate these

issues. One of the most important lessons we've learned is that it's critical to find the appropriate partners for an RPP at an early stage. Partners must have an authentic interest in bridging practice and research and using evidence to guide practice. They must also have deep knowledge of their organizations and communities so they can accurately represent the challenges, interests, and priorities of those organizations and communities during the process of setting the research agenda. Finally, the partners involved in developing the research agenda must have adequate authority to represent their organizations, and must be prepared to advocate for projects that come out of the research agenda. In short, unless the right people participate in defining an RPP's direction, the partnership may find it tough to pursue its research agenda and to produce research that has meaningful benefits for the intended constituents.

Another important but often underappreciated point is that the process of developing and pursuing a collaborative research agenda is rarely neat and linear. It's often messy, with many twists and turns. We've described how conditions like partners' changing priorities and diverging needs have led PEER to reevaluate and adjust some of our approaches. When collaborating with partners to create a research agenda, we recommend that RPP leaders be prepared to encounter challenges, remain open to the opportunities presented by these challenges and use them to improve the partnership, and avoid becoming too attached to their initial agenda. Just as applied researchers advocate using evidence for continuous improvement, RPPs can continuously seek input from partners and use it to make the partnership more effective.

Endnotes

1. Cynthia E. Coburn, William R. Penuel, and Kimberly E. Geil, Research-Practice Partnerships: *A Strategy for Leveraging Research for Educational Improvement in School Districts* (New York: William T. Grant Foundation, 2013).
2. Connecticut Data Collaborative, accessed Feb 2, 2019, http://ctdata.org/.
3. Julie Reed Kochanek, Natalie Lacireno-Paquet, and Rebecca Carey, *Developing a Coherent Research Agenda: Lessons from the REL Northeast & Islands Research Agenda Workshops* (Washington, DC: US Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands, July 2014), https://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/REL_2014014.pdf.
4. Connecticut Data Collaborative.
5. Council of the Great City Schools, "Urban School Superintendents: Characteristics, Tenure, and Salary: Eighth Survey and Report," *Urban Indicator* (Fall 2014): 1–10.

Developing Decision-Making Tools through Partnerships

Amanda Williford, Jason Downer, Kate Miller-Bains, Jenna Conway, and Lisa Howard

Summary

In this article, Amanda Williford, Jason Downer, Kate Miller-Bains, Jenna Conway, and Lisa Howard tell us how a university research center, an early education advocacy group, and a state department of education joined forces in a research-practice partnership to develop and implement a more comprehensive assessment of young Virginia children's readiness for kindergarten. The Virginia Kindergarten Readiness Program, or VKRP, as the assessment they built is called, added measures of math, self-regulation, and social skills to complement Virginia's existing statewide assessment of prekindergarten children's literacy. The aim was not only to better assess children's readiness to enter school, but also to guide teachers' instructional practice and help the state target support.

The partnership produced many benefits: for policy makers, a statewide snapshot of children's readiness; for researchers, on-the-ground feedback from teachers; and for the education department, joint review and interpretation of data patterns to aid decision-making. But at times, the fast pace of statewide implementation affected the university partners' ability to pursue their research aims, at least in the short term, highlighting a recurring theme of this issue—the challenges of balancing researchers' and partners' needs.

Amanda Williford is a research associate professor and CASTL associate director for early childhood education at the University of Virginia School of Education and Human Development. Jason Downer is a professor of education and the director of CASTL at the University of Virginia School of Education and Human Development. Kate Miller-Bains, now at Oak Ridge Associated Universities, was a research assistant professor at the University of Virginia School of Education and Human Development. Jenna Conway is the chief school readiness officer at the Virginia Department of Education. Lisa Howard is president and CEO of E3: Elevate Early Education.

In this article, we describe how the Virginia Kindergarten Readiness Program, a unique research-practice partnership between the Center for Advanced Study of Teaching and Learning at the University of Virginia, Elevate Early Education (an issue-based advocacy group), and the Virginia Department of Education, was developed to answer a key question: how best to assess the readiness skills of incoming Virginia kindergartners. Across the country, departments of education are recognizing the need to understand how children are entering kindergarten and to define the readiness gap between children from low-income backgrounds and their higher-income peers.[1] In recent years, many states have established comprehensive assessment tools as the first step toward understanding children's readiness on a larger scale.[2] These assessments help teachers, school divisions (Virginia's term for districts), and policy makers, and they may be used in a variety of ways depending on the breadth and depth of the data they produce.[3] For example, teachers can use the information to guide and differentiate instruction to fit each student's strengths and weaknesses. School divisions may use the information to target interventions and improve student outcomes. Advocates use it to drive strategic investments in early education, and policy makers use it to align funding for interventions and to understand the impact of the investments they've made.

Yet there's no clear or perfect approach to assessing readiness, and measuring young children's skills is especially challenging.[4] The process is further complicated by the fact that large-scale readiness assessment tools are relatively new, so there aren't a lot of established programs from which to choose. As a result, districts and states have taken a variety of approaches when establishing their own assessment tools. Some have implemented or modified off-the-shelf assessments, others have developed their own systems, and yet others have collaborated across state agencies.[5]

Here we discuss the partnership among the Center for Advanced Study of Teaching and Learning (CASTL), Elevate Early Education (E3), and the Virginia Department of Education (VDOE) to develop the Virginia Kindergarten Readiness Program (VKRP) as a statewide readiness assessment. Through this partnership, researchers brought their expertise in measurement to the task of establishing a sound estimate of children's readiness skills that could be used by the state's advocates, policy makers, and other leaders to guide the creation of an expanded assessment system that would complement an existing statewide assessment system focused solely on early literacy.

Readiness Assessments in the United States

Thanks to an increased awareness of the importance of early childhood learning experiences and federal awards like the Race to the Top—Early Learning Challenge, at least 40 states have instituted or are piloting kindergarten readiness assessments.[6] In practice, these assessments often satisfy multiple needs: identifying students as ready or not ready, influencing classroom instruction at the start of kindergarten, targeting resources for both educators and students, guiding early childhood policies and programs, and more. States have used a range of formats, so their readiness assessments vary both in methodology and in the scope of their chosen measures.[7] Some 30 states have adopted entry assessments, and more

than half of these have chosen observation-based programs, such as commercially available measures like Teaching Strategies GOLD, the Desired Results Developmental Profile—School Readiness, or the Work Sampling System.[8] Connecticut, Ohio, and Maryland have partnered with other departments of education and/or institutes of higher education to develop their own assessments, which use a mix of methods.[9] Oregon, on the other hand, worked with local university researchers to adapt a widely used computerized assessment of math and reading skills and to incorporate an existing rating scale of children's emotional and social skills.[10]

Without assessment, it's hard to tell whether a particular program or policy is meeting its objectives.

Until fall 2019, Virginia lacked a statewide, multidimensional kindergarten readiness assessment. The early childhood education (ECE) partnership among CASTL, E3, and VDOE led to the development and implementation of a kindergarten readiness assessment that's unique in the United States. Choosing VKRP's assessment method was a critical element of the research-practice partnership (RPP). Below, we describe how the partnership influenced this decision and how tensions among the partners were handled along the way.

Assessments as Decision-Making Tools

Student assessment has long been recognized as a powerful tool to guide education programs and policies.[11] Without assessment, it's hard to tell whether a particular program or policy is meeting its objectives. Student assessments serve different purposes depending on when and how they're used. They may provide baseline information about participants' needs, measure the extent to which students are receiving an intervention as it was conceived and planned (that is, intervention *fidelity*), or gauge whether key outcomes have been achieved.[12] No single assessment can meet every need, so choosing an assessment tool requires weighing the trade-offs.[13]

Another consideration is how the information will be collected. Depending on how the data are intended to be used and by whom, departments of education may set different requirements.[14] When states are planning to use the information for accountability, for example, or to make comparisons across schools and divisions, they often require all teachers to administer the same assessment to ensure consistency. Alternatively, if states prioritize using the data for local decision-making, they might let schools or districts choose their own assessments, finding ones that best complement the schools' or districts' initiatives.

Different types of assessments have different advantages and disadvantages. Assessments vary in the extent to which they're consistent when administered across a range of settings (known as *reliability*), as well as the extent to which they provide complete or meaningful information about the skills of interest (known as *validity*). Moreover, these two qualities are often in tension with one another.[15] One benefit of a strong RPP is that it can help states understand the benefits and drawbacks of different assessment tools.

The Virginia Kindergarten Readiness Program

VKRP is an initiative to better understand school readiness and success in Virginia. As an assessment system, VKRP added measures of math, self-regulation, and social skills to complement Virginia's statewide assessment of literacy skills—the Phonological Awareness Literacy Screening, or PALS.[16] With VKRP, Virginia can establish a consistent and more comprehensive statewide baseline of readiness, do more to help teachers and principals meet kindergartners' needs, and better engage families to support young learners who are entering school. E3 conceived the initiative as a way to define the state of school readiness in Virginia and use the data collected to advocate for a stronger investment in high-quality early childhood education.

The interests, expertise, and missions of the VKRP partners intersect and complement one another in several important ways. As a statewide bipartisan issue-advocacy organization that promotes strategic data-driven investments in early education, E3 sought to define and understand the scope of kindergarten readiness through a larger-scale comprehensive assessment. VDOE likewise had a clear stake in estimating children's kindergarten readiness so that it could support teachers and students. VDOE's mission is to ensure that the state has a quality public education system that meets students' needs and helps them become educated, productive, responsible, and self-reliant citizens.[17] VDOE also oversees the state's largest preschool program, the Virginia Preschool Initiative, serving nearly 18,000 four-year-olds annually; thus the department would benefit from a deeper understanding of how preschool participation relates to readiness data. CASTL offered expertise in understanding and measuring children's development, an established record of conducting research in early childhood settings, and experience working in and with Virginia schools. CASTL is a research and development center, and its core mission involves bringing together the best of developmental and education science to guide educational practice at scale. The VKRP partnership gave CASTL an opportunity to engage in a research-to-practice process, from developing and piloting an assessment all the way to statewide implementation. In sum, this partnership gathered all the resources needed to implement a large-scale readiness assessment for divisions, schools, and classrooms across the state.

Partnership History

Virginia defines school readiness as:

> the capabilities of children, their families, schools, and communities that best promote student success in kindergarten and beyond. Each component—children, families, schools and communities—plays an essential role in the development of school readiness. For Virginia's youngest citizens, a ready child is prepared socially, personally, physically, and intellectually in the areas of literacy, mathematics, science, history and social science, physical and motor development, and personal and social development.[18]

More than 90,000 Virginia children enter kindergarten each year. Recent VKRP data indicate that approximately 40 percent—or 36,000 students—may lack the literacy, math, self-regulation, and/or social skills they need to succeed in the classroom. Among children

from low-income backgrounds, the situation is even more concerning: nearly half aren't fully ready, meaning they start behind their peers from higher-income backgrounds.[19]

In the past, Virginia tested children's readiness skills only in literacy. The Early Intervention Reading Initiative, enacted in 1997, gave Virginia schools the resources to assess students' literacy skills when they enter school; the vast majority of school divisions use the PALS assessment for this.[20] But the state knew very little about children's skills in other essential areas. The lack of a consistent, comprehensive measure of kindergarten readiness made it hard to quantify and then address the opportunity gap at the start of kindergarten.

In 2011, E3 set out to make public investment in ECE a priority. An E3-sponsored study found that Virginia legislators, educators, and division leaders reported needing more data on children's kindergarten readiness beyond literacy in order to make decisions about early childhood investments. As a result, E3 decided to partner with CASTL and VDOE to create the VKRP and define the readiness gap more broadly, using a combination of state and private funding. The partners established the following goals:

- Select an assessment tool that can be used statewide to accurately assess children's incoming school readiness across a range of skills.
- Create a snapshot of Virginia's entering kindergartners' readiness skills.
- Define the school readiness skills gap in Virginia and indicate the extent to which estimates of readiness may be different for children in different subgroups.
- Guide the implementation of a statewide, more comprehensive readiness assessment.
- Equip education leaders, legislators, advocates, and other decision makers with information that can be used to guide public policy and funding decisions in early childhood education.

During the first phase of the partnership, the team decided to pilot a commercially published and widely used observation-based assessment to measure kindergarten readiness. This system, which covered a broad range of skills, was being adopted by many states as a kindergarten readiness assessment, and there was some evidence that it was reliable in early childhood.[21] But its usefulness for kindergarten classrooms hadn't yet been examined. CASTL pressed the partnership to test the assessment's reliability and validity in a small sample before proceeding further. This decision exemplifies a unique outcome of the partnership. Without CASTL's involvement, the test pilot likely wouldn't have occurred, because most practitioners and policy makers might assume that a widely used assessment is a good one. But the pilot showed that the tool posed several challenges for assessing kindergarten readiness in Virginia, including lengthy administration time; redundancy in the area of literacy (because Virginia kindergarten teachers were already assessing literacy skills); highly correlated scores across different skill areas (literacy, language, and math) that limited teachers' understanding of how children's skills were differentiated; and high intra-class correlations relative to direct

assessments of the same skills, indicating that the tool wasn't a good choice to provide unique skill profiles of children within a classroom.[22]

Given these findings, CASTL advised that this tool wasn't suited to assess Virginia children's school readiness across key indicators. The research team's recommendation created some tension in the partnership, as the observation-based tool was widely used by other states, fully comprehensive, and appealing to the valued partners in Virginia's early childhood advocacy community. Dismissing this tool also meant moving away from a more naturalistic approach in favor of one that was more standardized and scientifically sound but less often used. CASTL, E3, and VDOE collaborated to present the pilot data clearly and objectively to interested parties. CASTL decided that in the next phase, where the goal was to provide a comprehensive estimate of the readiness gap in Virginia, VKRP would use a combination of measures known to be valid and reliable. CASTL integrated the literacy data already collected (a teacher-administered direct assessment) with measures of math (also a teacher-administered direct assessment) and of self-regulation and social skills (using teacher-rating scales). This approach revealed that the proportion of students entering kindergarten without key readiness skills was larger than had previously been estimated using literacy data alone.[23]

In addition to establishing a statewide estimate of readiness, CASTL issued a report to the Virginia General Assembly that made several recommendations for the statewide rollout of a more comprehensive readiness assessment system.[24] The most critical aspect involved building off the infrastructure of the state's existing literacy assessment so that teachers, administrators, and policy makers could work with a system they knew well and obtain useful data across multiple readiness skills. VKRP hired a contractor to program math and social-emotional measures into an online application that would interface with the existing literacy assessment system. Thus, teachers could use a single link and login to upload their student rosters, access all the assessments, see integrated readiness reports, and acquire instructional resources. CASTL's report also recommended comprehensive training for educators and school leaders on how to administer the new assessments and how to interpret and use the data.[25]

Over the next three years, CASTL implemented a voluntary rollout in which division leaders could choose whether to adopt VKRP. CASTL continued to work with teachers, divisions, and VDOE to improve the assessment system, online application, reports, and available resources. It was unusual for CASTL researchers to take the lead during the rollout, rather than VDOE, but it offered a big advantage: researchers who are deeply involved in implementation (beyond just providing capacity) will get a more accurate perspective on what's happening in classrooms, so they can see where the implementation is working and where it's falling short. Thus CASTL could use an iterative approach, regularly gathering feedback from teachers, principals, and other practitioners and using it to revise the assessment system substantially each year. The process was more intense than what VDOE could have done alone. For example, when teachers asked for a spring assessment to capture growth during kindergarten, the research team was the first to hear their request. The researchers quickly applied for outside funds to further develop the tool, and

they were ready to implement it shortly after VDOE presented the formal request for a spring assessment.

One reason the practice and policy world has been slow to adopt the use of science to guide decision-making is that scientists are often at least an arm's length away from day-to-day complexities.

CASTL's implementation of VKRP has had other benefits as well, such as the team's day-to-day responsiveness to educators. CASTL interacts regularly with educators, and its response systems (chat, phone, email) are available whenever the online system isn't working well or when the educators don't know how to access or interpret data. When 99.5 percent of the data are complete, a few missing bits may mean little to a researcher who's viewing the data in aggregate. But losing a child's data because the server became unstable means a great deal to a teacher who just spent 25 minutes assessing a student. The CASTL team's care and commitment to data from the level of the child to that of the state has boosted CASTL's credibility with both VDOE and education practitioners. By leading the implementation of data collection, CASTL understands the data's strengths and limitations, which provides critical context for explaining data patterns when presenting results to VDOE.

One reason the practice and policy world has been slow to adopt the use of science to guide decision-making is that scientists are often at least an arm's length away from day-to-day complexities. Thus researchers' recommendations can be seen as (and may well be) out of touch, and are treated with skepticism. Having CASTL directly implement VKRP removed this barrier and gave the team more street cred with our practice partners, so that the assessment system balanced practicality with good science.

But having the research team so closely engaged in implementation also has a major drawback. Researchers who are deeply involved in day-to-day operations tend to become invested in the particular assessment system they're overseeing. Researchers are often included in ECE partnerships to provide independent, clear-eyed advice and insight. If an organization leads implementation year in and year out, it may lose sight of opportunities to innovate and adapt to meet the changing needs of schools or children.

Challenges of the Work

For CASTL, a significant challenge has been the need to quickly bring VKRP to scale across the state, and to do so in the context of very public data sharing. This task collided with the need to choose assessments that are scientifically sound and to build data systems that maintain the data's integrity. When the state funded VKRP's voluntary rollout in 2015, the CASTL team was asked to build an online system that integrated with the state's literacy platform, to create online reports, and to develop instructional resources. This work had to be completed within months so that more than 500 kindergarten teachers could administer the new VKRP assessments, alongside the existing literacy assessments, to almost 10,000 students across 21 school

divisions. The first version was clunky—the assessments were slow, which frustrated teachers; the reports weren't interactive; and the instructional resources weren't embedded in the reports for easy access. We made vast improvements to the assessment system even while we were rapidly expanding it into more Virginia school divisions. In retrospect, this constituted an iterative approach. We learned from teachers and principals, and we built a better system because of it.

A second significant challenge was that day-to-day development and operation of the assessment, along with the need to provide summary reports to the state and other interested parties, left little time to use the vast amount of data collected for research purposes. Fundamentally, CASTL participated in VKRP to serve the state, and traditional research for academic purposes has had to take a back seat. CASTL regularly provides data summaries to divisions, VDOE, and the state legislature. On the one hand, this helps define the relationship as a true partnership. But this sort of work—a state-funded school readiness initiative implemented by CASTL in partnership with the state—isn't valued in academia in the same way that work done under the auspices of a research grant would be. And not using the population-level data acquired by VKRP to advance the science of school readiness may also be a missed opportunity.

As VKRP continues to move toward full statewide implementation, more work lies ahead for all involved. It's hard to implement a change in practice across thousands of classrooms; kindergarten teachers, school leaders, and families need support to ease the transition. CASTL and VDOE have tried to ensure the transition goes well—phasing in the assessment over time; communicating regularly; offering in-person and online training and technical assistance; and providing resources that include a website, a blog on instructional resources, and customized professional development. This support will need to grow as every kindergarten teacher in the state comes on board.

VDOE has identified another critical challenge: positioning VKRP in the context of a much broader understanding of school readiness. When Virginia gathered educators, leaders, advocates, and others to define school readiness for the state, the aim was not just to focus on the skills of children entering kindergarten, but also to directly acknowledge that communities, families, and schools must be "ready" and "prepared" to support the transition to school.[26] VKRP expands Virginia's assessment of children's readiness skills, but it's not comprehensive even in its measures of children's early learning. For example, VKRP doesn't measure such crucial areas of learning as language and critical thinking. We had to make tradeoffs between breadth (measuring all areas of early learning) and depth (providing enough precision to guide instruction) while prioritizing feasibility and practicality. But this decision comes with the risk that any unmeasured readiness skills may be perceived as less important.

The question of how to report the results also presents challenges. Like many assessments, VKRP scores children on a scale to capture the variability in math, literacy, social, and self-regulation skills. A natural question, then, is what point on the scale indicates that a child is ready to take advantage of the learning opportunities presented by kindergarten. This issue involves myriad technical measurement questions, many of

which can only be answered after repeated use of the assessment with many children over time. But one of the most pressing needs of practitioners, education leaders, and advocates is to use VKRP data to identify which children are *least* ready for kindergarten, and thus to guide investments in those children so that more of them enter kindergarten with the foundational skills they need. It's common practice to establish benchmarks (often called thresholds or cut points) to determine where students fall in comparison to a standard, and VKRP has done so based on a combination of data and theory. Benchmarks are a quick way to interpret a student's standing. For instance, a student who scores well above the benchmark likely possesses a high level of skills in that area. And teachers should be concerned about a student whose scores fall well below the benchmark in an area. Yet a benchmark is an imprecise estimate, and this can be problematic for students who fall just above or below it. Thus, the VKRP team has been careful to tell teachers that being above or below the benchmark on a VKRP assessment shouldn't be the sole criterion for understanding a child's readiness when it comes to that skill. Continual progress monitoring plays a critical role, because students develop skills at different rates and respond differently to instruction and support.

Another challenge, for VDOE, CASTL, and E3 alike, is to make sure that VKRP data are understood in a broader context, not simply as a set of scores that represent skills internal to a child or group of children. This means developing careful reports that aggregate VKRP readiness data up to classroom, school, and division levels to represent how well communities are preparing children for school. We've also held fast to the notion that these data must be actionable, not just for decision-making at the state and local levels, but also for teachers who must individualize their instruction because children enter school with varied skills and experiences. Family reports have been carefully crafted so that teachers convey children's strengths as well as the challenges they face.

All the partners have identified myriad challenges inherent to developing an assessment system that can be used for multiple purposes—for teachers in their classrooms; for divisions making professional development decisions; and for monitoring progress at the school, division, or state level.[27] It's appropriate and prudent to use VKRP data (and other sources of ECE information) for these purposes, as well as to identify readiness gaps, track system-level trends, and effectively allocate education resources. But VKRP data could be misused, particularly for punitive, high-stakes purposes. Although the VKRP can provide reliable estimates of readiness across a variety of contexts, it wasn't designed for a high-stakes accountability environment, and it wouldn't be appropriate for determining consequences for students, teachers, or programs. Rather, the data are primed to help key players in classrooms, schools, divisions, and government make data-driven decisions about how to best meet the needs of Virginia's youngest students and invest strategically in early childhood initiatives. Many school division leaders were hesitant to be among the first to participate, as they worried that publishing their division's data might lead to unfavorable comparisons. They were also concerned that the data would be used for accountability purposes. So CASTL, VDOE, and E3 worked to communicate the limits to using VKRP data for accountability purposes, and they've

continued to encourage policy makers to see this information as evidence of students' and educators' needs—not of the shortcomings of individual students, teachers, schools, or programs.

VKRP's Successes

Alongside the challenges, the VKRP partnership has seen important successes that likely wouldn't have occurred if CASTL, E3, and VDOE hadn't taken risks and built a relationship. First, the crosscutting partnership has brought increased attention to improving early childhood education in Virginia, especially preschool. Beyond the kindergarten classroom, VKRP can help Virginia connect individual readiness to longer-term outcomes, such as third-grade test results and high school graduation rates. Demonstrating the relationship between kindergarten readiness and longer-term outcomes helps emphasize the importance of early investments, and paints a clearer picture of student achievement over time. Without a consistent statewide assessment at school entry that measures more than just literacy, it's difficult to analyze how schools can best promote student growth, especially in the early elementary grades.

By identifying school readiness gaps, VKRP sheds light on inequities in quality or access, helping policy makers and practitioners deploy resources strategically.

VKRP can help guide improvement across the early childhood system—that is, the diverse set of programs where children are cared for and educated before kindergarten. Virginia children currently lack equitable access to high-quality early childhood care and education. Seventy percent of children from birth to five years of age don't have access to affordable childcare.[28] Thirty percent participate in public programs whose quality varies because it's not measured consistently. VKRP can help create a sense of urgency that will compel policymakers and practitioners to work together to unify and strengthen the early childhood system so that more Virginia children can enter kindergarten ready for school. Specifically, VKRP can show where quality early childhood programming is associated with better child outcomes, thus highlighting the return on these investments. By identifying school readiness gaps, VKRP sheds light on inequities in quality or access, helping policy makers and practitioners deploy resources strategically. Along with other important sources of data about the early childhood system, VKRP also promotes continuous quality improvement at the community level.

This data-driven approach to understanding children's readiness at school entry has secured largely bipartisan support, resulting in recent state investments to improve Virginia's early childhood programs. In 2017 Virginia released a legislative report titled *Improving Virginia's Early Childhood Development Programs.* In response to some of the report's findings, E3, alongside policy makers and CASTL, developed a Virginia state House-led 2018 budget package that was designed to advance high-quality early education, with $6 million in targeted funds. This investment represents a shift in focus from expanding access to ensuring high quality. The legislative results included:

- A mandate that the VKRP be implemented in all kindergarten classrooms, and expanded to assess students in both fall and spring of the kindergarten year.
- Professional development for teachers and school divisions to help them use the data effectively to improve teaching and learning in the classroom.
- Increases in per-pupil funding for the state-funded preschool program.
- Funding to observe the quality of teacher-child interactions in each state-funded preschool classroom, and to provide professional development to improve that quality.
- Ensuring that every state-funded preschool program uses a comprehensive, evidence-based curriculum package.

Beyond the benefits to the state and its young children who are preparing for school, the VKRP partnership has been a positive learning experience for CASTL, as a university-based research center with a mission of bridging the gap between science and practice. We've made tremendous gains in understanding how to bring science and data to conversations with early childhood advocates such as E3, and with policy makers such as the state legislators in the joint House and Senate preschool subcommittee. By participating in these conversations in the state capitol, CASTL's scientists had a rare opportunity to share evidence-based practices from the field of early childhood education; this approach has ensured that state policy makers are basing their ECE decisions on sound, developmentally appropriate data about readiness skills.

Through these interactions with advocates and legislators, we've been learning how to disseminate the science in easily digestible ways so that it will be heard, understood, and acted on. A 15-minute presentation to a legislative subcommittee meeting sounds nothing like its counterpart at a national research conference. It must be brief, clear, and expressed in nontechnical language, with graphs and figures that a wide audience can understand. We've leaned heavily on CASTL's instructional technology and design team to hone our messaging through the best data visualization strategies.[29] The same can be said for legislative reports; these must be concise and to the point, responding to the key questions of policy makers who are deciding about future investments. The science and evidence must be precise and thorough; in other words, take-home messages must be straightforward, easy to follow, and organized in easily digestible chunks, but they must be backed by extensive tables and supporting materials in appendices to substantiate the rigor of the effort.

VKRP has also had an enormous impact on CASTL's relationship with school districts across the state. None of the VKRP work would be possible if teachers and school leaders hadn't been willing to embark on a joint mission to improve how we assess school readiness and use the data. Fostering relationships with more than 130 school divisions has been both daunting and energizing. Access to the varied experiences and contexts of districts statewide—urban, rural, linguistically and racially diverse,

and more—has helped us understand the concerns of frontline educators in unprecedented ways.

CASTL aims to bring science to problems that matter to educators and policy makers, and the VKRP partnership has created additional opportunities to identify shared research agendas whose findings can guide future state policies and investment decisions. For example, we learned quickly that teachers and administrators vary considerably in their ability to interpret assessment data across school divisions—that is, to make sense of VKRP's school readiness data and then use it for decision-making. We viewed this as an opportunity to develop and pilot data-use training with embedded feedback loops from our district partners. From a scientific standpoint, we saw a chance to conduct several small, low-cost experiments when we piloted these interventions, giving us evidence of what worked and what didn't. The first of these experiments, conducted in a single school division, investigated the effects of one-on-one data consultations between teachers and trained research staff, relative to no additional supports, on teachers' perceptions and use of the VKRP data.[30] Though the results suggested that these one-shot data conversations improved teachers' understanding of the VKRP assessments, it wouldn't be feasible to provide such one-on-one help at scale. The following year, we tested other formats and delivery methods to see if the consultations could be just as effective when conducted remotely and/or with groups of teachers from all participating divisions. This second experiment gave us two useful pieces of information: that not all schools and divisions were interested in or capable of using such services, and that one-on-one, remote consultations akin to hotlines could deliver results similar to those produced by in-person sessions.[31]

These experiences also enhanced our own approach to research. We've learned how to involve practice partners and how to ask ourselves difficult questions about the feasibility of the work at scale. After all, a proof of concept with a stellar evidence base does little good if it ultimately has no chance of being successfully implemented in the field. That may seem obvious, but researchers can find it easy to rest on the principles of the scientific method and ignore issues of scope and practicality. Successful public-university partnerships can pave the way for universities to value more highly the kind of scholarship where scientists work alongside others to infuse research evidence into public policy decision-making.

Recommendations for Researchers and Policy Makers

In developing Virginia's statewide tool to measure young children's readiness skills, all members of the partnership learned how to collaborate effectively to achieve common and distinct goals. Many of the lessons learned have broader implications for using RPPs to develop and implement assessment tools that can guide decision-making.

One lesson is that project goals must be transparent within and across the partners. It's also crucial that all parties understand the benefits they can expect and the challenges they'll face. For example, when CASTL and E3 began working together, it was made clear that even though E3 was seeking help from CASTL because of its measurement expertise, this wasn't a research project. E3's goal was to answer a specific question to advance its advocacy agenda: How many Virginia children enter kindergarten "not

ready"? E3 was also straightforward in conveying that CASTL would be entrusted with designing and implementing a pilot study to provide the most reliable and valid answer with the funds available. We at CASTL found the VKRP project appealing because it allowed us to examine the science, determine a set of procedures, pilot those procedures, examine the data, and revise based on what the data told us. Transparent goals also helped us develop relationships with Virginia's school divisions. As VKRP grew, CASTL clearly described to each school division how the data would be used. We explained that we would use the data collected through VKRP both to understand how to improve the system and to conduct related research; as such, the procedures employed in these endeavors might include additional surveys and randomization into piloting of procedures and interventions that go beyond business as usual for schools and educators.

In hindsight, it's easy to describe the VKRP partnership as smooth and successful. But anyone considering joining an RPP should be prepared to face a multiyear roller coaster ride. Each partner's staff, leadership, and resources must be aligned and committed to supporting the partnership for the long haul. Funding for VKRP is a good example of this up-and-down ride. E3, a strong advocate for the initiative, assertively articulated a five-year plan to decision makers in the Virginia legislature and VDOE, and it secured state and private funding to get VKRP off the ground. If E3 hadn't successfully argued for continuous funding, VKRP would never have moved to scale statewide. Even now, sustained funding isn't a given, and all parties must be prepared for what the next phase requires. CASTL secured its own funding to support VKRP. Anticipating that VKRP would be expanded to include spring assessments, CASTL lined up internal funding to pilot more than 200 math items with 900 students in preschool through first grade. This allowed us to select a diverse subset of items with strong evidence of reliability and validity as we expanded to assess in the spring and in additional grades. We continue to apply for foundation and federal funding to support research activities that aren't part of VKRP but still complement the state's agenda.

Relatedly, VDOE's investment in VKRP increased as state funding became more stable and as VKRP moved from a voluntary pilot into mandatory statewide implementation. As a result, the partnership between CASTL and VDOE has become stronger. So far, CASTL has been responsible for VKRP's implementation. But with VKRP going statewide, VDOE will likely take more ownership over time, and CASTL's role may shift. Thus a successful partnership doesn't develop in a linear way, and participants must maintain their commitment in the face of instability and be open to role changes.

[People's] aversion to change creates all sorts of risk for partners, and it makes robust communication essential.

Partners also need to be clear-eyed about risk. VKRP now has wide but certainly not universal support. In its early phases, some decision makers strongly opposed the idea of a statewide, more comprehensive assessment at the start of kindergarten. CASTL's key role in developing and implementing the assessment meant that the University of

Virginia School of Education and Human Development, of which CASTL is a part, would be associated with an initiative that might be unpopular among the state's ECE decision makers. Thus CASTL needed the education school's support for this high-profile, high-stakes, and potentially controversial initiative. Without the dean's support, CASTL faculty wouldn't have taken on the partnership.

Implementing something new or fundamentally changing an existing process is difficult; often, people (and the organizations they belong to) don't like change. This aversion to change creates all sorts of risk for partners, and it makes robust communication essential. In VKRP, VDOE and E3 help CASTL present the data—tables, figures, and text—in ways that are clear and easily digestible, so that teachers, school leaders, and decision makers will understand the information and be more likely to use it. E3 helps frame the conversation to ensure that the data can ultimately be used to make strategic investments in early education that focus on improving programs for young children. This often involves careful planning about who needs access to the data and how to ensure that VKRP remains focused on data-driven decision-making without taking on a high-stakes or punitive component. As we've already mentioned, VDOE has worked to make sure that VKRP data is valuable to everyone in the pre-K–12 system.

It's also critical to stay connected to the front line, which in our case means early childhood classrooms and especially the interactions between teachers and young children. Infants, toddlers, preschoolers, and kindergartners learn through relationships in which they feel supported, encouraged, and challenged to be curious, take on new tasks, and think critically.[32] Working at the state level allows academics to influence policy making and the distribution of resources. But like academia itself, state-level work is a step removed from the classroom and day-to-day interactions between adults and young children. Similarly, assessments can sometimes be implemented in a way that leads to an overly academic focus, pushing teachers, school leaders, and families to emphasize rote skills rather than robust learning and development. When assessments prioritize certain skills over others, we can fail to grasp the importance of teaching the whole child in a comprehensive and integrated way.

Our efforts to connect to teachers, children, and the classroom experience have no doubt helped us gain support from decision makers for the implementation of VKRP. CASTL provided in-person training to all kindergarten teachers whenever a new school district adopted VKRP. CASTL also gave teachers all the technical assistance they needed, conducted professional development workshops, undertook classroom observations during assessment windows, and gathered direct feedback from teachers via satisfaction surveys and focus groups. E3 also implements VKRP assessments in its own model demonstration early childhood program; it then presents the data to teachers and parents, and gives teachers training and feedback to improve their practice. VDOE works to ensure that the assessments are tied to what's happening in kindergarten classrooms and beyond. The department has helped to clearly articulate how the assessments are aligned with Virginia's preschool development framework and kindergarten standards of development and learning.

Conclusions

The VKRP partnership seeks to ensure that Virginia's youngest children, from birth through preschool, get the support they need to reach their potential as they enter the state's public school system. VDOE, CASTL, E3, and others used each organization's expertise to develop a kindergarten entry assessment system that's practical, scalable, and evidence-based. Beyond this shared goal, the partnership also produced individual benefits for each party: for policy makers, a statewide snapshot of children's readiness; for CASTL, on-the-ground feedback from teachers to guide revisions to the assessment interface; and for VDOE, joint review and interpretation of data patterns to aid decision-making. Of course, to maximize the impact of the partnership, each organization had to be willing to adapt, but never to the point of undermining its own fundamental mission. This underscores an important point: an organization must be prepared to operate out of its comfort zone when joining a partnership that's focused on local, state, or national data-based decision-making and assessment initiatives in early childhood. But the combined strengths of researchers, practitioners, advocates, and policy makers can produce a technically sound approach that's feasible to implement and that targets the needs of a variety of end users.

Endnotes

1. Elliot Regenstein, Maia Connors, and Rio Romero-Jurado, "Valuing the Early Years in State Accountability Systems under the Every Student Succeeds Act," *Ounce Policy Conversation No. 5* (February 16, 2016), https://startearly.org/resource/valuing-the-early-years-in-state-accountability-systems-under-the-every-student-succeeds-act/.

2. Ibid.

3. Elliot Regenstein et al., "Uses and Misuses of Kindergarten Readiness Assessment Results," *Ounce Policy Conversation No. 6* (February 22, 2017), http://www.theounce.org/wp-content/uploads/2017/03/PolicyConversationKRA2017-1.pdf; Kyle Snow, *Developing Kindergarten Readiness and Other Large-Scale Assessment Systems: Necessary Considerations in the Assessment of Young Children* (Washington, DC: National Association for the Education of Young Children, 2011); Catherine E. Snow and Susan B. Van Hemel, *Early Childhood Assessment: Why, What, and How* (Washington, DC: National Academies Press, 2008).

4. Sally Atkins-Burnett, "Measuring Children's Progress from Preschool through Third Grade" (Washington, DC: Mathematica Policy Research, 2007), https://www.mathematica.org/-/media/publications/pdfs/measchildprogress.pdf; Elliot Regenstein, "An Unofficial Guide to the Why and How of State Early Childhood Data Systems," *Ounce Policy Conversation No. 7* (August 22, 2017), https://startearly.briteweb.com/app/uploads/2020/09/PUBLICATION_An-Unofficial-Guide-to-the-Why-and-How-of-State-Early-Childhood-Data-Systems.pdf.

5. J. Ackerman, *Real World Compromises: Policy and Practice Impacts of Kindergarten Entry Assessment-Related Validity and Reliability Challenges* (Princeton, NJ: Educational Testing Service, 2018), https://doi.org/10.1002/ets2.12201.

6. Center on Standards and Assessment Implementation, *State of the States: Pre-K/K Assessment*, 2017, http://www.csai-online.org/sos.

7. Center on Standards and Assessment Implementation, *State of the States;* Amanda P. Williford et al., *The Virginia Kindergarten Readiness Project: Executive Summary & Legislative Report, Fall 2014, Phase II* (Charlottesville, VA: University of Virginia, CASTL), https://vkrponline.org/wp-content/uploads/sites/3/2020/09/VKRP_Executive_Summary_and_Legislative_Report_2015_01_21_updated_1.pdf.

8. Samuel J. Meisels, Xiaoli Wen, and Kristy Beachy-Quick, "Authentic Assessment for Infants and Toddlers: Exploring the Reliability and Validity of the Ounce Scale," *Applied Developmental Science* 14 (2010): 55–71, https://doi.org/10.1080/10888691003697911; California Department of Education, "Introduction to Desired Results" (2016), http://www.cde.ca.gov/sp/cd/ci/desiredresults.asp; Cate Heroman et al., *Objectives for Development and Learning: Birth through Kindergarten* (Washington, DC: Teaching Strategies, 2010).

9. Ackerman, *Real World Compromises*; Regenstein et al., "Uses and Misuses."

10. Megan McClelland et al., "A Research Perspective on Oregon's Kindergarten Assessment," letter to Oregon Early Learning Council, February 13, 2014, https://www.oregon.gov/ode/educator-resources/assessment/Documents/ka_research_perspective.pdf.

11. Laura Hamilton, "Assessment as a Policy Tool," *Review of Research in Education* 27 (2003): 25–68, https://doi.org/10.3102/0091732X027001025; National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, *Types of Evaluation* (Atlanta, GA: Centers for Disease Control and Prevention, 2014), https://www.cdc.gov/std/program/pupestd/Types%20of%20Evaluation.pdf.

12. Center for Standards, Assessment, and Accountability, "State of the States," https://csaa.wested.org/tools/state-of-states/; Ackerman, *Real World Compromises*.

13. Snow, "Developing Kindergarten Readiness"; Snow and Van Hemel, *Early Childhood Assessment*.
14. Center on Standards and Assessment Implementation, *State of the States*.
15. Snow and Van Hemel, *Early Childhood Assessment*.
16. Phonological Awareness Literacy Screening, "Background of PALS," University of Virginia, School of Education and Human Development, https://pals.virginia.edu/rd-background.html.
17. Virginia Department of Education, "About VDOE," http://www.doe.virginia.gov/about/index.shtml.
18. Virginia Department of Education, "School Readiness," http://www.doe.virginia.gov/instruction/early_childhood/school_readiness/.
19. Virginia Department of Education and the Center for Advanced Study of Teaching and Learning, *VKRP Status Report for the Chairmen of House Appropriations and Senate Finance Committees* (Charlottesville, VA: University of Virginia, CASTL, October, 2019).
20. Phonological Awareness Literacy Screening, "Background of PALS."
21. Richard G. Lambert, Do-Hong Kim, and Diane C. Burts, "Using Teacher Ratings to Track the Growth and Development of Young Children Using the *Teaching Strategies GOLD*® Assessment System," *Journal of Psychoeducational Assessment* 32 (2014): 27–39, https://doi.org/10.1177/0734282913485214; Richard G. Lambert, Do-Hong Kim, and Diane C. Burts, "The Measurement Properties of the Teaching Strategies GOLD® Assessment System," *Early Childhood Research Quarterly* 33 (2015): 49¬–63, https://doi.org/10.1016/j.ecresq.2015.05.004; Ackerman, *Real World Compromises*; GG Weisenfeld, *Assessment Tools Used in Kindergarten Entry Assessments* (KEAs): State Scan (CEELO FastFact) (New Brunswick, NJ: Center on Enhancing Early Learning Outcomes , 2017).
22. Katherine L. Miller-Bains et al., "Examining the Validity of a Multidimensional Performance-Based Assessment at Kindergarten Entry," *AERA Open* 3 (2017), https://doi.org/10.1177/2332858417706969; Amanda P. Williford, Jason T. Downer, and Bridget K. Hamre, *Virginia Kindergarten Readiness Project—Phase 1 Report: Concurrent Validity of Teaching Strategies GOLD* (Charlottesville, VA: University of Virginia, CASTL, 2013).
23. Williford et al., *Readiness Project—Phase II*; Martha B. Bronson et al., *Child Behavior Rating Scale* (Cambridge, MA: Abt Associates, 1990); Herbert P. Ginsburg and Sandra Pappas, "Invitation to the Birthday Party: Rationale and Description," *ZDM Mathematics Education* 48 (2016): 947–60, https://doi.org/10.1007/s11858-016-0818-4.
24. Williford et al., *Readiness Project—Phase II*.
25. Ibid.
26. Virginia Department of Education, "School Readiness."
27. Regenstein et al., "Uses and Misuses"; Ackerman, *Real World Compromises*.
28. Billy Cannaday, "Virginia is for Learners: Advisory Committee Recommendations," presentation, Virginia Department of Education Board Meeting, January 2019.
29. National Forum on Education Statistics, *Forum Guide to Data Visualization: A Resource for Education Agencies* (NFES 2017-016) (Washington, DC: National Center for Education Statistics, US Department of Education, 2016), https://nces.ed.gov/pubs2017/NFES2017016.pdf.
30. Sadie L. Hasbrouck, "Using Kindergarten Entry Data to Guide Teachers' Skills, Beliefs, Self-Efficacy, and Practices Through Instructionally Focused, Data-Driven Consultation" (PhD dissertation, University of Virginia, May 2016), https://libra2.lib.virginia.edu/downloads/fb494876g?filename=Hasbrouck_Sadie_2017_PHD.pdf.

31. Katherine L. Miller-Bains et al., "Digging into Data: An Evaluation of Supports to Help Kindergarten Teachers Interpret Entry Assessments," unpublished manuscript, University of Virginia, March 2019.

32. Bridget K. Hamre, "Teachers' Daily Interactions with Children: An Essential Ingredient in Effective Early Childhood Programs," *Child Development Perspectives* 8 (2014): 223–30, https://doi.org/10.1111/cdep.12090.

Scaling Early Childhood Evidence-Based Interventions through RPPs

Laurie Brotman, Spring Dawson-McClure, Dana Rhule, Katherine Rosenblatt, Kai-ama Hamer, Dimitra Kamboukos, Michelle Boyd, Michelle Mondesir, Isabel Chau, Erin Lashua-Shriftman, Vanessa Rodriguez, R. Gabriela Barajas-Gonzalez, and Keng-Yen Huang

Summary

In 2014, New York City launched its Pre-K for All program, which rapidly tripled the number of children in free, full-day prekindergarten. Two years later, the city rolled out ThriveNYC, a citywide mental health initiative with a focus on early childhood.

By this time, a team from New York University's medical school had partnered for nearly two decades with the city's Division of Early Childhood Education, during which time they developed, tested, and refined ParentCorps, an intervention aimed at supporting the parents and teachers of prekindergarten children. They were thus well positioned to take on the citywide scale-up of their proven intervention. Nonetheless, the partnership was challenged by the scope of the scale-up and by the need to modify the intervention, which had been developed and tested in schools, for use in the community-based organizations that house many of the city's prekindergarten programs. In this article, Laurie Brotman and colleagues describe how their long-established partnership principles—for example, ParentCorps's commitments to racial equity, centering parents' voices, and continuous learning— helped guide their actions, their strategy development, and ultimately their plan for scaling ParentCorps locally and nationally.

www.futureofchildren.org

At the NYU Grossman School of Medicine, Laurie Brotman is Bezos Family Foundation Professor of Early Childhood Development in the Department of Population Health; Spring Dawson-McClure is an assistant professor in the Department of Population Health; Dana Rhule is a clinical assistant professor in the Department of Population Health; Katherine Rosenblatt is a clinical assistant professor in the Department of Population Health; Dimitra Kamboukos, Vanessa Rodriguez, and R. Gabriela Barajas-Gonzalez are assistant professors in the Department of Population Health; Keng-Yen Huang is an associate professor in the Department of Population Health; Kai-ama Hamer is ParentCorps associate director; Michelle Boyd is a ParentCorps specialist; and Michelle Mondesir, Isabel Chau, and Erin Lashua-Shriftman are ParentCorps managers.

The authors would like to acknowledge the more than 100 ParentCorps coaches, trainers, research faculty and staff from the Center for Early Childhood Health and Development who have shown up in deep and courageous ways to center children and families of color. We are indebted to our New York City Department of Education partners and the thousands of pre-K program leaders, teachers, mental health professionals, parent support staff, and families who have shared their knowledge, perspectives, and most importantly, their stories.

Many early childhood interventions show great potential in early testing but then fail to scale effectively.[1] Likewise, school districts and other public systems may introduce evidence-based interventions (EBIs) to improve the quality of their programs and promote better child and family outcomes, but without adequate support for implementation and continuous improvement, such interventions often remain underused or fail to achieve their aims.[2] According to the Society for Prevention Research's Mapping Advances in Prevention Science (SPR MAPS) IV Translational Research Task Force, achieving scaled impact "remains one of the most vexing challenges facing prevention science."[3]

Researchers have argued that traditional models of moving from efficacy trials to effectiveness trials to scaling haven't worked, in part due to a failure to consider key aspects of the system (such as policies and practices) in which the EBI is being scaled.[4] Scholars describe the "pipeline paradox" as based on faulty assumptions about the development and scaling of interventions being a linear process.[5] Specifically, there's a common belief that once an intervention is tested, refined, and shown to work (and considered to be an EBI), the only step remaining is for people and systems to use it. In this linear model, the EBI developer's role is quite limited, since the EBI is considered complete as tested; the researchers serve only as independent evaluators of the scaled EBI's effectiveness.[6] Importantly, such *accountability* evaluations, which intentionally restrict the relationships among researchers, policy makers, and practitioners, often provide minimal benefits to the system or little information that can be used to improve systems-level policies and practices. Many recent reviews about scaling EBIs in public education systems have concluded that researchers, practitioners, and policy makers must work together in new ways.[7]

In this article, we describe a research-practice partnership (RPP) developed to support the scaling of an early childhood EBI, ParentCorps, as part of New York City's Pre-K for All. Key partners who have contributed to the RPP over the past five years include ParentCorps developers; implementation leaders and researchers from the Center for Early Childhood Health and Development at New York University's Grossman School of Medicine; and policy makers, program leaders, and researchers from the Division of Early Childhood Education (DECE) of the New York City Department of Education (DOE). By bringing together researchers, practitioners, and policy makers in new ways, this RPP aims to support ParentCorps implementation in prekindergarten (pre-K) programs; strengthen impact at scale; and produce sustainable improvements to ParentCorps and DOE policies, practices, and programs.*

Background

In 1998, the first author of this article, Laurie Brotman, along with Esther Calzada, developed ParentCorps as a preventive intervention for culturally diverse families (with respect to race, ethnicity, immigrant

* The RPP includes leaders from the New York City DOE's Research and Policy Support Group and the DECE Mental Health & Wellness, Teaching & Learning, Data & Analytics, and Performance teams. Principal research partners from the NYU Grossman School of Medicine's Department of Population Health who are not authors on this article include Alexandra Ursache, Samrachana Adhikari, and Andrea Troxel.

status, nationality, religion, and more) living in historically disinvested neighborhoods. As clinical psychologists and prevention scientists, they brought both research and practice perspectives to their work. Their vision for ParentCorps was influenced by cultural adaptation efforts that sought to ameliorate the pervasive and persistent patterns of limited access to culturally relevant mental health services for communities of color. Cultural adaptation aims to enhance engagement and effectiveness through a process of "systematic modification of an EBI to consider culture and context in such a way that it is compatible with the client's cultural patterns, meanings and values."[8] Ideally, cultural adaptation follows a series of steps that are guided by a deep familiarity with the research evidence and its limitations; by a commitment to preserving core components of fruitful interventions alongside a critical analysis of the extent to which interventions are aligned with white middle-class values (or consider white culture as normative); and by partnership with community members, extensive pilot work, and tests of efficacy.[9] Meta-analyses generally show that cultural adaptation can successfully meet the needs of people of color.[10]

In the spirit of cultural adaptation, Brotman and Calzada reviewed the research on parenting interventions for young children to identify core components: that is, a set of behavioral parenting strategies and adult behavior-change techniques (such as role play or home practice). As they developed the content and delivery model, they partnered with a respected Black-led community-based organization to center family voices and to engage community stakeholders and cultural informants, including Black and Latino parents, educators, and mental health professionals.[11]

Looking ahead to implementation at scale, the developers saw the need to design a program that addressed the complexities of city life. Many urban centers are highly segregated, with tremendous variation throughout nonwhite areas, ranging from racial/ethnic enclaves to multicultural neighborhoods.[12] The characteristics of urban populations also shift over time due to changing patterns of marriage, fertility, and immigration (for example, from 2000 to 2010, 6 percent of New York City's approximate 29,000 census blocks changed from predominantly white to predominantly Latino or Asian American). In these ever-evolving communities, individual experiences of culture also change—the acculturation of each adult in a family follows a unique path, adding further nuance to the family's cultural characteristics.[13]

ParentCorps is not for one particular cultural group; instead, it embraces a broad definition of culture. In the parenting program and professional development, parents and teachers are asked to reflect on their own values and beliefs and to consider how these are influenced by their cultural identity—for example, a Puerto Rican mother with pronounced familistic values; an African-American father with strong racial identity; a Jamaican grandfather who immigrated as an adult; or a third-generation, English-speaking Dominican parent.

ParentCorps includes three components that help parents and pre-K teachers create safe, nurturing, and predictable environments for children:

- professional development for pre-K teachers and leaders on culturally responsive family engagement and social-emotional learning;
- a parenting program for families of pre-K students; and
- a social-emotional learning program in pre-K classrooms.

ParentCorps aims to place culture at the center of each program component by honoring every family's culture as important and adaptive.[14] Discussions and activities elicit cultural and contextual themes as they relate to parenting and child development. At the start of the 14-session parenting program, for example, parents share their cultural values and beliefs and contemplate what has influenced their own parenting choices. Parents then set goals for their children, grounded in a "whole child" view of development (that is, social, emotional, behavioral, physical, and cognitive) and in the context of their cultural values and beliefs. For example, a mother who values respect and obedience may want to help her children feel confident and work hard in school, even when they're frustrated. These culturally informed goals are a focal point of later sessions where parents assess the fit and relevance of each parenting strategy.

The process is collaborative, allowing for the mutual transfer of expertise; parents examine their cultural values and beliefs in response to strategies introduced by the group's facilitator. Driven by the unique characteristics of their children, families, and contexts, parents make their own decisions about whether and how to use "the science of parenting" (that is, cross-culturally robust strategies linked to positive child outcomes). The facilitator supports the parents' autonomy to do so. Each strategy session is introduced through a consistent structure, including evocative questions such as, What might your grandmother say about praising children for good behavior? Did your parents or other important adults play with you when you were a child? What would they think now if they saw you down on the floor playing? Facilitators invite parents to express skepticism, but also encourage them to consider whether each new strategy could help them meet any of their goals or handle certain situations. This approach can lead parents to open up to a strategy they'd perceived as being at odds with a prominent value. The developers hypothesized that placing culture at the center would support parents' participation in further sessions and skill practice at home, and increase the extent to which parents find the sessions relevant and respectful—and see the strategies as helpful in reaching their goals.[15]

For 10 years, Brotman and colleagues implemented, tested, and improved the original version of ParentCorps for families of pre-K students living in historically disinvested New York City neighborhoods. Studies involved a pilot with 40 families in partnership with leaders, practitioners, and community members from the Harlem Children's Zone; a randomized controlled trial involving 171 families in eight elementary schools in a Brooklyn community school district, with a short-term follow-up; and a randomized controlled trial with 1,050 pre-K families in 10 elementary schools from two community school districts in Brooklyn, with follow-up studies through the transition to middle school.[16]

These studies found that ParentCorps worked as intended: it promoted self-regulation in early childhood by

strengthening parents' and teachers' capacity to support children's skill development.[17] Specifically, ParentCorps improved important aspects of the home and classroom environments, leading to increased knowledge and use of effective practices (such as positive reinforcement and setting clear expectations) and more nurturing adult-child interactions. It also strengthened family engagement as perceived by both parents and teachers.[18] Through its impact on social-emotional development in pre-K, the ParentCorps program helped prevent mental health problems, including both emotional and behavioral problems at school, through second grade.[19] It also led to improved academic achievement by the end of kindergarten, an impact that lasted at least through second grade.[20] In addition, among children who entered pre-K without strong behavior-regulation skills, ParentCorps reduced early behavior problems and prevented the development of obesity and unhealthy behaviors through second grade.[21] Long-term follow-up showed that ParentCorps even reduced chronic absenteeism from third through sixth grades. Based on implementation costs and outcomes from the two randomized controlled trials, a study found that in high-poverty, urban schools, compared to standard pre-K programming, ParentCorps saved $4,387 per student and substantially increased each individual's quality-adjusted life expectancy.[22]

In 2009, Spring Dawson-McClure, a clinical psychologist and prevention scientist, partnered with Brotman and Calzada to incorporate into ParentCorps new basic science findings in the areas of children's eating habits, physical activity, and sleep. The revised program was piloted in six high-poverty elementary schools with pre-K programs. A study with 91 families found that the revised program had the same positive impacts on parenting practices and child behavior found in previous studies, and the study also suggested new areas of impact: on child nutrition and physical activity knowledge, preferences, and health behaviors, including sleep health.[23] By 2014, ParentCorps implementation leaders had developed a portfolio of user-friendly, culturally relevant materials to support fidelity and efficiency and to shift responsibility for facilitating the ParentCorps programs to teams based in schools (mental health professionals, teachers, and parent support staff).

Many factors—including impact on parent and child outcomes, evidence of cost effectiveness, and a strong history of working collaboratively to help pre-K programs implement ParentCorps—provided a convincing rationale to invest in scaling ParentCorps to enhance pre-K programs in historically disinvested neighborhoods. In 2015, the New York State Office of Mental Health and several family foundations with a deep interest in scaling early childhood EBIs committed to fund the scaling of ParentCorps in New York City's newly expanded universal pre-K programs. The ParentCorps team hoped that the approach to scaling ParentCorps would serve as a model for other cities in the state and across the country.

Context for Scaling

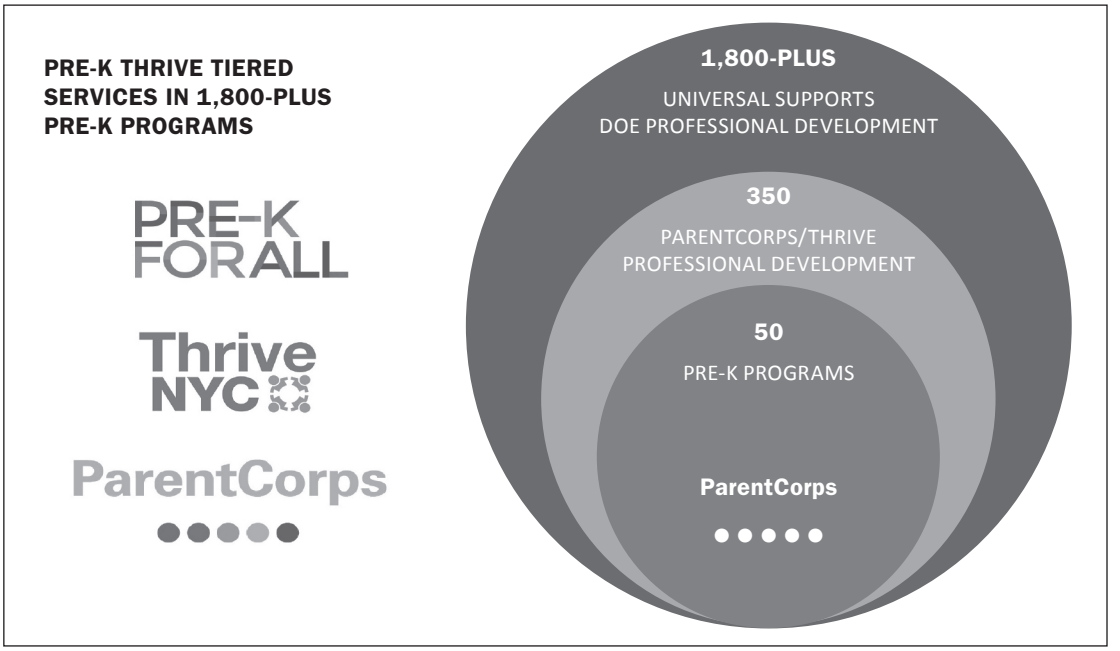
New York City's DOE is the largest school district in the country, serving more than 1.1 million students. About three-fourths of these students are economically disadvantaged, and more than 80 percent are children of color. Twenty percent are students with disabilities, and 14 percent are multilingual

learners or English language learners. The DECE launched Pre-K for All in 2014, relying on research showing that high-quality early childhood education helps children get a strong start in school and in life. Since then, the city has more than tripled the number of children in free, full-day pre-K; Pre-K for All now serves nearly 70,000 four-year-olds annually. About 40 percent of the 1,800-plus pre-K programs that make up Pre-K for All are in elementary schools, with the rest in community-based organizations (CBOs). DECE sets policy and quality standards for all pre-K programs, whether they're in schools or CBOs, and uses a centralized data system and procedures for enrollment and assessment of program quality. In addition, DECE supports the pre-K programs with a large, centralized workforce of early childhood social workers and instructional coordinators, and a comprehensive system of professional development for teachers, principals, and program directors.

The DOE's large investment in universal pre-K, and its commitment to using data to guide policies and practices, created an ideal context for RPPs to drive high-quality programming for children and their families. In 2015, the Mayor's Office established ThriveNYC, a citywide mental health initiative across 12 city agencies, including the DOE. Consistent with ThriveNYC's key focus on early childhood and EBIs, the DECE received multiyear funding to procure a vendor to provide evidence-based services and resources that would promote family engagement and social-emotional learning in Pre-K for All. Brotman and her colleagues proposed using an RPP to study the implementation and impact of scaled services, and to improve policies and practices systemwide.

Responding to DECE's urgent need to increase support for family engagement and social-emotional learning in pre-K, the ParentCorps team designed a tiered

Figure 1. ParentCorps Three-Tier Nested Model



service delivery model that was grounded in ParentCorps's evidence and its approach to behavioral change. Over three years, the three-tier nested model (see figure 1) would unfold as follows: tier 1, "fun with feelings" products to support social-emotional learning for all 70,000 families of pre-K students and professional development for DECE social workers, implemented in all 1,800-plus pre-K programs; tier 2, an aligned professional development series for early childhood teachers and leaders, implemented in 350 pre-K programs in high-poverty areas; and tier 3, the full ParentCorps model, implemented in 50 pre-K programs (selected from the 350 in tier 2). In 2016, Brotman and her colleagues were awarded the contract.

The Research-Practice Partnership Approach

The Pre-K Thrive contract provided three years of public funding to scale ParentCorps (the contract was later extended for three more years, through 2022). This funding was necessary, but not enough to ensure ParentCorps's successful and sustainable implementation across the school district. As described in a recent report by the SPR MAPS Task Force, other critical factors included the degree to which:

- the public system enacts policies (that is, statutes, regulations, and guidance) requiring or recommending EBIs;
- leadership and community stakeholders support EBIs;
- EBIs are ready for scale-up;
- there is a skilled workforce capable of delivering EBIs; and
- the system and/or partners have capacity to support implementation, data monitoring, and evaluation.[24]

The RPP focused on each of these factors at different phases.

The first six months of the contract involved establishing the RPP and collaborative planning for scaling. This meant ensuring alignment with DECE policies, establishing buy-in from DECE leadership at all levels, and understanding the logistics and nuances of the education system that might impact delivery and adoption. The RPP created work groups with members from ParentCorps and DECE (and sometimes from the DOE's Research and Policy Support Group) to establish shared understanding of implementation activities and develop ways to measure accountability. The work groups would also design and plan a series of evidence-building activities, including three randomized controlled trials that included more than 175 pre-K programs in schools and early education centers (EECs) (see figure 2).

Drawing from two decades of collaboration among researchers, practitioners, and community members, the ParentCorps team relied on a set of five principles to guide its involvement in the RPP work with the DECE (see table 1). Below, we illustrate how we applied these guiding principles in a set of interrelated projects that all took place in EECs in CBOs.

The Challenge

Brotman and colleagues rigorously tested ParentCorps in pre-K programs in public elementary schools in historically disinvested neighborhoods. We have strong

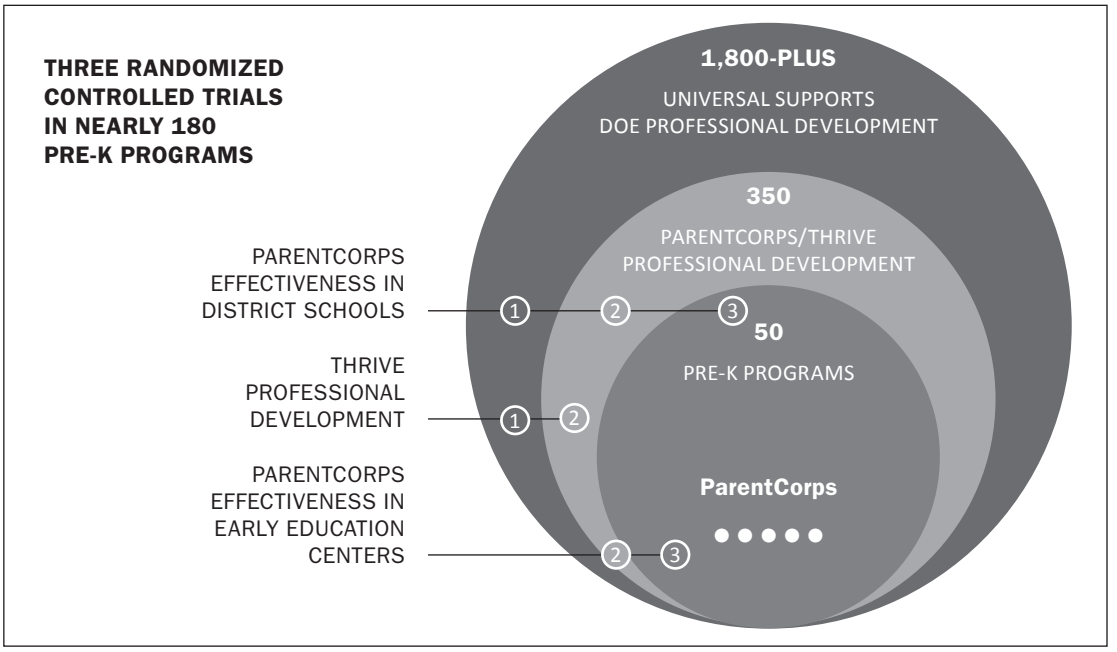
Table 1. ParentCorps's Five Principles for RPP Work

1. ParentCorps is committed to racial equity.	As an organization, we take a population health perspective and strive to interrupt internalized and systemic racism in all aspects of our work. We are committed to investing in our team's own professional and personal development so that all team members are equipped to be effective racial equity leaders and population health advocates. We commit to taking an anti-racist approach with all aspects of ParentCorps planning, implementation, and research, in order to advance racial equity as well as population health.
2. ParentCorps is committed to centering parent voices, especially the voices of parents of color.	We recognize the importance of engaging parents as key stakeholders in all aspects of our growth, including the initial planning phase with new partners. The centering of parent voices is essential for a culturally responsive, family-centered program that seeks to reach and support parents experiencing adversity related to poverty, racism, and discrimination—both to achieve scaled impact on child outcomes and to begin addressing structural inequity.
3. ParentCorps is committed to continuous learning, improvement, and innovation.	To achieve scaled impact at the population level, we must create opportunities for measuring implementation and impact, fully digest and process what we learn, and improve ParentCorps programs, strategies, products, and processes. In addition, we are uniquely positioned to advance the field by describing the inputs of successful RPPs and disseminating our key findings.
4. ParentCorps is committed to understanding the needs and priorities of school leaders, teachers, mental health professionals, and other school staff.	We believe that principals are experts on their schools and teachers are experts on their classrooms, just as parents are experts on their children. We recognize that educators bring their whole selves to their interactions with children and families, and have their own social and emotional needs. We are committed to honoring the voices of these professionals and supporting them as needed to forge meaningful relationships with families.
5. In advancing new scaling relationships, ParentCorps is committed to value alignment and careful systems thinking.	In considering whether to engage with a new public system, we will carefully consider our values and lessons learned from past scaling efforts. For example, are the key elements for success (such as an established mental health workforce) in place, or is there willingness to build capacity? To what extent is the system committed to racial equity, family voices, and continuous learning? Lastly, we will critically consider how to embed ParentCorps in the larger ecosystem of policies, quality standards, and budgets, and we will focus on providing programs with a real potential for sustainable implementation at scale.

evidence for ParentCorps's impact in school-based pre-K programs, and the ParentCorps team has extensive experience in designing, testing, and improving implementation supports for elementary schools. Before submitting our proposal for scaling the full ParentCorps program in schools, the ParentCorps team made sure that all its components (professional development, parenting program, and social-emotional learning program) aligned with the schools' contexts, including staffing, finances, technologies, policies, and practices. For instance, all the schools employed a full-time parent coordinator to engage families, and the vast majority of high-poverty elementary schools employed at least one full-time mental health professional. These individuals, who participate in multiyear professional development (including one-on-one coaching) along with pre-K teachers, are responsible for key aspects of ParentCorps's implementation.

Though ParentCorps was designed and tested in schools, in the early planning phase

Figure 2. ParentCorps Randomized Controlled Trials



of the RPP, DECE leaders required that ParentCorps also be offered in EECs in CBOs, which served nearly 60 percent of the families in pre-K programs. Early discussions about resource allocation were guided by principle 1, *commitment to racial equity.* This was especially important as the ParentCorps team grappled with strongly held values for both equity and evidence. On the one hand, it was an opportunity to expand our reach to families in CBOs. And given that children in CBOs are more likely to be from immigrant families, and that teachers are more likely to be women of color who are paid less than their counterparts in schools, it was also a potential opportunity to counter systemic racism and the marginalization of educators and families. On the other hand, considering the intervention's theory of change, it wasn't a straightforward process to apply ParentCorps's evidentiary foundations, developed in elementary schools, to CBOs. Specifically, the theory of change says that long-term impacts on children's mental health and academic achievement are expected to result from sustained changes in children's social-emotional skill development and self-regulation, parents' involvement in children's learning, and parenting more generally. The question of whether changes during pre-K are robust enough to be sustained through the transition into kindergarten could have a different answer for families entering kindergarten from a CBO, as compared to families continuing in the same elementary school building, with trusted pre-K teachers down the hall and a group of supportive parents who had developed relationships during the previous year. Indeed, the potential for sustained relationships is one reason that ParentCorps focused on schools. Similarly, the loss of supports at the transition to kindergarten has been suggested as an explanation for the fadeout of gains made in pre-K, seen by researchers in several early childhood intervention trials.[25] ParentCorps researchers wanted DOE partners to have a

realistic appraisal of this possibility, and to then collaboratively consider a full range of evidence-building options.

The decision to scale ParentCorps in EECs raised some critical questions for the RPP: What additional adaptations or implementation supports might be needed to build staff capacity in CBOs, especially related to facilitating the parenting program (typically done by a mental health professional based in a school) and outreach to parents (typically done by the school's parent coordinator)? Does ParentCorps produce meaningful family and child outcomes beyond the pre-K year after children transition to kindergarten in a different building?

RPP Solutions

Guided by our fourth principle, *commitment to understanding the needs and priorities of school leaders, teachers, mental health professionals, and other school-based staff*, the ParentCorps team honored the needs and priorities of DECE leadership. In collaboration with DECE, we set out to more fully understand how to support evidence-based policies and practices in the context of EECs in historically disinvested neighborhoods. Over four years, the RPP:

1. adapted the delivery model to fit the EEC context;
2. created and tested new implementation supports, including testing new parent outreach strategies;
3. conducted a qualitative study of EEC pre-K teachers to better understand their own social and emotional needs; and
4. designed and carried out a hybrid implementation-effectiveness randomized controlled trial in 23 EECs, with follow-up of children and families through the end of kindergarten.

Adapting the delivery model to fit the EEC context. Because mental health professionals were needed to facilitate the ParentCorps parenting program in EECs, DECE committed to allocating social workers from its centralized workforce. Previously, DECE social workers had facilitated the program only in the rare cases when school-based mental health professionals weren't available. Allocating social workers to the parenting program not only solved a critical problem, it also held promise for a sustainable solution and further institutionalization of ParentCorps throughout Pre-K for All. At the same time, it raised a new set of challenges for the RPP. Reallocating a subset of social workers meant that DECE was burdened with managing a new role for mental health professionals who must shift between multiple responsibilities to support the entire pre-K system. In addition, using a centralized workforce required a host of new processes for communication and decision-making to ensure that pre-K programs had the right facilitator, at the right time, who was fluent in the languages that met the needs of most parents. All this required system-level adjustments to training, supervision, and monitoring among ParentCorps coaches, DECE supervisors of social workers, and pre-K program leaders.

To conduct the role assumed in schools by parent coordinators, the RPP created the functional position of "ParentCorps champion," which could be carried out by a range of people working in EECs (such as pre-K teachers, administrative staff, or family

service workers in Head Start programs). Among the ParentCorps champion's responsibilities were coordinating logistics for the parenting program (determining the best time and the primary language for families), ordering food from local vendors, and managing reimbursement for program expenses paid by DECE. The champions also led outreach efforts, inviting parents to participate in the program.

These adaptations were complicated and time-consuming, but they produced a more comprehensive and inclusive model for ParentCorps delivery that can be applied to a range of settings in New York City and elsewhere.

Creating and testing a package of parent outreach strategies. Studies of ParentCorps have shown that participation in the parenting program is intricately tied to achieving positive parent and child outcomes.[26] Therefore, the success of ParentCorps in settings like EECs required parent outreach strategies that were effective and culturally relevant. To improve these strategies and ensure that they fit with EEC settings and could be easily used by a range of ParentCorps champions, the ParentCorps research and implementation leaders partnered with three experts in behavioral economics (BE) at NYU: Lisa Gennetian, Zoelene Hill, and Michelle Spiegel. Guided by principle 2, *commitment to centering parent voices*, the RPP team valued the BE framework because it does center family voices, and it helped us understand the experiences of families of color who are facing stress from poverty, racism, and discrimination. Specifically, the BE framework considers parents' in-the-moment decision-making in the context of their lived experiences. The RPP anticipated that BE might also offer insights on low-cost strategies to increase participation in the parenting program.

With principle 3 in mind (*commitment to continuous learning, improvement, and innovation*), the RPP used an iterative process to translate BE concepts into outreach materials that incorporated input from parents and practitioners.[27] For example, new materials, including a "Real Talk" brochure insert and a new tagline—"Together We: Parent. Share. Learn. Grow."—were designed to reduce stigma that parents may perceive related to accepting parenting support. ParentCorps's commitments to racial equity and centering parent voices (principles 1 and 2) were important themes in the design and testing of the BE-infused outreach materials. The RPP conducted a randomized experiment to test the feasibility of the new outreach package and to estimate its impact relative to ParentCorps outreach materials and strategies.[28]

Centering the voices of EEC teachers of color. To better understand the social-emotional needs of teachers in EECs, the RPP, led by Vanessa Rodriguez of NYU (a former NYC school teacher and qualitative researcher), carried out in-depth cognitive interviews with 18 pre-K teachers—most of them women of color—from 10 EECs.[29] The interviews explored the teachers' social-emotional awareness through a developmental perspective. Data analysis used Rodriguez's "Five Awarenesses of Teaching Framework" and identified three key themes.[30] First, the cognitive capacities relevant to teachers in EECs were highly consistent with those of other teacher populations. Second, we found an underlying conflict between teachers' keen awareness of their students'

social-emotional development and active suppression of their *own* social and emotional wellness. Third, most of the teachers believed that their own race and ethnicity weren't as important as those of their students; nearly half denied that their race or ethnicity had any influence on their teaching process.

To enhance professional development and other resources for teachers, our RPP continues to explore the findings that teachers suppress their own emotional needs for the perceived benefit of their students' social-emotional learning, and that they fail to acknowledge how their own racial/ethnic identity and racialized lived experiences influence their teaching practice.

Conduct a hybrid implementation-effectiveness randomized controlled trial in EECs. To answer critical questions about ParentCorps's implementation and impact in EECs, the RPP designed a hybrid implementation-effectiveness RCT that we carried out in 23 centers. As in previous RCTs and implementation experiences, the ParentCorps team expected to achieve replication of impacts for parents and children by the second year, after a year of coaching to support high-quality implementation and integration within the centers and classrooms.[31] Therefore, in the first year the study focused on documenting and monitoring implementation and testing aspects of the outreach model for engaging families in the parenting program, given the new setting and roles for outreach detailed above. In the second year, we continued to assess implementation and conducted the qualitative study of teachers described above. We also enrolled a cohort of children and families to follow over time so we could assess the impact of ParentCorps on these factors: children's learning and development, the use of evidence-based parenting practices, the quality of parent-teacher relationships, parents' involvement in their children's learning, and parents' wellbeing. This study followed 323 families across 19 centers (originally 23 centers, but one closed in the first year of implementation, and the leaders of three others declined to participate). We tracked the children from pre-K through kindergarten, and continued to collect administrative data through fifth grade.

The design and conduct of this study required the RPP members to work together to solve a range of challenges, including:

1. randomization design within the context of the services contract;
2. engagement of EECs from a pool of leaders participating in professional development (tier 2) and joint communication from the RPP;
3. commitment from the DECE to prioritize study centers for systematic classroom observations repeated at meaningful intervals, given the service provision and study time lines;
4. development of culturally responsive recruitment materials and strategies that represented the RPP and study purpose in a transparent and autonomy-supporting manner; and
5. development of a culturally relevant, strengths-based assessment battery for teachers, parents, and children that was feasible and sensitive to intervention.

Here, too, the guiding principles of commitment to racial equity and centering

parent voices were particularly important as the RPP team sought solutions for each of these challenges.

Applying the experiences gained and lessons learned from this collaborative RCT, the RPP designed and executed a three-arm RCT in 80 elementary schools in historically disinvested neighborhoods (comparing the three nested tiers of the service delivery model; see figure 2). A primary aim of this ongoing study is to test variation of implementation and impact on pre-K teachers, classrooms, families, and children (from pre-K through fifth grade) when implemented at scale and in more diverse populations.

The RPP and COVID-19

When the COVID-19 pandemic began in early 2020, the RPP was well established after four years of collaborating, shared decision making, compromising, and navigating big and small challenges. Grounded in our guiding principles, we'd already had many conversations about race, including personal and institutional racism, and were experienced at working toward centering the voices of families of color in resource allocation, program improvements, and evidence-building. The RPP work groups and leadership had learned to adapt nimbly in the context of a large, complex system. We were able to use the partnership to extend ParentCorps to support children, families, and school staff during the pandemic, and to gather data to support the ongoing crisis response.

Facing tremendous uncertainty in the first US epicenter of the pandemic, our RPP focused on the most immediate and critical needs of families and educators leading up to and through the abrupt, unprecedented school closures. We moved rapidly to virtual delivery of professional development for pre-K teachers (tier 2) and the parenting program (tier 3, operating in more than 50 pre-K programs), and we did our best to foster connection, community, and support when it was most needed. The ParentCorps team expanded the portfolio of culturally responsive social-emotional learning and family engagement products, including new tools to help families and teachers discuss the pandemic together.

To plan for the new school year, we adapted ParentCorps to the many challenges and traumatic experiences faced by families of color in historically disinvested neighborhoods. For example, the ParentCorps team drew on the core elements of our rigorously tested parenting program to create a virtual program, Parenting through the Pandemic, with four sessions that focused on grief, loss, and cultivating predictability for children in an unpredictable world. The RPP developed a plan for the ParentCorps team to train and support the DECE's early childhood social workers to deliver this virtual program to families in pre-K programs in the hardest-hit neighborhoods. We're now working on evidence-building strategies to assess the value of these responsive programs and approaches.

When the pandemic hit, the RPP was collecting research data from teachers and parents as part of two RCTs. After an initial pause, we opted to keep collecting data by phone and online surveys, with modifications to express care and offer support. Thoughtful engagement produced three major advances. First, our standing RPP meetings gave us the opportunity to share themes from parents' spontaneous descriptions of their experiences with remote learning, job loss, and illness—

giving the DECE insights to guide its myriad decisions through the early months of the pandemic. Second, collecting data remotely allowed us to contribute to the science on early childhood adversity and the pandemic's enormous physical, emotional, and educational impacts on children, families, and educators. Third, the RPP is poised to offer policy and practice implications to school leaders across the country who are striving to address children's social-emotional learning needs and family engagement. This work is more important than ever, as communities of color face disproportionate illness, death, stress, and trauma from the pandemic, and as the country reckons with centuries of state-sanctioned anti-Black violence.

Next Steps

The ParentCorps team is made up of intervention developers, implementers, and researchers who have participated in different types of partnerships. As such, we recognize the many advantages, challenges, and compromises of scaling within an RPP. Through this multiyear process, which managed to unite interested parties who are usually disconnected from one another, we at ParentCorps individually and collectively strengthened our commitment and skills necessary to center the voices of people of color and promote equity. At the same time, we built an authentic understanding of the constraints under which our public partners operate in large, complex systems that by design are slow to change. Looking ahead,

Table 2. Learning Agenda Themes and Sample Questions

Themes of Inquiry	High-Level Questions
Replication of impact at scale	Are impacts on children's achievement, mental health, and physical health achieved when ParentCorps is implemented independently by schools and pre-K programs at scale?
Unique impacts at scale	With ParentCorps's demonstrated impacts on child development, parenting, and teaching practices, what are its cascading benefits for parents' and teachers' mental health?
For whom?	Consistent with prior evidence, is ParentCorps meeting the needs of children who enter pre-K at the highest risk for difficulties in school? Are families and teachers participating and benefiting comparably across racial and ethnic groups?
How?	What core pathways does ParentCorps use to promote young children's health and development (mapping links from the refined theory of action, which specifies essential program elements and aspects of adult capacity that promote child outcomes)?
In what settings?	What unique strengths and challenges across different pre-K settings may require adaptations to help ParentCorps fit into the context?
Optimizing reach and impact	As ParentCorps offers a portfolio of tools, unbundled program elements, and digital adaptations to meet social-emotional and family engagement needs, what is the uptake, use, and perceived benefit? What innovations are needed?
Ensuring fidelity and quality at scale	What empirical thresholds for implementation fidelity and quality are sufficient for ParentCorps to produce meaningful benefits for children?

we'll strive to build upon our foundational RPP experience: we'll engage in new partnerships in new cities in the United States and globally, and we'll share critical knowledge to help more EBIs successfully scale in large public systems, in the service of children and families. To guide this work, we established a learning agenda with a series of questions, many of which can be answered from the RPP-led activities in NYC and can contribute to advances in implementation and dissemination science in education (see table 2).

In 2020, the ParentCorps team completed a strategic growth plan for scaling ParentCorps nationally. This planning process built on the many lessons we learned through our RPP in New York City, including the nuts and bolts needed to work toward transformative scale in large urban school districts, the importance of engaging key community partners to advance the learning and the need for value-aligned partners throughout the public system, especially in terms of equity, evidence and excellence.

Final Thoughts

A strategic plan is a critical step in RPP-driven scaling. After developing the initial plan together, the partners need to keep thinking strategically throughout its implementation. This means paying attention to the many factors that affect scaling, and adjusting the strategy as necessary. The first version of the RPP scaling strategy can quickly become obsolete as circumstances change, or because of the emergence of factors you didn't adequately consider in the planning process. But the initial plan, with its shared set of understandings and guiding principles, is the foundation for the adjustments to be made intentionally by the partnership as scaling proceeds.

The RPP scaling of ParentCorps, conducted in this country's largest school district and early childhood education system even during an unprecendented pandemic, provides an example of how this approach can advance opportunities and outcomes for children and families of color from historically disinvested neighborhoods.

Endnotes

1. Kenneth A. Dodge, "Toward Population Impact from Early Childhood Psychological Interventions," *American Psychologist* 73 (2018): 1117–29, https://dx.doi.org/10.1037/amp0000393.
2. Abigail A. Fagan et al., "Scaling Up Evidence-Based Interventions in U.S. Public Systems to Prevent Behavioral Health Problems: Challenges and Opportunities," *Prevention Science* 20 (2019): 1147–68, https://doi.org/10.1007/s11121-019-01048-8.
3. Ibid., 1148.
4. Dodge, "Toward Population Impact."
5. Virginia Knox, Carolyn J. Hill, and Gordon Berlin, "Can Evidence-Based Policy Ameliorate the Nation's Social Problems?," *Annals of the American Academy of Political and Social Science* 678 (2018): 166–79, https://doi.org/10.1177/0002716218769844.
6. Jennifer L. Brooks, Beth A. Boulay, and Rebecca A. Maynard, *Empowering Practitioners to Drive the Evidence Train: Building the Next Generation of Evidence* (Boston: Project Evident, 2019), https://www.projectevident.org/updates/2019/6/4/empowering-practitioners-to-drive-the-evidence-train-building-the-next-generation-of-evidence.
7. Action: Building Networked Improvement Communities in Education," in *Frontiers in Sociology of Education*, ed. Maureen T. Hallinan (Dordrecht, Netherlands: Springer, 2011), 127–62; Marisa Cannata et al., "Using Teacher Effectiveness Data for Information-Rich Hiring," *Educational Administration Quarterly* 53 (2017): 180–222, https://doi.org/10.1177/0013161X16681629; David A. Chambers and Susan T. Azrin, "Partnership: A Fundamental Component of Dissemination and Implementation Research," Psychiatric Services 64 (2013): 509–11, https://doi.org/10.1176/appi.ps.201300032; Lora Cohen-Vogel et al., "Implementing Educational Innovations at Scale: Transforming Researchers into Continuous Improvement Scientists," *Educational Policy* 29 (2015): 257–77, https://doi.org/10.1177/0895904814560886; Keng-Yen Huang et al., "Unpacking Partnership, Engagement, and Collaboration Research to Inform Implementation Strategies Development: Theoretical Frameworks and Emerging Methodologies," *Frontiers in Public Health* 6 (2018): article 190, https://doi.org/10.3389/fpubh.2018.00190.
8. Guillermo Bernal et al., "Cultural Adaptation of Treatments: A Resource for Considering Culture in Evidence-Based Practice," *Professional Psychology: Research and Practice* 40 (2009): 361–8 (quote p. 362), https://doi.org/10.1037/a0016401.
9. Felipe González Castro, Manuel Barrera Jr., and Charles R. Martinez Jr., "The Cultural Adaptation of Prevention Interventions: Resolving Tensions between Fidelity and Fit," *Prevention Science* 5 (2004): 41–5, https://doi.org/10.1023/B:PREV.0000013980.12412.cd; José Rubén Parra Cardona et al., "Culturally Adapting an Evidence-Based Parenting Intervention for Latino Immigrants: The Need to Integrate Fidelity and Cultural Relevance," *Family Process* 51 (2012): 56–72, https://dx.doi.org/10.1111/j.1545-5300.2012.01386.x.
10. Steven G. Benish, Stephen Quintana, and Bruce E. Wampold, "Culturally Adapted Psychotherapy and the Legitimacy of Myth: A Direct-Comparison Meta-Analysis," *Journal of Counseling Psychology* 58 (2011): 279–89, https://doi.org/10.1037/a0023626; Timothy B. Smith, Melanie M. Domenech Rodríguez, and Guillermo Bernal, "Culture," in *Psychotherapy Relationships That Work: Evidence-Based Responsiveness*, ed. John C. Norcross (New York: Oxford University Press, 2011): 316–35, https://doi.org/10.1093/acprof:oso/9780199737208.003.0016; Ana A. Baumann et al., "Cultural Adaptation and Implementation of Evidence-Based Parent-Training: A Systematic Review and Critique of Guiding Evidence," *Children and Youth Services Review* 53 (2015): 113–20, https://doi.org/10.1016/j.childyouth.2015.03.025.

11. Melissa B. Caldwell et al., "Community Involvement in Adapting and Testing a Prevention Program for Preschoolers Living in Urban Communities: ParentCorps," *Journal of Child and Family Studies* 14 (2005): 373–86, https://doi.org/10.1007/s10826-005-6850-6; Esther J. Calzada et al., "Training Community Members to Serve as Paraprofessionals in an Evidence-Based, Prevention Program for Parents of Preschoolers," *Journal of Child and Family Studies* 14 (2005): 387–402, https:/doi.org/10.1007/s10826-005-6851-5.

12. Jessica Epperly, "UCLA Report Finds Changing U.S. Demographics Transform School Segregation Landscape 60 Years After Brown v Board of Education," press release, Civil Rights Project, University of California, Los Angeles, May 15, 2014, https://civilrightsproject.ucla.edu/news/press-releases/2014-press-releases/ucla-report-finds-changing-u.s.-demographics-transform-school-segregation-landscape-60-years-after-brown-v-board-of-education.

13. Esther Calzada et al., "Early Childhood Internalizing Problems in Mexican- and Dominican-Origin Children: The Role of Cultural Socialization and Parenting Practices," *Journal of Clinical Child & Adolescent Psychology* 46 (2017): 551–62, https://doi.org/10.1080/15374416.2015.1041593.

14. Cynthia Garcia Coll et al., "An Integrative Model for the Study of Developmental Competencies in Minority Children," *Child Development* 67 (1996): 1891–914, https://doi.org/10.2307/1131600.

15. Spring Dawson-McClure, Esther J. Calzada, and Laurie M. Brotman, "Engaging Parents in Preventive Interventions for Young Children: Working with Cultural Diversity within Low-Income, Urban Neighborhoods," *Prevention Science* 18 (2017): 660–70, https://doi.org/10.1007/s11121-017-0763-7.

16. Caldwell et al., "Community Involvement"; Calzada et al., "Training Community Members"; Laurie M. Brotman et al., "Promoting Effective Parenting Practices and Preventing Child Behavior Problems in School among Ethnically Diverse Families from Underserved, Urban Communities," *Child Development* 82 (2011): 258–76, https://doi.org/10.1111/j.1467-8624.2010.01554.x; Laurie M. Brotman et al., "Cluster (School) RCT of ParentCorps: Impact on Kindergarten Academic Achievement," *Pediatrics* 131 (2013): e1521–9, https://doi.org/10.1542/peds.2012-2632.

17. Brotman et al., "Promoting Effective Parenting Practices"; Brotman et al., "Cluster (School) RCT of ParentCorps."

18. Spring Dawson-McClure et al., "A Population-Level Approach to Promoting Healthy Child Development and School Success in Low-Income, Urban Neighborhoods: Impact on Parenting and Child Conduct Problems," *Prevention Science* 16 (2015): 279–90, https://doi.org/10.1007/s11121-014-0473-3.

19. Laurie M. Brotman et al., "Effects of ParentCorps in Prekindergarten on Child Mental Health and Academic Performance: Follow-Up of a Randomized Clinical Trial through 8 Years of Age," *JAMA Pediatrics* 170 (2016): 1149–55, https://doi.org/10.1001/jamapediatrics.2016.1891.

20. Brotman et al., "Cluster (School) RCT of ParentCorps"; Dawson-McClure et al., "Population-Level Approach."

21. Laurie M. Brotman et al., "Early Childhood Family Intervention and Long-Term Obesity Prevention among High-Risk Minority Youth," *Pediatrics* 129 (2012): e621–8, https://doi.org/10.1542/peds.2011-1568.

22. Negin Hajizadeh et al., "Potential Return on Investment of a Family-Centered Early Childhood Intervention: A Cost-Effectiveness Analysis," *BMC Public Health* 17 (2017): article 796, https://doi.org/10.1186/s12889-017-4805-7.

23. Spring Dawson-McClure et al., "Early Childhood Obesity Prevention in Low-Income, Urban Communities," *Journal of Prevention & Intervention in the Community* 42 (2014): 152–66, https://doi.org/10.1080/10852352.2014.881194.

24. Fagan et al., "Scaling Up."

25. Drew Bailey et al., "Persistence and Fadeout in the Impacts of Child and Adolescent Interventions," *Journal of Research on Educational Effectiveness* 10 (2017): 7–39, https://doi.org/10.1080/19345747.2016.1232459.

26. Brotman et al., "Promoting Effective Parenting Practices"; Brotman et al., "Cluster (School) RCT of ParentCorps."

27. Sally Thigpen et al., "Moving Knowledge into Action: Developing the Rapid Synthesis and Translation Process within the Interactive Systems Framework," *American Journal of Community Psychology* 50 (2012): 285–94, https://doi.org/10.1007/s10464-012-9537-3.

28. beELL-ParentCorps, "Supporting Outreach Efforts and Engagement in a Parenting Program to Help Preschoolers Succeed," http://www.beell.org/beell_parentcorps.php; Zoelene Hill et al., "Behavioral Economics and Parent Participation in an Evidence-Based Parenting Program at Scale," *Prevention Science* (forthcoming).

29. Vanessa Rodriguez et al., "With Awareness Comes Competency: The Five Awarenesses of Teaching as a Framework for Understanding Teacher Social-Emotional Competency and Well-Being," *Early Education and Development* 31 (2020): 940–72, https://doi.org/10.1080/10409289.2020.1794496.

30. Ibid.

31. Brotman et al., "Promoting Effective Parenting Practices"; Brotman et al., "Cluster (School) RCT of ParentCorps."

Fast-Response Research to Answer Practice and Policy Questions

Christina Weiland, Jason Sachs, Meghan McCormick, JoAnn Hsueh, and Catherine Snow

Summary

Research-practice partnerships often face a fundamental tension: well-designed, high-quality research takes time, but practitioners and policy makers need answers to pressing questions as soon as possible.

In this article, Jason Sachs, Meghan McCormick, JoAnn Hsueh, and Catherine Snow discuss this mismatch between the tight timelines of educational decision makers and the typically longer timelines of researchers who are pursuing rigorous analyses. They tell us how, in a partnership between researchers and the Boston Public Schools Department of Early Childhood, they've worked to make fast-turnaround research as rigorous as they can, while also conducting longer-term causal studies.

Because policy makers and practitioners typically aren't highly trained in study design and causal inference, a key responsibility for researchers is communicating the strengths and limitations of fast-turnaround work in ways that can be easily understood by their partners—and making it clear that fast-response analyses should be viewed as only one piece of evidence for guiding a decision.

www.futureofchildren.org

Christina Weiland is an associate professor at the University of Michigan School of Education. Jason Sachs is director of the Boston Public Schools Department of Early Childhood. Meghan McCormick is a research associate at MDRC. JoAnn Hsueh is the director of the Family Well-Being and Children's Development Policy Area at MDRC. Catherine Snow is the Patricia Albjerg Graham Professor of Education at the Harvard Graduate School of Education.

The authors wish to thank the Boston Public Schools, Brian Gold, Anne Taylor, the BPS Department of Early Childhood coaches and staff, the BPS Office of Data and Accountability, Michelle Maier, Rebecca Unterman, the ExCEL study team, the Sustaining the Boost study team, Hirokazu Yoshikawa, Nonie Lesaux, Richard Murnane, John Willett, Abt Associates, and the Wellesley Centers for Women. Thanks also to the participants at the Princeton *Future of Children* conference for this issue for helpful comments, particularly from Pamela Morris, Daphna Bassok, Jenna Conway, and Jon Wallace.

Well-executed research-practice partnerships can have many benefits. For the practitioner, partnering with researchers can provide data and analysis to guide important decisions within often-short windows of policy attention; create a narrative of challenges and successes that can bring new district leaders up to speed quickly; and bring in scarce resources and expertise for research. For the researcher, partnering with practitioners can lead to more relevant research questions; access to and buy-in from key groups, such as principals, teachers, and parents; opportunities to assess classroom practice and child development in a way that balances theory and implications for real-world practice; and the ability to conduct studies with generalizable samples that would not be available otherwise. However, as this issue of *Future of Children* illustrates, such partnerships share common tensions, such as navigating the different needs and cultural worlds of researchers and practitioners or balancing the slow pace of rigorous (that is, well-designed and high-quality) research against practitioners' desire for timely information.[1]

In 14 years and counting, our partnership has enjoyed many benefits while managing most of the inevitable tensions. The partnership comprises a continuous relationship between Jason Sachs and the Boston Public Schools Department of Early Childhood, on the one hand, and Christina Weiland (formerly at the Harvard Graduate School of Education and now at the University of Michigan), on the other; project-specific relationships with the Wellesley Centers for Women and research and consulting firm Abt Associates; and, in its latest form, an expansion to include nonprofit research organization MDRC and the Harvard Graduate School of Education. In this article, we describe the history of our partnership and then focus on our most difficult challenge—balancing rigor against timeliness in a large public school district.

Balancing rigor and timeliness has been our central challenge in large part because of the complexity of large educational systems. The Department of Early Childhood manages programs for thousands of prekindergarteners through second-graders in Boston and oversees professional development and other support for hundreds of teachers each year. Its oversight responsibilities include deciding where new prekindergarten classrooms will open, what curricula to use, what professional development model to use, and how teachers will assess student learning. Each program element has many possible directions and could be the subject of a careful, years-long academic study.

But the district lacks the luxury of time. Most of these decisions have to be made within a few months and then quickly acted on and communicated. Thus the district often has to make decisions based on imperfect but quickly accessible evidence or no evidence at all. At the same time, rigor still matters. As education research has shown time and again, correlation is not causation.[2] Moreover, districts need measurement approaches that generate reliable and valid data, and research samples that can be generalized to a study's population of interest. With less rigorous studies, we risk getting the answer wrong because of the limitations posed by research design, measurement approaches, and sampling constraints. Further, Boston—and early education researchers and practitioners more broadly—have benefitted from lengthy,

careful studies that have guided policy and practice. We illustrate here how we balance competing demands for rigor and timeliness by trying to make our fast-turnaround work as rigorous as we can; by communicating the strengths and limitations of that work clearly and loudly to policy makers and practitioners; and by mixing fast-turnaround work with longer-term causal studies so that we get the right answers to our most pressing questions.

Background

The BPS Prekindergarten Program

The Boston Public Schools (BPS) created the Department of Early Childhood (DEC) in 2005 to oversee the City of Boston's public prekindergarten programs for three- and four-year-olds. Then-mayor Thomas Menino and then-BPS superintendent Tom Payzant pushed for a city-funded prekindergarten program because they believed it could better prepare children for school and could help attract families to the Boston Public Schools who might otherwise choose other options. Jason Sachs (a co-author of this article) has led the DEC since its inception. Sachs came from a research background. He held a doctorate from the Harvard Graduate School of Education and, as the director of research at the Massachusetts Department of Early Care and Education, he had used data and research extensively.

From a national perspective, the program was unusual. It was based entirely in the public schools, paid teachers on the same scale as K–12 teachers, subjected teachers to the same educational requirements as K–12 teachers (for example, achieving a master's degree within five years of being hired), and was open to any child in the city, regardless of income. The program grew quickly in its early years, from 750 students in 2005 to 1,206 in 2006, 1,467 in 2007, and 1,900 in 2008.[3] In recent years, the DEC's work has grown to cover curriculum and instruction through second grade. The city has also continued to expand the program, which doesn't yet have enough seats for all children in the district who apply.

Early Use of Research

From its first years under Sachs's leadership, data and research were a key part of the DEC's strategic planning and decision making. Most notably, in the program's first year, the DEC hired the Wellesley Centers for Women, a research institute at Wellesley College, to measure quality in a random sample of its classrooms. The study was meant to be a needs assessment—that is, it was meant to provide baseline data to guide the DEC's efforts to increase quality. The *Boston Globe* displayed the study's findings prominently on its front page under the headline "Boston preschools falling far short of goals, study says: Teacher quality, site safety faulted," followed by "Boston's public preschool and kindergarten programs are hobbled by mediocre instruction, unsanitary classrooms, and dangerous schoolyards, according to a first-ever study of the programs."[4]

The very public nature of those findings spurred the district to sharpen its focus on improving quality. DEC leaders asked experts to review the evidence on preschool curricula broadly, determine which curricula were being used around the district, and find out how teachers and principals viewed these curricula.[5] Based on their findings, the DEC decided to use Opening the World of Learning in its prekindergarten classrooms. This curriculum targets children's early language and literacy skills; each unit also

embeds a social-skills component, in which teachers discuss social-emotional issues and introduce emotion-related vocabulary words.[6] The DEC also chose Building Blocks, an early mathematics curriculum that covers both numeracy and geometry, with a heavy focus on verbal mathematical reasoning.[7] In other studies, both curricula have shown positive effects on children's outcomes, though the evidence for Building Blocks is stronger than that for Opening the World of Learning.[8] Teachers received training in both curricula, as well as bi-weekly to monthly coaching focused on helping them troubleshoot problems with classroom management, differentiating instruction for children with special needs and dual language learners, and implementing the curricula. Importantly, this professional development model matches the science of adult learning and the existing evidence on helping teachers improve.[9] Coaches prepared detailed guides that showed teachers how to implement the two curricula in tandem and also made sure classrooms were well stocked with the many materials and supplies necessary to carry out the curricula as intended by the developers.[10] To improve quality in both prekindergarten and kindergarten and to ensure classrooms had adequate start-up funding, the district also sought accreditation from the National Association for the Education of Young Children (NAEYC).

More than a decade since the program's founding, Boston's structural and programmatic choices remain exceptional among the nation's public programs, which tend not to require that teachers have master's degrees, not to pay prekindergarten teachers on the same scale as K–12 teachers, not to use a proven, consistent curriculum, and not to provide coaching.[11]

Partnership History

Our research-practice partnership (RPP) began in the summer of 2007, following the decisions to implement the new curricula and coaching model and pursue NAEYC accreditation. Our partnership was sparked by an internship. Christina Weiland, the first author of this article, had just completed one year of doctoral studies at the Harvard Graduate School of Education and was interested in quantitative methods and early childhood education, particularly in the universal prekindergarten movement. When she approached Sachs about spending the summer working for the DEC, he decided the price was right (Weiland had outside funding that made her work free). Sachs saw value in having someone map out available data and create a research plan. Weiland spent that first summer on two tasks: developing relationships with district staff and coaches, and learning which BPS departments held administrative data relevant to the prekindergarten program, the quirks of the available data, the program's components, and the DEC's goals and questions. Weiland was supported by her adviser, Professor Hirokazu Yoshikawa, who provided guidance and expertise as she learned the ropes (and who continues to contribute to research in Boston).

At the end of that first summer, Weiland and Yoshikawa prepared a memo listing all the data collected by the district that were relevant to the DEC.[12] They also highlighted study designs that could be appropriate for answering different kinds of questions that the DEC wanted to ask. This memo helped create a blueprint and timeline for the questions that our research partnership would address. Some questions were centered on monitoring progress

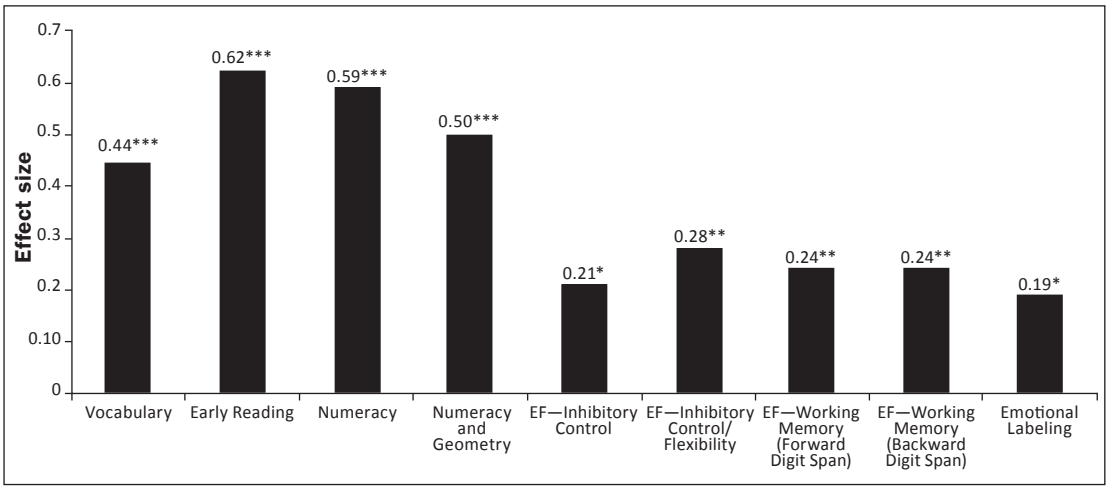
and improving quality and others on determining how the program affected instruction and children's learning. Some could be answered by reviewing previous research, but others required new data and analysis.

One key early issue was whether the program was ready for an impact study and, if so, how to fund it. The district wanted to identify the program's strengths and find out where it could be improved. The research team wanted to learn about how the impacts of a program that was not typical in the national landscape matched or differed from those of other programs. From other studies, we determined that two years of full program implementation was generally considered enough time to determine whether the new model was working. We identified a potential funding opportunity from the Institute of Education Sciences. Because the district wouldn't allow random assignment of students to the program, we looked for an approximation of experimental conditions in the real world (known to researchers as a *natural experiment*) and found one in the program's age cutoff. As in many districts, a child had to turn four by a given date (on or before September 1 in Boston) to attend the prekindergarten program that year. This created a natural experiment, one previously used in several other contexts, for estimating the causal effect of attending the program.[13] Children just one day apart on either side of the age cutoff for enrollment in prekindergarten are equivalent in their background characteristics. Yet the cutoff assigned children born on September 1 to be eligible for prekindergarten one year earlier than children born on September 2. This effectively randomized children who were right at the age cutoff. Provided that all statistical requirements of the design were met, the difference in outcomes between children just at or just below the cutoff would represent the program's causal impact on children's school readiness.

Writing the grant was Weiland's course project in Richard Murnane and John Willett's spring 2008 causal inference class at Harvard. These two respected methodologists agreed to be part of the research team for the grant submission, as did Nonie Lesaux, an expert on dual language learners who is also a Harvard professor; Yoshikawa was the lead researcher. We applied in summer 2008 and received funding to start in spring 2009. The study we conducted included 2,018 children enrolled in 238 classrooms at 67 schools. We found that the BPS model had meaningful impacts on language, literacy, math, and socio-emotional skills—precisely the kindergarten readiness outcomes that were directly targeted by the program. It also had positive impacts on children's executive function skills—that is, their working memory, flexible thinking, and response inhibition skills. The Boston program didn't directly target executive function, but it is developmentally linked to growth in other domains, particularly math.[14] These impacts are shown in figure 1.

RPPs necessarily evolve over time. Today, our partnership has broadened to include the DEC, the University of Michigan, MDRC, and the Harvard Graduate School of Education. The work has extended to include the full prekindergarten-to-third-grade (P–3) span in the Boston Public Schools. As our research aims expanded, we needed to be able to collect and analyze more detailed data.

Figure 1. Boston Prekindergarten Impacts on Children's School Readiness



Notes: EF = executive function; *p<.05; **p<.01; ***p<.001; estimates come from a study of 2,018 children who attended Boston's prekindergarten program in 2008–09.

Source: Christina Weiland and Hirokazu Yoshikawa, "Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills," *Child Development* 84 (2013): 2112–30, https://doi.org/10.1111/cdev.12099.

Partnership Principles

Our partnership today has largely the same goals and philosophy as it did in its early days, even as the DEC's reach within BPS has extended through second grade and as the research team has added institutions. Our goal is simple: we use data and research to improve educational experiences and outcomes for young children. Our principles reflect how we do so:

1. *District questions are the priority.* The research team and the DEC work together to identify priority areas for research. Over the years, the DEC has resisted investigating questions it considered "too academic," that is, likely to answer questions that interest scholars but at a cost of too high a burden on the district without direct benefit. We've learned to select research topics at the intersection of the DEC's practical challenges and unanswered questions among researchers.

2. *The DEC is in the driver's seat in designing its programs.* In our partnership, the research partner's role is not to design interventions or tell the DEC what to do programmatically. We view the practitioners as the experts on implementation and the BPS context. The researcher's role is to support the DEC's agenda by offering insights from the broader scholarship when appropriate, while designing studies to guide quality improvement and key decisions in the BPS context.

3. *Coaches' and teachers' voices matter.* We share research findings first with coaches and then with teachers. We ask for their input on whether the findings seem right to them and on the story behind the findings. For example, why is a given aspect of quality lower than other aspects? How could supports for teachers be changed to improve quality in a given area? From our experience, sharing data and findings also improves

buy-in for research and reinforces a quality-improvement culture.

4. *What you don't do is as important as what you do.* The DEC philosophy is that teachers should focus on teaching. The DEC doesn't ask teachers to rate children's skills the way most US prekindergarten programs do. As part of our partnership work, we reviewed research on such rating systems and found scant rigorous evidence that they provide reliable, valid data or that they change teachers' practices.[15] But they do require a lot of time from teachers. Instead, for research, we largely rely on outside assessors who collect data on a sample of children and classrooms. To guide instruction, teachers do collect short direct assessments of children's language and literacy on well-validated, reliable measures. Because formal assessments don't measure skills like critical thinking and problem solving, the DEC also encourages and supports teachers to regularly collect data on student work and learning processes so they have authentic artifacts of learning to shape both their reflections on their practice (alone and with others) and their teaching strategies in real time.

5. *Trust is fundamental.* Our relationship's longevity has allowed us to develop trust. The DEC trusts that the research team understands the BPS context and the history of the program, won't misrepresent the findings, and will prioritize accurate and balanced dissemination of results. The DEC also trusts that when it would like research guidance for new issues or to guide decision-making, the research team will respond quickly. The research team trusts that the DEC will grapple with the findings and take seriously any implications about changing course.

Overview of Our Work Together

Following these principles, we've collected and used data in a variety of ways.[16] Table 1 offers an overview of the types of data we've used, how frequently and why we collect them, and how we've used them to drive change. The table broadly summarizes these factors so as not to overwhelm readers with every data type and wave of data collection.

Funding and funders' agendas have also influenced our data collection; grants to support programs often require evaluation as well. We've also won several large-scale federal evaluation grants that have funded additional data collection, including work to create and use measures in district classrooms that assess the degree to which curricula are implemented as intended (known as *fidelity*) and to expand child outcome data to important domains beyond language and literacy, such as mathematics and executive function. Our only constant has been administrative data collected routinely by the district, such as child demographics, who enrolls where, teacher characteristics, and children's basic literacy skills.

Our collective knowledge of the BPS data, schools, context, and DEC staff, gained over more than a decade of work, allows us to consolidate different data sources and different findings and to identify what's new. Without an RPP, the district would have had to do more work to launch each new research effort, and less research would have been done. For example, the partnership means that our researchers already know about the district's structure,

Table 1. Types of Data Collected, Frequency of Collection, and Use (Summary)

Data source	When collected	Purpose	Use
Classroom quality and curriculum fidelity observational scores	About every two years	Changes as program evolves; in 2012, for example, data collection focused on K–2 due to concerns about quality of education after prekindergarten	Determine program gaps, needs, and strengths; guide professional development and programmatic decisions
Administrative data	Continuously	Tracking important programmatic data like child attendance, enrollment, demographics, teacher education, certification, and experience	Answer questions about program use and take-up; describe population and how it changes over time; use as control variables in analyses to limit participant burden
Teacher survey	About every two years	Gathering richer data on teacher background, experience of professional development, and opinions/desires related to current offerings	Understand teacher population in more depth; guide professional development and program decisions
Early reading skills and prekindergarten vocabulary	Three times per year by teachers	Monitor children's early literacy and language skill development; identify supports as needed	Describe BPS population; useful as outcomes in evaluation studies
Broader set of child outcomes	When external funding is available or when a research study requires them	Examine children's levels and growth on a broader set of important outcomes, like math, executive function, and socio-emotional skills	Describe BPS population; used as outcomes in evaluation studies

Note: A version of this table appears in Betty Bardige, Megina Baker, and Ben Mardell, *Children at the Center: Transforming Early Childhood Education in the Boston Public Schools* (Cambridge, MA: Harvard Education Press, 2018).

what data it collects, and its core model. Sachs has also had five superintendents and eight immediate supervisors in 13 years. Strong research on the program quickly brought these bosses up to speed, showcased the DEC's strengths, and helped persuade BPS leaders to continue the DEC's direction rather than instituting new reforms and risk "reform weariness" among the district's early childhood teachers. For the researchers, the buy-in and support from the district has been invaluable. Launching large-scale, complex studies would have been harder without a close working relationship with district staff. Also, our research questions have been more relevant to practice and policy because we frequently hear from administrators and practitioners who touch the program daily and we set our research agenda together with the DEC.

Navigating Tension between Rigor and Timeliness

Across projects funded by three large federal grants from the Institute of Education

Science, by small local grants, and by donated time from researchers, we've always had to navigate the natural tension between rigor and timeliness. Rigorous work tends to be painstakingly slow; policy decisions can be head-spinningly fast. The three examples that follow—covering decisions about summer learning, NAEYC accreditation, and professional development for teachers—illustrate how we navigated this tension.

Summer Learning

Research shows that among children from low-income families, growth in academic skills stagnates in the summer, leading to increases in income-based achievement gaps during the summer months.[17] High-quality summer enrichment programs can help stem this problem.[18] Well aware of this research, the BPS offered its own summer reading program for kindergarten and first-grade students. In fall 2010, facing a tight budget, the district considered whether to continue offering this program and, if so, whether to extend it to incoming prekindergarten students. Designing an experimental study for summer 2011 was not an option; the decision had to be made quickly, based on data the district already had. Weiland and Sachs identified data from the 2009 district summer program that could influence the decision (specifically, data on which families chose to send their children to the program, attendance, and student outcomes) and framed the key research questions. Weiland began analyses in mid-November 2010 and worked with Sachs to summarize results by mid-January 2011—a much tighter turnaround than the prekindergarten impact study displayed in figure 1, which took five years from grant writing to publication.

Without an experimental design, we had to solve a central "rigor" challenge caused by student selection into the program. In trying to estimate the effects of the summer program, the risk was that any results, whether positive or negative, could have had to do with characteristics of the students themselves and not those of the program (that is, *selection bias*). For example, imagine that only students who were already strong in reading attended the program—students who might have spent much of their summer choosing skill-building reading experiences on their own, even without the program. Now imagine that we compared their summer gains to those of their peers with weaker reading skills who didn't attend the program. We would find that the stronger students made larger gains than their peers. We would therefore risk concluding that the program was effective, when in fact the stronger students might have shown the same growth even without it.

To help overcome this potential problem, we compared the students who attended the summer program to two other groups of students: 1) students who applied to the program but did not attend; and 2) all other students in the same regular-school-year schools as those who attended the summer program. The students in the first control group were more similar to those who attended summer school, as they were drawn from a pool of those interested in the summer program. The second control group allowed us to compare children who attended the same schools and therefore might have shared background characteristics and classroom reading experiences with students who attended the summer program. Both of these comparison groups were imperfect; because students were not randomly assigned to attend the program or not, we couldn't

definitively conclude that the program itself, rather than other factors, *caused* any student gains. But having two control groups allowed us to examine whether our answer about the program's effectiveness depended on our choice of the control group. If so, our results would have been less trustworthy. If not, we would have greater confidence in our answers to the district.

We found that program attendance was strong. Eighty percent of students had attendance rates of 73 percent or higher, indicating that if it were offered, many district families would send their young children to the program. We also found that the program reached children who needed it more; before the program began, participants had lower literacy skills than their peers, and they were statistically significantly more likely to have repeated a grade. Controlling for their end-of-year literacy scores and background characteristics, students who attended the program showed stronger summer gains in literacy skills than did their peers in either of the two control groups (on one of four measures when compared to appliers and on three out of four measures when compared to all other students in the same schools).

Sachs presented this evidence in district meetings. We were careful to communicate its limitations and also to emphasize other data relevant to the decision, such as feedback from teachers in the program, to ensure that the student-level analysis would be seen as just one piece of evidence and not as decisive by itself. Though it was imperfect, in our view the student-level analysis helped the district make a more informed decision than it could have otherwise. Ultimately, the district decided to continue offering the program and to begin serving entering kindergarteners as well (particularly those who had not attended preschool). As the summer program has continued and matured, it has become an important place for the DEC to pilot new curricula and has been expanded through second grade. The program has served over 3,000 students and is now part of the district's core summer programming.

And our evaluation work on summer has continued. In summer 2019, members of our RPP evaluated the program's effects on kindergarteners and first-graders using a more rigorous randomized trial. We were able to do so because more families applied for slots in the program than there were seats available. Within our relatively small samples (157 kindergarteners and 114 first-graders), we found benefits in reading skills for those who attended the program, particularly in phonics for kindergarteners.[19] We expect that as the district and the nation chart a way forward after COVID-19, our results point to a concrete and proven approach to address summer learning loss.

NAEYC Accreditation

Another critical decision the district faced was whether to pursue accreditation from the National Association for the Education of Young Children (NAEYC) for all district elementary schools. In the early childhood field, NAEYC accreditation is widely considered to be a marker of quality. The accreditation process is intended to improve program quality by ensuring that participating programs meet 10 standards covering four domains: children, teachers and staff, management and administration, and family and community relations.

To become accredited, a program must complete four steps. First, staff members

evaluate whether their program meets NAEYC's 10 standards and then make changes where they deem necessary. In the second and third steps, the program gathers and submits evidence to prove that it meets the standards. In the final step, an NAEYC assessor observes the program and recommends whether it should be accredited. To retain accredited status, a program must submit yearly reports documenting that it is maintaining quality. Program accreditation must be renewed every five years.[20]

Meeting the 10 standards and going through the four-step process is hypothesized to increase classroom quality and thereby improve the developmental outcomes of enrolled children.[21] Yet when we reviewed the research, we found limited empirical evidence that NAEYC accreditation affects classroom quality and child outcomes. Several studies found that NAEYC-accredited programs were of higher overall quality as measured by a commonly used classroom quality measure, the Early Childhood Environment Rating Scale–Revised (ECERS–R), than were programs not seeking NAEYC accreditation. Compared to unaccredited centers, NAEYC accredited centers have a greater proportion of college-educated teachers and staff with degrees or certifications in early childhood education and lower staff turnover rates. Compared to staff in unaccredited programs, staff in NAEYC-accredited programs have also been found to have interactions with children that are more positive.[22] But these studies weren't experimental, meaning that the results may be subject to selection bias—that is, they may be due to some other factor than the NAEYC accreditation process. Further, in some studies, being accredited didn't necessarily ensure high quality; some accredited programs had mediocre scores on an observational measure of quality.[23] Only one prior study had examined the relationship between NAEYC accreditation and child outcomes, and that study was not experimental.[24]

The DEC had two primary reasons to pursue accreditation for its public elementary schools: to improve programs' structural elements so that they would be ready for deeper curriculum and instruction work and to have a tool to begin improving kindergarten. Supporting schools to help them achieve and sustain NAEYC accreditation costs BPS roughly $5,000 per classroom per year in coaching, materials, and facility work and takes three years to complete, on average. This is a considerable investment to make without clear-cut evidence that those resources would best be spent on NAEYC accreditation rather than another need in the district.

Using district data, we examined whether undertaking NAEYC accreditation was associated with higher classroom quality and with larger gains in children's vocabulary skills, comparing early adopters of the approach in the district to other district classrooms. Importantly, schools chose whether to seek accreditation, posing a major threat to rigor via selection bias. It could have been the case that the strongest schools or the schools most motivated to change had chosen to be early adopters. Students in these schools might have shown more growth across time due to factors other than NAEYC accreditation. But an experimental study was out of reach, because of time constraints, cost, and other practicalities. Even so, we aimed to maintain rigor by using reliable and valid classroom measures and including a sample of classrooms large enough to make

some generalizations about the district more broadly. Ultimately, we used data collected by the Wellesley Centers for Women in 2008 as part of the district's biannual progress monitoring program.

From a sample of 119 prekindergarten and kindergarten classrooms, the results overall were positive. For example, on the language and reasoning subscale of the ECERS–R (a commonly used observational measure of quality), classrooms in accredited schools scored 0.55 points higher than non-accredited schools (about half of a standard deviation, which is a relatively large difference in quality). However, we found no association between NAEYC accreditation and another measure of classroom quality, the Classroom Assessment Scoring System (CLASS).

We also found that preschool and kindergarten children in BPS schools that were involved in the NAEYC accreditation process had statistically significantly higher vocabulary score gains from fall to spring than did their peers in other BPS schools. This link remained after controlling for children's vocabulary scores in the fall and characteristics such as race/ethnicity, eligibility for the free/reduced-price lunch program, and language spoken at home. When we controlled for the global quality of the classroom as measured by the ECERS–R, preschool and kindergarten children in BPS classrooms undertaking the NAEYC accreditation process significantly outscored children in other classrooms.

Along with feedback from coaches, principals, and teachers, the district used our analysis as one piece of evidence in making the decision to expand NAEYC accreditation to more district schools. We were careful to explain that our research couldn't identify causal effects and that factors other than NAEYC accreditation might have led to the gains we saw in classrooms that undertook the accreditation process. We did examine whether schools with NAEYC-accredited preschool and kindergarten classroom programs differed from those with unaccredited programs when it came to school-level factors such as third-grade test scores, school size, availability of wraparound services (such as before- and after-school care), principals' participation in an early childhood fellowship program, and the proportion of early childhood teachers with master's degrees. We found no statistically significant differences, though we always underscore that selection bias is a possible explanation for the results of a study like this one. Ultimately, despite the limitations of the research, the district had to make a decision about NAEYC accreditation. Our imperfect evidence was better than none at all.

We've since repeated our NAEYC analyses using data collected in 2010 and 2015 as part of the district's biannual progress monitoring. The analyses using 2010 data largely replicated our 2008 findings. In 2015, however, we found little association between NAEYC accreditation and three separate measures of classroom quality (though we didn't use the ECERS–R in 2015). We discussed these findings with our coaching team, which had expanded considerably since 2008. We found that they were spending much of their time meeting NAEYC requirements and not enough time on instructional quality specifically. These findings led to a shift in the DEC's approach to the NAEYC process; coaches have since placed a high priority on cognitively demanding instruction and tasks for students. Specifically, coaches were asked

to spend at least 50 percent of their NAEYC coaching time on instruction, particularly on supporting teachers' use of cognitively demanding tasks.

The district now has 47 NAEYC-accredited schools (of 80 elementary schools) and has invested $8 million in accreditation. We are once again examining associations between NAEYC accreditation and student gains using data collected in 2017 and 2018. This analysis will be used to make decisions about further NAEYC work in the district.

Professional Development

BPS is part of the IES Early Learning Network, an effort across five states to identify malleable factors at home, in the classroom, and in schools that can increase children's success from prekindergarten through third grade.[25] Like others in the IES network, we've focused on three research questions:

1) What are the district's goals and outcomes for P–3, and which state, local, and district policies either facilitate or hinder this vision?

2) What malleable classroom-level features, processes, and practices can predict within-year gains in students' outcomes?

3) How do students' cumulative experiences in their classrooms, homes, after school, and during the summer influence their P–3 developmental skill trajectories?

Beginning in fall 2016, we recruited a cohort of prekindergarten children to follow through the end of third grade. In the same cohort, we also added a group of children who didn't attend prekindergarten, to be followed beginning in their kindergarten year. We're assessing the children's language, literacy, mathematics, social-emotional, and executive function skills across time. We're also measuring students' classroom experiences each year using CLASS, the Individualizing Student Instruction (ISI) measure, and curriculum fidelity measures; conducting surveys of parents and teachers; attending teacher training; interviewing key district leaders; conducting teacher focus groups; and reviewing relevant district and state documents. Finally, we're using administrative data on students, families, teachers, and schools, both from the district and, for students who leave BPS, from other districts in Massachusetts. Because of the COVID-19 pandemic, we were unable to assess children in second and third grade and now plan to follow them into fourth grade and possibly beyond.

As we conduct rigorous work to answer our primary research questions, we're also aiming to respond quickly to the DEC's need to identify the strengths and shortcomings of its curriculum reforms that aim to align instruction within and across prekindergarten to second grade, and to help teachers improve their practice. These added goals answer two needs. First, while measures like CLASS have helped the DEC improve overall quality, the feedback they offer teachers is somewhat broad and have not been predictive of children's gains. For example, a teacher might be told that language modeling isn't one of his or her strengths. But this guidance doesn't identify a specific weakness (for example: Too many closed-ended questions? Too little elaborative talk? And if so, when? In small groups? In the whole group?). We wanted to pinpoint more specific feedback that teachers could act on. Second, the

research we reviewed offers few specific recommendations for instruction, including how to allocate instructional time.[27] BPS coaches and teachers report that after subtracting transitions, bathroom breaks, lunch, recess, and special activities like music and art, about 3.5 hours are available for instruction out of a 6.5-hour school day. Administrators and teachers alike feel the need to make sure they use this precious instructional time judiciously. How much small group time should teachers have? How long should students be in centers? How long should we focus on a content-specific area to teach skills, and how much time should we allot for integration and synthesis? Furthermore, should our calculations and practices change for children who have experienced poverty and/or students whose first language isn't English? More detailed data on classroom practices could help with these hard choices.

Finally, for about five years the DEC has been working on creating and implementing its own aligned, play-based, interdisciplinary P–2 curriculum supported by district-developed training and coaching. The DEC undertook this work for three reasons: no P–3 curriculum models have been proven by experimental evaluation to improve students' third-grade outcomes; internal data showed that Boston's kindergarten through third-grade classrooms were of lower quality than its prekindergarten classes; and growing evidence suggests that exposure to high-quality learning environments after prekindergarten can help prevent fadeout of the boost children get from attending preschool.[28] The BPS curriculum, professional development, and coaching models are being revised based on lessons learned during their implementation (that is, via an iterative process), and data from the IES study could help the DEC with these revisions.

But progress has been slow in meeting our fast-response goals of identifying the strengths and shortcomings of the P–2 curriculum reforms in real time and giving teachers guidance to help improve their practice. It took time to put our data infrastructure and analysis systems in place. It also took time to build cohesion and understanding among the DEC staff about expectations for curriculum adherence and fidelity and for them to agree on an acceptable level of implementation. Our fidelity measure combined measures created by the curriculum developers, fidelity measures from past research in BPS, and new or revised items for the current project.[29] To maximize accuracy, we aimed to collect at least two two-hour observations per classroom.

In August 2017—about a year after data collection began, about three months after fidelity data were collected, and just before our cohort was to start kindergarten—we sent our first results memos to the district. These were essentially a set of means and frequencies of the prekindergarten curriculum fidelity data, and they came with the caveat that the work was very preliminary. To make the data comparable across classrooms and usable to teachers and the district, we undertook an iterative process to identify key constructs for assessing fidelity and to examine the measurement properties of those constructs. We also discussed extensively how to give data back to teachers. Ultimately, for each curriculum component, our measure captured *dosage* (whether a given component was observed and how long teachers engaged in a component),

adherence (the degree to which teachers implemented the curriculum as intended), and *quality* (how well the curriculum was implemented); these were also averaged across components.[30] We also looked for measures that cut across components, grouping items into four constructs that were easy to talk about with coaches and teachers, and for which we found some support in our data analysis: use of rich vocabulary, making connections, scaffolding and differentiation, and building and extending children's thinking.

After an initial analysis of the fidelity data, we first shared key results with coaches. Next, we worked with coaches in large and small groups to figure out the right venue and format for sharing data with teachers. Coaches were worried that teachers with less-than-stellar scores might become more difficult to enlist in improvement efforts. Together, we planned to share the data in September 2018, when our cohort was entering first grade. Sachs would present key study findings to an audience of about 500 prekindergarten and kindergarten teachers at a district-wide training session to start

Box 1. Vocabulary Items from Fidelity Measure Used to Guide Professional Development

- Teacher uses vocabulary words as related to the unit book(s) and/or small group activity.
- Teacher embeds vocabulary in language (from Centers, Intro to Centers, Read Aloud components)
- Teacher defines vocabulary words (from Centers, Intro to Centers, Read Aloud components)
- Teacher uses a variety of vocabulary words that are sophisticated or advanced (from Centers, Intro to Centers, Read Aloud components)
- Teacher is intentional in which vocabulary words are used and how they are defined (from Intro to Centers, Centers, Small Group Read Aloud components)

Box 2. Professional Development Session Descriptions Shared with Teachers

Explicit & Embedded Vocabulary Instruction in Kindergarten Classrooms
In this session, attendees will briefly discuss early childhood vocabulary acquisition before delving into specific opportunities for vocabulary instruction within the K1 and K2 curricula. Through watching classroom videos of successful vocabulary instruction, attendees will become more familiar with best practices for explicit and embedded vocabulary instruction. Finally, attendees will analyze texts and select vocabulary most appropriate for instruction, and conclude the session by drafting preliminary plans for vocabulary instruction during the coming school year.

Making Connections
"Integrated curricula," "holistic approach," and "interdisciplinary learning" describe the instructional practice of making connections. This practice, implemented with quality, consistency, and as part of a classroom's discourse, is correlated with positive outcomes for children's gains in executive function and flexible thinking.

Scaffolding & Differentiating for High Quality Instruction
In this session, teachers will learn and share useful differentiation strategies that address the range of learning needs in the classroom. Teachers will learn appropriate scaffolds for a variety of learners. The session will hone in on scaffolding and differentiating during three key components of the Focus Curricula: Centers, Literacy Small Group and Math Small Group.

Building & Extending Children's Thinking Through Conversations, Questions, and Interactions
This session looks at the teacher's role in helping children think more deeply about their ideas and understanding of the world through the use of conversations and questioning.

the school year. To keep teachers who had weaker reports engaged in this work, we framed the presentation as highlighting strengths and weaknesses for all of BPS, district administrators included. We also worked together to develop individual data reports for teachers, to be shared only with the teachers themselves, consistent with the consent forms teachers had signed.

Coaches also worried that teachers might learn their weaknesses without clear advice about how to improve. Accordingly, they developed professional development sessions guided by our study findings and details of the four constructs (see box 1 for examples of vocabulary items and see box 2 for descriptions of the sessions shared with teachers to guide their selections).

In his presentation to teachers, Sachs started with the good news, particularly that students generally improved from fall to spring and that their gains on many tests averaged above the national norm. Despite these gains, we still saw large differences between white and nonwhite students and between children from low-income and better-off families. He then shared the CLASS findings, which showed teachers' scores for instructional support were substantially lower than for emotional support and classroom organization. This finding is typical nationally, but the data also showed that kindergarten teachers scored lower than prekindergarten teachers on conceptual development and language modeling.[31] Then, Sachs introduced the four cross-cutting fidelity constructs and went over key findings from the fidelity study. He explained that teachers could pick up their individual data reports and discuss them with study team members one on one. He also explained that teachers should use their reports as just one piece of information for guiding their practice in the classroom. And most immediately, they could use it to choose professional development sessions on the following day, where activities would focus on the four fidelity constructs. He invited them to speak with district administrators about the patterns in the data and what steps could be taken to change them.

Our work on the fidelity measure continues. We have analyzed the kindergarten data collected in 2018 and used them to create measures that assess the extent to which children's instruction was aligned across prekindergarten and kindergarten. We also adapted the fidelity tool for use in first and second grade, incorporating adjustments to the curricular components made for older students. In doing so, we aimed to collect a consistent set of measures across grades while also working collaboratively with DEC staff to make the tools relevant to the district's ongoing need for classroom observation and coaching, in addition to using the data for research purposes. We collected the first-grade data in 2018–19 and a subset of second-grade classrooms in 2019–20 before the start of the pandemic. We aim to be as rigorous as possible—careful about sampling and the quality of the measurement—while also being able to share descriptive findings of interest to the district as soon as they become available. Rather than waiting until the end of the project to use the data to guide instruction, we're applying lessons from the prekindergarten and kindergarten years to improve Boston's Focus on Early Learning curriculum now. (The real-world tension in this work—localities need decisions quickly, but building valid,

reliable measures takes a great deal of time—is also reflected in the article in this issue by Amanda Willford and Jason Downer.)

Takeaways

The three vignettes above offer several broad lessons. First, when it comes to tradeoffs between timeliness and rigor, the potential cost for researchers is underappreciated. Meeting a practice partner's needs will almost certainly result in some work that, on its own, isn't publishable in peer-reviewed journals. Yet such publications are essential for academic researchers' careers. As an example, none of the work we discussed in the vignettes has appeared in a peer-reviewed journal. The summer-school study, for example, used an identification strategy that was too weak for economics-oriented journals, and we had too little valid, descriptive observational data on the components and quality of delivery of the program to satisfy educational research journals. The 2008 results of the NAEYC work were published in NAEYC's non-peer-reviewed publication for practitioners.[32] We tried to include the 2010 work in a paper on quality measures in BPS prekindergarten, but peer reviewers told us to cut the NAEYC work from the paper due to concerns that it wasn't sufficiently rigorous.

When it comes to tradeoffs between timeliness and rigor, the potential cost for researchers is underappreciated.

To be sure, the RPP model also can facilitate scholarship. For example, between 2012 and 2020, members of our RPP team published 19 peer-reviewed articles that used Boston data, and four others are currently under review. Our unpublished work for the district helped us by building trust, giving us important insights into how the district makes decisions, and developing good will that supported some of our slower, more rigorous research. But early career researchers in particular need to carefully balance fast-turnaround RPP work with more publishable, slower work. Early career faculty members may not yet have PhD students who can help with data and analysis (and who in turn can benefit from being trained to apply research methods in real-world contexts). Such students require supervision and training, but they can greatly reduce the amount of time faculty spend analyzing RPP data.

Ultimately, fast-turnaround work—which can have some of the largest impacts on partners and programs—carries a time cost no matter how it's managed. Academic/researcher incentive systems generally don't recognize this cost, discouraging researchers from investing in fast-turnaround work. To encourage more scholars to get involved in RPPs, many academic institutions would need to change their evaluation systems to give such work additional weight, particularly for early career researchers for whom the stakes of producing peer-reviewed publications quickly are particularly high. This is even more true for researchers who work with smaller school districts, where there may not be enough classrooms and children for the kind of large-scale randomized trials that are generally easier to find funding for and publish about. We need research on programs and contexts beyond large cities like Boston and New York to better

guide policy and practice, yet too little such work has been done, particularly on rural prekindergarten programs.

Second, funding structures can exacerbate the tension between rigor and timeliness. Most funders want to support either research or programming, but not both. We've cobbled together funding as best we can, sometimes combining several smaller grants into a common evaluation effort. When a program is funded but research isn't, it's much more difficult to design a carefully planned research study that produces rigorous, usable knowledge. To ensure stronger studies and greater learning, programmatic and research funding need to be better intertwined. (See article in this issue by Jacqueline Jones for a perspective on how funders are thinking through the role of philanthropy in supporting RPPs.)

Third, fast-response approaches are better suited for some research questions than for others. Once we decided that the Boston prekindergarten program was ready for an impact evaluation, for example, we rejected a fast-response approach. The impact evaluation and associated extension studies took five years.[33] In that case, we were evaluating the program as a whole, not just a single component. Less intensive fast-response approaches, in our view, are best for guiding decisions about specific model components or policies. Not every element of a program or program policy can be investigated separately via a rigorous randomized trial; programs and children's and family's needs are simply too complex. Moreover, the whole can be greater than the sum of the parts. Investigating each component separately may not lead to the best answers about building the strongest program.

Fourth, in building RPPs with capacity for both fast-turnaround and longer-term work, we advise starting simple. Faster work on questions like "Is NAEYC Accreditation associated with stronger classroom quality and strong child language gains?" helped build trust in our RPP, helped establish how the members of the partnership would get work done, and trained the researcher in using the district's data systems. Many of the questions asked and answered by RPPs may yield unflattering or null findings, and these can lead to adaptations to improve the program as well as revised research questions. It takes time going back and forth to get to the more nuanced questions that are directly relevant to the challenges faced by the school district. We're now asking questions such as: Under what conditions does the prekindergarten boost last?[34] How does quality vary across children in P–2 classrooms? How do we define and measure alignment from prekindergarten to second grade? Simple analysis of basic data helps build the foundation for a strong RPP. The trust developed through the RPP then allows researchers to disseminate their findings regardless of the results, which is imperative for building actionable knowledge and generating broader learning about early childhood education.

Finally, a core goal of fast-response work is to balance speed with the most rigorous approach to generating results. Accordingly, it's important to keep the limitations of fast-response studies front and center and to combine such quantitative analyses with other data, including interviews with teachers and program administrators. Policy makers typically aren't highly trained in study design and causal inference. The researcher bears the responsibility of making clear that quantitative fast-response analyses should

be viewed as only one piece of evidence for guiding a decision.

The Path Forward

Most of the eventual achievement gap between wealthier and poorer children is in place by the first day of kindergarten.[35] Closing the gap requires our collective best thinking on how to create high-quality early educational programs and improve existing ones.[36] For Boston, our path forward has included a strong RPP that has helped shape teacher practice and district policies while contributing to the broader scholarship. After 13 years of our relationship and counting, we believe deeply that despite their inevitable challenges, RPPs in early childhood are essential to chart the way forward.

Endnotes

1. Cynthia E. Coburn, William R. Penuel, and Kimberly E. Geil, *Research-Practice Partnerships: A Strategy for Leveraging Research for Educational Improvement in School Districts* (New York: William T. Grant Foundation, 2013).
2. Richard J. Murnane and John B. Willett, *Methods Matter: Improving Causal Inference in Educational and Social Science Research* (New York: Oxford University Press, 2010).
3. Betty Bardige, Megina Baker, and Ben Mardell, *Children at the Center: Transforming Early Childhood Education in the Boston Public Schools* (Cambridge, MA: Harvard Education Press, 2018).
4. Tracy Jan, "Boston Preschools Falling Far Short of Goals, Study Says," *Boston Globe*, April 7, 2007, A1; Christina Weiland and Jason Sachs, "The Role of Data, Research, and Evaluation in the Process of Change [appendix]," in Bardige, Baker, and Mardell, *Children at the Center*, 195–210.
5. Bardige et al., *Children at the Center*.
6. Judy Schickedanz and David Dickinson, *Opening the World of Learning* (Iowa City, IA: Pearson, 2005).
7. Douglas H. Clements and Julie Sarama, *SRA Real Math, PreK-Building Blocks* (Columbus, OH: SRA/McGraw-Hill, 2007).
8. Mary K. Ashe et al., "Opening the World of Learning: Features, Effectiveness, and Implementation Strategies," *Early Childhood Services* 3 (2009): 179–91; Douglas H. Clements and Julie Sarama, "Effects of a Preschool Mathematics Curriculum: Summative Research on the Building Blocks Project," *Journal for Research in Mathematics Education* 38 (2007): 136–63; Douglas H. Clements et al., "Mathematics Learned by Young Children in an Intervention Based on Learning Trajectories: A Large-Scale Cluster Randomized Trial," *Journal for Research in Mathematics Education* 4 (2011): 127–66, https://doi.org/10.5951/jresematheduc.42.2.0127.
9. Matthew A. Kraft, David Blazar, and Dylan Hogan, "The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence," *Review of Educational Research* 88 (2016): 547–88, https://doi.org/10.3102/0034654318759268; Christina Weiland et al., "Preschool Curricula and Professional Development Features for Getting to High-Quality Implementation at Scale: A Comparative Review across Five Trials," *AERA Open* 4, no. 1 (2018): https://doi.org/10.1177/2332858418757735.
10. Greg J. Duncan and Richard J. Murnane, *Restoring Opportunity: The Crisis of Inequality and the Challenge for American Education* (New York: Russell Sage and Cambridge, MA: Harvard Education Press, 2014).
11. Allison H. Friedman-Krauss et al., *The State of Preschool 2017* (Rutgers, NJ: National Institute of Early Education Research).
12. Weiland and Sachs, "Role of Data."
13. William T. Gormley Jr. et al., "The Effects of Universal Pre-K on Cognitive Development," *Developmental Psychology* 41 (2005): 872–84, https://doi.org/10.1037/0012-1649.41.6.872; Vivian C. Wong et al., "An Effectiveness-Based Evaluation of Five State Pre-Kindergarten Programs," *Journal of Policy Analysis and Management* 27 (2008): 122–54, https://doi.org/10.1002/pam.20310.
14. Christina Weiland and Hirokazu Yoshikawa, "Impacts of a Prekindergarten Program on children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills," *Child Development* 84 (2013): 2112–30, https://doi.org/10.1111/cdev.12099.
15. Jaclyn M. Russo et al., "Examining the Validity of a Widely-Used School Readiness Assessment: Implications for Teachers and Early Childhood Programs," *Early Childhood Research Quarterly* 48 (2019): 14–25, https://doi.org/10.1016/j.ecresq.2019.02.003.

16. For more details, see Weiland and Sachs, "Role of Data."

17. Paul T. von Hippel, , Joseph Workman, and Douglas B. Downey, "Inequality in Reading and Math Skills Forms Mainly before Kindergarten: A Replication, and Partial Correction, Of 'Are Schools the Great Equalizer?,'" *Sociology of Education* 91 (2018): 323–57; Meghan P. McCormick et al., "Can Center-Based Care Reduce Summer Slowdown Prior to Kindergarten? Exploring Variation by Family Income, Race/Ethnicity, and Dual Language Learner Status," *American Educational Research Journal* 58 (2020): 420–55, https://doi.org/10.3102/0002831220944908.

18. Geoffrey D. Borman and N. Maritza Dowling, "Longitudinal Achievement Effects of Multiyear Summer School: Evidence from the Teach Baltimore Randomized Field Trial," *Educational Evaluation and Policy Analysis* 28 (2006): 25–48, https://doi.org/10.3102/01623737028001025; Doris R. Entwisle and Karl L. Alexander, "Summer Setback: Race, Poverty, School Composition, and Mathematics Achievement in the First Two Years of School," *American Sociological Review* 57 (1992): 72–84, http://dx.doi.org/10.2307/2096145; Brian A. Jacob and Lars Lefgren, "Remedial Education and Student Achievement: A Regression-Discontinuity Analysis," *Review of Economics and Statistics* 86 (2004): 226–44, http://dx.doi.org/10.1162/003465304323023778.

19. Gloria Maldonado,and Christina Weiland, "The Impacts of Boston's Summer Focus Program on the Literacy Skills of Kindergarten and First-Grade Students," unpublished manuscript, 2020.

20. National Association for the Education of Young Children (NAEYC), "Accreditation Process for Early Learning Programs," accessed June 20, 2019, https://www.naeyc.org/accreditation/early-learning/process; Marcy Whitebook, Laura Sakai, and Carollee Howes, *NAEYC Accreditation as a Strategy for Improving Child Care Quality: An Assessment by the National Center for the Early Childhood Work Force* (Washington, DC: National Center for the Early Childhood Workforce, 1997).

21. Marcy Whitebook.,"NAEYC Accreditation as an Indicator of Program Quality: What Research Tells Us," in *NAEYC Accreditation: A Decade of Learning and the Years Ahead*, ed. Sue Bredekamp and Barbara A. Willer (Washington, DC: NAEYC, 1996), 31–46.

22. Paula J. Bloom, "The Quality of Work Life in Early Childhood Programs: Does Accreditation Make a Difference?," Bredekamp and Willer, *NAEYC Accreditation*, 13–24; Carollee Howes and Ellen Galinsky, "Accreditation of Johnson & Johnson's Child Development Center," in Bredekamp and Willer, *NAEYC Accreditation*, 47–60; Whitebook "NAEYC Accreditation as an Indicator"; Whitebook, Sakai, and Howes, *NAEYC Accreditation as a Strategy*; Marcy Whitebook, Carollee Howes, and Deborah Phillips, *Who Cares? Child Care Teachers and the Quality of Care in America. Final Report, National Child Care Staffing Study* (Oakland, CA: Child Care Employee Project, 1989).

23. Whitebook "NAEYC Accreditation as an Indicator"; Whitebook, Sakai, and Howes, *NAEYC Accreditation as a Strategy.*

24. Minnesota Department of Human Services, *School Readiness in Child Care Settings: A Developmental Assessment of Children in 22 Accredited Child Care Settings* (St. Paul, MN: Minnesota Department of Human Services, Children and Family Services, 2005).

25. Meghan McCormick et al., *The Challenge of Sustaining Preschool Impacts* (New York: MDRC, 2017).

26. Paola Guerrero-Rosada et al., "Null Relations between CLASS Scores and Gains in Children's Language, Math, and Executive Function Skills: A Replication and Extension Study," *Early Childhood Research Quarterly* 54 (2020): 1–12, https://doi.org/10.1016/j.ecresq.2020.07.009; Christina Weiland et al., "Associations between Classroom Quality and Children's Vocabulary and Executive Function Skills in an Urban Public Prekindergarten Program," *Early Childhood Research Quarterly* 28 (2013): 199–209, https://doi.org/10.1016/j.ecresq.2012.12.002.

27. Nell K. Duke and Heidi Anne E. Meiser, *Phonics Faux Pas: Avoiding Instructional Missteps in Teaching Letter-Sounds Relationships* (Washington, DC: American Federation of Teachers, 2018); Judith A. Schickedanz and Catherine Marchant, *Inside PreK Classrooms: A School Leader's Guide to Effective Instruction* (Cambridge, MA: Harvard Education Press, 2018).

28. Deborah Phillips et al., *Puzzling It Out: The Current State of Scientific Knowledge on Pre-Kindergarten Effects—A Consensus Statement* (Washington DC: Brookings Institution, 2017).

29. Weiland and Yoshikawa, "Impacts"; Monica Yudron, Christina Weiland, and Jason Sachs, "The Importance of Adjusting for Context: Learning from a Pilot Expansion of the Boston Public School Pre-K Model to Community-Based Settings," unpublished manuscript, 2018.

30. Christopher S. Hulleman, Sara Rimm-Kaufman, and Tashia Abry, "Innovative Methodologies to Explore Implementation: Whole-Part-Whole—Construct Validity, Measurement, and Analytical Issues for Intervention Fidelity Assessment in Education Research" in *Applying Implementation Science in Early Childhood Programs and Systems*, ed. Tamara Halle, Allison Metz, and Ivelisse Martinez-Beck (Baltimore, MD: Brookes Publishing, 2013), 65–93.

31. Ajay Chaudry et al., *Cradle to Kindergarten: A New Plan to Combat Inequality* (New York: Russell Sage, 2017).

32. Jason Sachs and Christina Weiland, "Boston's Rapid Expansion of Public School-Based Preschool: Promoting Quality, Lessons Learned," *Young Children* 65, no. 5 (2010): 74–7,

33. Christina Weiland, M. Clara Barata, and Hirokazu Yoshikawa, "The Co-Occurring Development of Executive Function Skills and Receptive Vocabulary in Preschool-Aged Children: A Look at the Direction of the Developmental Pathways," *Infant and Child Development* 23 (2014): 4–21, https://doi.org/10.1002/icd.1829; Weiland et al., "Associations"; Christina Weiland et al., "Early Mathematics Assessment: Validation of the Short Form of a Prekindergarten and Kindergarten Mathematics Measure," *Educational Psychology* 32 (2012): 311–33, https://doi.org/10.1080/01443410.2011.654190; Christina Weiland and Hirokazu Yoshikawa, "Does Peer Socio-Economic Status Predict Children's Gains in Receptive Vocabulary and Executive Function in Prekindergarten?," *Journal of Applied Developmental Psychology* 35 (2014): 422–32; Weiland and Yoshikawa "Impacts."

34. Rebecca Unterman and Christina Weiland, "Quantifying and Predicting Variation in the Medium-Term Effects of Oversubscribed Prekindergarten Programs," working paper, MDRC, 2019; Christina Weiland et al., "The Effects of Enrolling in Oversubscribed Prekindergarten Programs through Third Grade," *Child Development* 91 (2020): 1401–22, https://doi.org/10.1111/cdev.13308.

35. Chaudry et al., *Cradle*; Sean F. Reardon and Ximena A. Portilla, "Recent Trend in Income, Racial, and Ethnic School Readiness Gaps at Kindergarten Entry," *AERA Open* 2 (2016): https://doi.org/10.1177/2332858416657343.

36. Daphna Bassok and Mimi Engel, "Early Childhood Education at Scale: Lessons from Research for Policy and Practice," *AERA Open* 5 (2019): https://doi.org/10.1177/2332858419828690; Christina Weiland, "Commentary: Pivoting to the 'How': Moving Preschool Policy, Practice, and Research Forward," *Early Childhood Research Quarterly* 45 (2018): 188–92, https://doi.org/10.1016/j.ecresq.2018.02.017.

Embedding Causal Research Designs in Pre-K Systems at Scale

Rachel Abenavoli, Natalia Rojas, Rebecca Unterman, Elise Cappella, Josh Wallack, and Pamela Morris

Summary

In this article, Rachel Abenavoli, Natalia Rojas, Rebecca Unterman, Elise Cappella, Josh Wallack, and Pamela Morris argue that research-practice partnerships make it possible to rigorously study relevant policy questions in ways that would otherwise be infeasible.

Randomized controlled trials of small-scale programs have shown us that early childhood interventions can yield sizable benefits. But when we move from relatively small, tightly controlled studies to scaled-up initiatives, the results are often disappointing. Here the authors describe how their partnership with New York City's Department of Education, as the city rapidly rolled out its universal pre-K initiative, gave them opportunities to collect experimental and quasi-experimental evidence while placing a minimal burden on educators.

They argue that this type of research can answer the most pressing ECE questions, which are less about whether ECE can make a difference and more about the conditions under which early interventions are effective at scale. They offer three recommendations for researchers, policy makers, and practitioners who are considering partnership work: build a foundation of trust and openness; carefully consider whether rigorous causal research or descriptive research is the right choice in a given situation; and be flexible, seeking opportunities for rigorous research designs that may already be embedded in early childhood education systems.

Rachel Abenavoli is a research assistant professor of applied psychology at New York University's Steinhardt School of Culture, Education, and Human Development. Natalia Rojas is a postdoctoral research associate at NYU Langone's Center for Early Childhood Health and Development. Rebecca Unterman is a research associate at MDRC. Elise Cappella is a professor of applied psychology at NYU Steinhardt and vice dean for research. Josh Wallack is deputy chancellor for Early Childhood Education and Student Enrollment in the New York City Department of Education. Pamela Morris is a professor of applied psychology at NYU Steinhardt and a faculty affiliate of the NYU School of Global Public Health.

Despite decades of research and substantial public investment in early childhood education (ECE), there is still a wide gap between education research and education practice. What happens in classrooms and school districts isn't always based on the latest research evidence, and research studies aren't always designed to solve key problems of practice or yield information with direct implications for practitioners or policy makers. Although tightly controlled research studies such as randomized controlled trials produce the strongest evidence of programs' efficacy, findings from these studies—though methodologically rigorous—aren't always relevant to fully scaled local systems with broader priorities, unique constraints, and large, diverse populations to serve.[1] Research-practice partnerships (RPPs) are a promising approach to make education research more relevant because they help align the focus of research and the needs and priorities of practitioners and policy makers.[2] They also provide unique opportunities to rigorously evaluate program models at scale and draw clear, causal conclusions about programs' effectiveness in local contexts.

This article describes the development of co-designed studies conducted in the context of a multiyear partnership between researchers at the Steinhardt School of Culture, Education, and Human Development at New York University (NYU) and early education leaders in New York City's Department of Education (DOE) after the city began an ambitious reform: an effort in just two short years to turn a preschool system that served about 19,000 children into one that served about 70,000. In the initial phase of our partnership, our teams developed capacity and infrastructure to monitor the rapid expansion of Pre-K for All, as the program is called. NYU brought existing research evidence and descriptive data to bear on the city's important policy questions, and the DOE shaped the emerging research agenda by identifying priorities, sharing information and data, and building internal capacity for collecting and using data. Building on this critical early work, regular communication across our teams, and a foundation of trust established in the early years of the partnership, we've developed a research agenda to test a core component of the system: the Pre-K for All professional learning (PL) model for supporting program quality. Our jointly designed research studies focus on the city's high-priority questions using a range of rigorous methods that yield causal, actionable evidence about PL and its effects on classroom processes and child outcomes. By working together to rigorously answer questions of high practical relevance as the programs are being implemented at scale, our partnership aims to generate the kind of information that is useful to policymakers and researchers in the city and beyond to best support the learning and development of young children through prekindergarten at scale.

Embedding Rigorous Evaluations in Systems at Scale

Prekindergarten, on average, can improve children's academic and social outcomes, at least in the short term.[3] But its effects vary as a function of program quality. Developmentally appropriate classroom curricula and teacher PL are hallmarks of high-quality prekindergarten programs, and program quality is associated with gains in children's school readiness skills.[4] As in other disciplines, randomized controlled trials meet the highest standards of scientific

rigor and provide the strongest evidence of effective approaches. Several tightly controlled evaluations have shown that specific curricula targeting math, literacy, or social-emotional skills, together with aligned support for teachers, lead to changes in teaching practice and children's skills in the targeted domains.[5] Using randomized trials to document programs' efficacy is a critical step in translating research to practice, but there's a large discrepancy between the magnitude of impacts produced by programs tested in relatively small, tightly-controlled studies and those produced by national, state, or district-wide programs implemented at scale in the real world.[6] This discrepancy may be due in large part to differences in program implementation and support (for example, less frequent or less intense coaching, more variable implementation in the classroom), either because programs are too costly for districts to fully implement across an entire system or because the initial developers lack the capacity to support large-scale dissemination of programs in their original form.[7] The gap between impacts produced in efficacy trials and those achieved at scale could also be due to differences in population (for example, differences among the children served, the teachers doing the implementing, or the coaches supporting them) or even to unintended negative consequences (for example, program quality that is poor or inequitably distributed at scale), especially if rapid scale-up outpaces the capacity to support dissemination.[8]

Given the disconnect between impacts documented in smaller efficacy trials and impacts produced at scale, we need research that can provide clear, causal, and actionable answers about how to create high-quality, large-scale prekindergarten programs. Randomization is the gold standard for evaluating program impacts, but when evaluating fully scaled programs already in operation, the advantages associated with randomization must be balanced against practical considerations.[9] Experimental research designs are challenging to embed in large, complex systems, and randomization may be difficult for a number of reasons, including infeasibility, competing priorities, or concern about withholding potentially effective approaches from sites or children who need them. For example, certain research designs may not be possible or ethical in districts committed to offering prekindergarten to all children or PL opportunities to all teachers, or when program implementation is already well under way. Or districts may decide to target services to specific subgroups or allow individuals to opt in to programs when capacity and resources are limited, rather than mandate policies or specific curricula across a large and diverse system. These choices may represent the best decisions from a policy perspective, but they may limit research opportunities and preclude traditional randomized designs.

An RPP offers a unique opportunity to navigate the challenges of evaluating large prekindergarten systems, embed rigorous research designs at scale, and answer policy-relevant questions.[10] Factors that characterize RPPs—such as frequent and regular communication, a foundation of trust, and commitment to a shared vision and long-term collaboration—pave the way for research and policy partners to find research solutions that align with district priorities and constraints while maximizing the rigor with which research studies are designed and conducted.[11] In contrast to university-led research projects, where practice and policy partners may have only

a small role in planning, research-practice partnerships are well-positioned to recognize hidden opportunities for research that may otherwise be missed, to make small adjustments to existing processes that may be neutral for policy but beneficial from a research perspective, and to quickly identify challenges and their solutions.

Our multiyear partnership between NYU researchers and DOE leaders overseeing Pre-K for All shows how RPPs can maximize methodological rigor when evaluating programs and policies at scale. We've worked in a large, complex system with unique priorities and constraints to take advantage of existing opportunities for research (such as capitalizing on randomization that has already occurred) and embed new opportunities that increase our ability to draw clear, causal conclusions about program impacts (for example, by making slight adjustments to assignment processes). Our recent work seeks to answer policy-relevant causal questions about PL, a key component of the system expected to support program quality and therefore child outcomes. To generate useful information to advance practice and policy, we've sought methodological approaches that could answer our high-priority questions, were feasible to embed in the existing system, and met standards of scientific rigor that would permit causal inferences and contribute to early childhood education science.

History and Evolution of the Partnership

New York City's is the largest school district in the nation, serving 1.1 million students from prekindergarten through 12th grade. Pre-K for All represents a commitment to providing free, full-day, high-quality prekindergarten to every four-year-old in the city, and it is arguably one of the most swiftly and broadly deployed educational policy initiatives in the nation. In just two years, the DOE hired thousands of new teachers and more than tripled the number of children in preschool, from 19,000 students in 560 programs before the 2014 rollout to about 70,000 children per year in over 1,800 sites starting in 2015–16. To expand as rapidly as it did, Pre-K for All built on an existing (though decentralized) public and private early childhood system. Today, Pre-K for All serves about 60 percent of its students through programs in community-based organizations (called New York City Early Childhood Centers or NYCEECs) that contract with the DOE; the remaining programs are in district schools and DOE-created prekindergarten centers. All are held to the same quality standards, are integrated across data platforms for functions such as child enrollment and screening, are assessed similarly for program quality, and are offered the same PL opportunities and onsite coaching.

The NYU-DOE partnership began in the spring of 2014 when the city announced the Pre-K for All expansion. Soon after the announcement, NYU faculty gave a briefing on the state of ECE research to the city, and they were invited to collaborate with the DOE and the mayor's office. NYU then raised resources to capitalize on the first 12 months of implementation as an opportunity to monitor the expansion (for example, through descriptive analyses and data visualization; see below) and build a foundation for data collection, data-based decision-making, and research. At all phases, the NYU team provided technical assistance to the DOE—for example, by supporting the use and interpretation of data and

bringing prior research to bear on issues of policy and practice. At the same time, the DOE involved NYU in developing key decision-making processes and discussing the possibilities and constraints in bringing prekindergarten to scale. Over time, NYU's role shifted from key resource for the city's expansion to partner in conducting research.

In 2014–15, the first year of the partnership, the NYU team supported the DOE's efforts to launch Pre-K for All, helping the city take the pulse of the rapidly developing system. The DOE opened and licensed centers, hired teachers, and reached out to families to tell them about and help them sign up for Pre-K for All. Simultaneously, the DOE expanded its use of data for decision-making, built its own data team internally, and created a system of data collection and review that could support prekindergarten quality. The city also invested in a large-scale study of the program, led by the private research firm Westat, to learn from parents, administrators, DOE staff, and others about how the program was perceived and to get an early read on how children were faring across the city in the new system. NYU worked closely with Westat in this first year, helping to develop and administer tablet-based assessments of children's school readiness. The NYU team also created maps that combined key neighborhood characteristics, the location of Pre-K for All expansion sites, and child assessment data. Presenting this information visually allowed us to identify "hot spots," such as neighborhoods where children were most academically at risk. It also supported quality assurance and helped the DOE deploy its resources strategically. And the city was able to answer such questions as "Do Pre-K for All classrooms and sites vary widely in the extent to which they serve children at greatest socio-demographic, academic, and behavioral risk?" and "Are the gains made by children during their Pre-K for All school year similar to or different from other large urban school systems?" This work together laid the foundation for the partnership.

In the partnership's second year, 2015–16, we made a coordinated effort to align Pre-K for All's quality infrastructure with research-based practice, using methods appropriate for a large and diverse system. NYU reviewed the research to help the DOE develop research-based quality standards and select evidence-based PL tracks or program models that were aligned to those quality standards (we discuss this process in more detail below). We also worked together to construct a data-based decision-making process to assign the sites to tracks and to allocate onsite coaching support (for example, how often sites would receive support and whether the coach would be an instructional coordinator, a social worker, or both). Our two teams communicated regularly via weekly phone calls, periodic in-person meetings, and frequent emails, and our partnership grew stronger.

In 2016–17, the partnership's third year, the NYU and DOE teams worked closely to develop observational and survey-based measures of teacher practice to guide implementation of the PL tracks. The NYU team also received funding from the Foundation for Child Development to provide information on the teacher workforce (see the article in this issue by Jacqueline Jones). As part of this work, NYU examined how teachers with different training and experience were distributed across the city and collected data on teachers' formal PL (training and coaching), informal PL (advice networks), and work climate (satisfaction,

support, and stress) to identify barriers and opportunities for bringing PL to classrooms.

Technical assistance, descriptive information, and data visualization were critical in the initial phase of the partnership, when the city's priorities included monitoring the Pre-K for All expansion and deciding how best to allocate support and resources. But as the city's chief concerns shifted from understanding the Pre-K for All landscape toward testing and strengthening components of the system, our research questions shifted from descriptive to causal. Given that PL is one of DOE's key supports to help Pre-K for All sites improve program quality, we jointly developed a set of research questions that focused on how PL affects teachers and children. We turn now to that work and our efforts to embed rigorous evaluations of PL in the system.

Embedding Rigor to Evaluate Professional Learning

As we said above, in its early phase the partnership developed a differentiated PL system that supports all prekindergarten teachers and leaders. Because needs vary among leaders and teachers and from site to site, the city decided to offer multiple PL options, called "tracks," each year. This approach better meets the needs of a large and diverse population and increases buy-in among program leaders and teachers, who have disparate training, experience, and interests. Each year, each site is assigned to one of several tracks, each with its own theme or focus. The first track to be introduced, in 2015–16, was NYC Pre-K Explore, which integrates the evidence-based math curriculum Building Blocks with interdisciplinary Units of Study that the DOE developed to support children's higher-order thinking and problem-solving.[12] It evolved from discussions in the early phase of the partnership, when the DOE and NYU together reviewed evidence-based models to support children's learning. Three other tracks were introduced in 2016–17:

- NYC Pre-K Thrive, in which sites learn evidence-based strategies to support children's social-emotional development, grounded in research on the evidence-based family engagement intervention ParentCorps;[13]
- NYC Pre-K Create, an arts-based approach that integrates visual art, dance, theater, and music into instruction to promote learning and engagement; and
- NYC Pre-K Inspire, now called Teaching Team Learning Communities, which focuses on topics aligned to the district's quality standards (for example, creating a positive classroom culture, engaging children in meaningful activity) that support instructional goals for early childhood education.

Each year, Pre-K for All program leaders and teachers attend three to four full-day PL sessions tied to their tracks. All sites also get onsite support from an instructional coordinator and/or a social worker; the frequency of visits and the makeup of the support personnel (that is, instructional coordinator, social worker, or both) depends on the site's needs. Explore sites also receive coaching designed to support implementation of Building Blocks and

its integration with the Units of Study, and Create sites work with a teaching artist to support implementation in the classroom.

Today, much of the partnership's work focuses on which approaches are effective and for which outcomes. We're not trying to find out whether prekindergarten "works" relative to no prekindergarten. Research along those lines wouldn't tell the DOE what it needs to know to operate a system in which all four-year-olds are guaranteed a prekindergarten seat and all teachers participate in some form of PL. Instead, we're working to understand the implementation and impacts of different PL approaches with distinct theories of change and targeted outcomes for teachers and children. This focus can not only help the DOE refine and strengthen its system in New York, but also contribute to early childhood education more broadly.

In an ideal world, we might randomize prekindergarten sites to different models of PL, which would give us useful causal evidence about how the models affect the desired outcomes. But with the system already fully scaled and well under way, simple randomization of children or sites to PL tracks isn't always feasible or consistent with the DOE's other principles or priorities. The DOE has prioritized choice in Pre-K for All as a way to meet the needs of its large and diverse population of about 70,000 prekindergarten children in over 1,800 sites. To increase access and equity for all children, families can apply to prekindergarten anywhere in the city, no matter where they live. At the same time, site leaders' preferences are strongly weighted in the algorithm the DOE uses to place sites in PL tracks; this aligns with the DOE's stance that differentiating PL by both needs and interests will maximize the benefits for teachers, and that having a choice in the PL assignment process will increase buy-in and implementation quality on the part of leaders and teachers. This is a sensible approach from a policy standpoint, but given the possibility of selection bias, it makes it harder to conduct research that permits strong causal inference. The strongest test of a program model is one that compares individuals who receive the program with individuals who don't receive the program *who are equivalent before the program begins* (and this is what randomization ensures); if different kinds of sites or families make different kinds of choices about what PL to receive or what prekindergarten program to attend, the assumption that teachers and children are equivalent from the start may not hold. Because traditional randomized designs are at odds with the city's commitment to choice, we've worked together to identify and/or embed randomization in the existing Pre-K for All choice processes, rather than impose a separate research design.

A Natural Experiment: The Explore Evaluation

As the DOE shifted in 2016 from expanding to strengthening the system, we had a series of conversations to map out new research efforts. Together, we decided to focus on how Pre-K for All's PL system affected outcomes at the setting level (for example, what changes do we see in teacher practices and classroom quality?) and to focus on the Explore PL track, the first track launched by the DOE. As we planned this work, we explored potential experimental and quasi-experimental designs (that is, designs in which researchers simulate a randomized controlled trial by identifying treatment

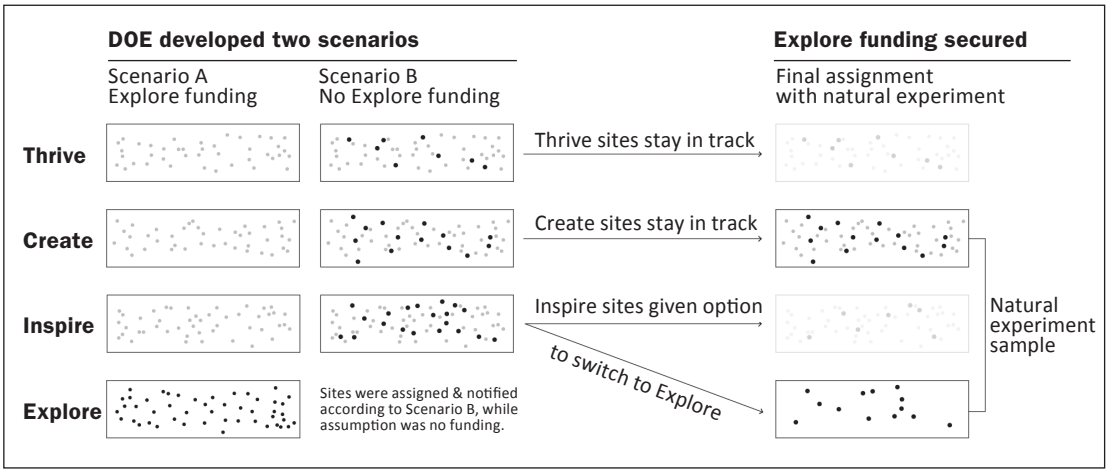
and control groups, even though true randomization didn't occur) that would meet our dual goals of using rigorous methods and honoring the city's emphasis on choice. This emphasis on choice was (and continues to be) reflected in the DOE's PL track assignment process: before the 2016–17 school year, the DOE developed a systematic method for assigning prekindergarten sites to the four PL tracks based on criteria that included the tracks' limited capacity (that is, for certain tracks, there is a maximum number of sites that can be served each year), program leaders' preferences among the PL tracks, and the sites' needs. Sites' needs were determined by factors such as the proportion of children at the site who were from high-poverty neighborhoods or living in temporary housing. Given the selection factors (preference and need) that could contribute to track assignment, it was critical to carefully shape our evaluation to circumvent these factors at the design stage or address them through our sampling or analytic approach.

We first considered the gold standard for causal research—random assignment of sites to PL tracks. Because only a limited number of slots were available in three of the PL tracks (Explore, Create, and Thrive), the DOE's assignment algorithm was designed to randomly assign sites to tracks if certain conditions were met (for example, if the number of sites wanting to participate in a given track and meeting other criteria exceeded the number of slots remaining). We considered whether we could capitalize on this randomization for the Explore evaluation but determined that doing so wouldn't be possible in the 2016–17 school year. Although some randomization occurred, too few sites were randomized for this approach to give us an adequate sample size. In addition, although the process for assigning sites to tracks was data-based and systematic, it was also multidimensional and complex. Thus it didn't easily lend itself to a research design used successfully in other evaluations where children above a clear cutoff receive an intervention (for example, children born before a certain date can enroll in prekindergarten) and are compared to similar children just below the cutoff who don't receive the intervention (this is called a *regression discontinuity* design).

Although these options weren't feasible, in the process of working closely together to develop our research plans, we discovered a hidden natural experiment of the sort frequently used in educational, economic, public health, and policy research. *Natural experiment* refers to an instance when two otherwise identical groups are affected differently by an event that is unrelated to either the treatment or the outcome and is outside of the researchers' control. Because a natural experiment mimics random assignment, it can answer casual questions while avoiding some of the practical and ethical challenges inherent in experimental research at scale.

The natural experiment we uncovered was the lucky result of a delay in decisions about funding the Explore PL track; the delay occurred due to factors outside the strict control of the DOE, NYU, and the programs themselves. When it came time to assign sites to PL tracks, the DOE didn't know whether funding for a new cohort of Explore sites would be available. So the department ran the site assignment process under two scenarios. Scenario A assumed funding for Explore was available, and sites could be assigned to Explore, Create, Thrive, or Inspire. Scenario B assumed no funding for Explore was available, and sites could be

Figure 1. Explore Natural Experiment as a Result of Funding Delay



assigned only to Create, Thrive, or Inspire (see figure 1). Because funding for Explore seemed unlikely to materialize, the DOE notified sites of their track assignment under scenario B; that is, under the assumption that the Explore track wouldn't serve a new cohort of sites. Shortly thereafter, though, the DOE secured funding for an additional Explore cohort. Sites that had been assigned to Inspire under scenario B but would have been assigned to Explore in scenario A were offered the opportunity to switch to Explore; sites that had been assigned to the Create and Thrive tracks under scenario B weren't given the choice to switch because the DOE had committed to serve a certain number of program leaders and teachers in those tracks.

As figure 1 shows, the Explore track ultimately comprised sites that both had and hadn't been assigned to Explore under scenario A. We selected a subset of Explore sites as our study sample. The treatment group in our natural experiment included sites that would have been assigned to Explore in scenario A and, in the end, were assigned to receive it after all. Our comparison group comprised sites that would have been assigned to Explore according to scenario A but were placed in a different track. This comparison group allowed us to examine what might have happened to sites in our Explore (treatment) group, on average, if they had been assigned to the other tracks. In our research, we use these inadvertent treatment and control groups to estimate the effects of Explore training and coaching on classroom quality and teachers' math instructional practices.

We were able to identify and take advantage of this natural experiment only because of the relationship we had established across our teams, which included open communication and frequent contact. Had the NYU team members been external researchers, we probably wouldn't have learned about the chain of events that led to the natural experiment; had the DOE not been willing to partner on this research and share details about the process—some of which may not have seemed relevant at first—we wouldn't have recognized the opportunity that the funding delay gave us to learn about Explore implementation and impacts. We expect that this is the case in other contexts as well. Natural experiments may not be as rare as

we might think, but they may go unnoticed when external researchers aren't in the know about such events and policymakers aren't on the lookout for such opportunities. RPPs are a way to avoid missed opportunities to rigorously estimate effects in the absence of a traditional randomized controlled trial.

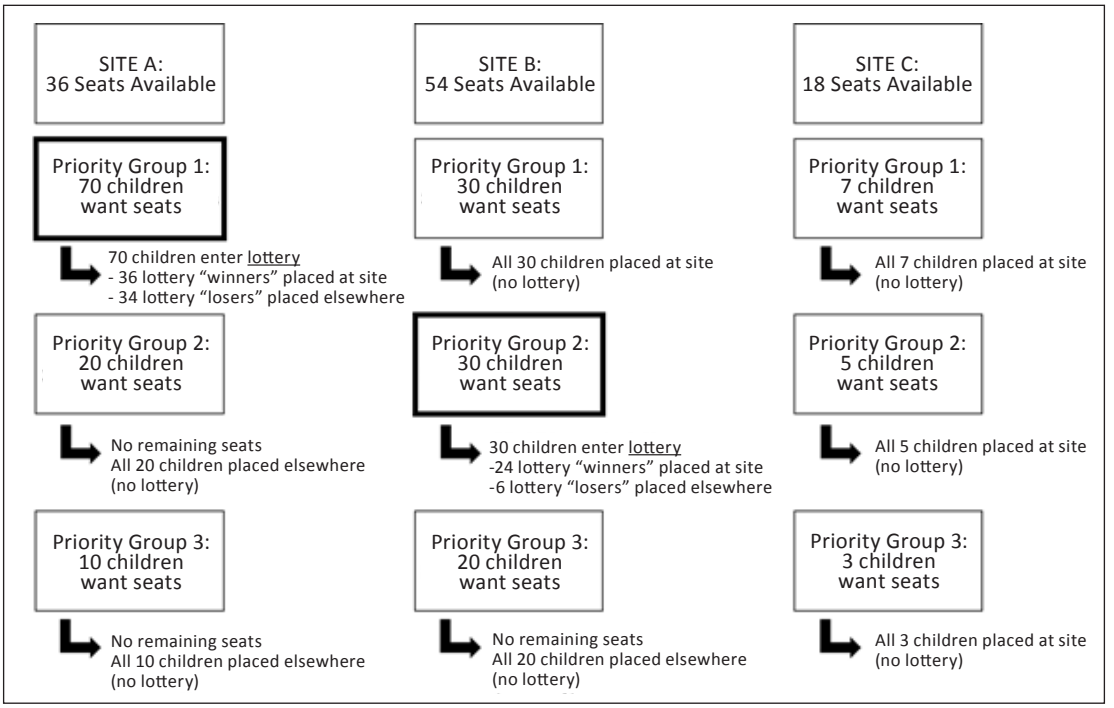
Before our main analyses testing Explore's effects on program quality and teacher practice, we conducted preliminary analyses to ensure that the natural experiment was a valid approach for evaluating Explore. First, we examined whether treatment and control sites were equivalent at the beginning of the school year, as would be expected in a randomized controlled trial. Documenting equivalence would give us confidence that any differences we observed at the end of the year were due to the treatment (that is, to Explore PL) and not to pre-existing differences between sites. This was important to investigate because factors other than the natural experiment (for example, site leaders' second choice among the PL tracks) contributed to the final track assignments, given the way sites were reassigned after the DOE secured Explore funding. We tested differences between our treatment and control groups across a range of child- and site-level characteristics and found a few significant and/or marginal differences between Explore and non-Explore sites; for example, control group sites had higher classroom quality scores on a widely used rating scale. In all cases, the differences we observed had the potential to bias estimates of treatment impact downward rather than upward. But they still raised the possibility that observed differences in teachers' practices in the spring could reflect pre-existing differences rather than participation in Explore.

Following our primary analyses testing Explore's effects, we also ran supplemental analyses to examine whether our findings were consistent across different subsamples and under different analytic assumptions. For example, we tested whether our findings were similar across the full sample, among district schools only, and among sites with the same PL preferences (for example, sites that listed Explore as their first choice and Create as their second choice) since this ultimately influenced their track assignment. Restricting the sample in these ways allowed us to estimate impacts among treatment and control sites we knew to be equivalent on key characteristics at the start of the year. This strengthened our confidence that any positive impacts we observed at the end of the year were due to participation in Explore and not pre-existing differences, especially when different analytic approaches produced results that were similar in pattern, magnitude, and significance.

Using Existing Randomization: Application Lotteries

We've just seen that the close partnership and regular communication between NYU and the DOE let us capitalize on a natural experiment to answer a policy-relevant question when the opportunity presented itself and to generate critical information about the Explore PL track. But new research questions emerged as the Pre-K for All system continued to expand and evolve over time. As an RPP with the mission of using data and research to strengthen the system broadly, we were well-positioned to expand our focus and answer new questions. For example, how do the other PL tracks affect children's outcomes? What are the effects among new cohorts of sites and children?

Figure 2. Lotteries in Three Hypothetical Sites



In the absence of natural experiments that could help us answer these new questions, we have considered other designs for understanding how PL affects teacher practice and child outcomes. One approach that's been used in New York at the high school level and increasingly in other settings at the prekindergarten level takes advantage of the fact that children are randomized to sites via lotteries that occur when a site receives more applications than seats; we call these sites *oversubscribed*.[14] In such cases, some children are randomly assigned to the oversubscribed site ("lottery winners") and other children are assigned to a site lower on their preference list ("lottery losers"). This lets researchers estimate both the effect of winning a lottery in and of itself (known as *intent-to-treat effect*) and the effect of winning a lottery and then enrolling in a given site (known as *local average treatment effect*).[15]

In New York, parents can apply to up to 12 Pre-K for All sites for their children, in order of preference. At the same time, the DOE places each child in a "priority group" for each site, which is based on criteria such as living in a certain zone or district, whether the student has a sibling at the site, and whether the child was a student at that site as a three-year-old (see figure 2). To understand whether we could use Pre-K for All application lotteries to answer our research questions about PL, we used a prior year of Pre-K for All application and enrollment data to examine issues of internal validity (the extent to which the child-level lottery design would allow us to establish the causal effect of PL) and generalizability (the extent to which our findings would apply to Pre-K for All children and sites outside our lottery sample). First, we identified the number and distribution of lotteries that occurred through the Pre-K for All

application process. Building on previous work that used lotteries to estimate the impacts of the city's Small High Schools of Choice, we developed an algorithm to identify sites where lotteries occurred, as well as the individual children who won or lost those lotteries.[16] To ensure that we accurately identified lotteries and lottery participants during this process, NYU researchers met several times with DOE partners to discuss the application process and nuances in the data. We then confirmed that the assignment process produced enough lotteries for each PL track to give us adequate sample size to test the PL tracks. We found that the lotteries were well distributed across the city and across different types of sites, indicating that lottery sites represented the full range of geographic and setting type diversity within the Pre-K for All system.

We also examined whether lotteries produced treatment and control groups (lottery winners and lottery losers, respectively) that were similar at the start of the school year. Because only limited demographic data were available on prekindergarten children and their families at the time of application, we used information about children's application choices themselves (for example, the number of choices listed, the proportion of choices in children's home districts, and the type and quality of children's top choices) as baseline characteristics, as well as demographic data that were available once children enrolled in Pre-K for All (but that weren't available for those who didn't enroll). Across these analyses, we found that the choice patterns and demographic characteristics of lottery winners and losers were generally well-balanced. This indicates that the lottery process did in fact create groups of treatment children and control children who were similar at the start of prekindergarten, a necessary condition for drawing causal inferences about the effects of PL.

To better understand the *treatment contrast* (that is, the difference between what treatment children experience and control children experience), we also examined the extent to which lottery winners and losers complied with their lottery assignments and the effect of winning a lottery on children's prekindergarten experiences. As expected, we found that winning a lottery for a site in a given track was indeed associated with greater exposure to that track during the prekindergarten year (for example, winning a lottery for an Explore site increased children's exposure to Explore). However, analyses also revealed that winning a lottery was associated with exposure to site characteristics other than PL (for example, higher program quality), making it difficult to isolate the effects of PL under this design. NYU, the DOE, and a team of methodological experts then considered multiple options for addressing the fact that lottery winners and lottery losers attended sites that differed in ways other than PL. For example, we considered using complementary quasi-experimental methods to test each track relative to a matched (but non-randomized) control group comprising only children in Inspire sites, and using secondary, descriptive analyses to examine how much of the treatment effect was attributable to PL versus other site characteristics. We also considered and ultimately switched to a different design altogether—a site-level randomized design—when we learned that sites would be assigned to PL track via the DOE's track assignment process. In our case, this design turned out to be better than the child-level lottery design because

it allowed us to experimentally isolate the effects of PL (our primary research question) and produce the clearest and most relevant evidence for NYC DOE.

Creating Randomization in Track Assignment

We've described two examples of how events or processes occurring in the education system can be used to conduct rigorous evaluations at scale: a natural experiment involving site-level assignment that occurred at a single point in time, and child-level randomization that occurs each year during the application process. In both examples, the existing system is taken as a given; that is, we capitalized on randomization that occurs as a by-product of the system and didn't manipulate the system for research purposes. In contrast, a third approach we've explored involves tweaking the system to create new opportunities for rigorous research.

Recall that Pre-K for All sites are assigned to PL tracks through a complex algorithm. Although the algorithm is revisited each year, assignment has always relied on program leaders' rank-ordered preferences among the PL tracks, site need, and each track's capacity. Our NYU and DOE teams have worked together to explore whether and how this PL assignment process might be adjusted to ensure more randomization to enable rigorous research in the future.

In our discussions, we considered the advantages of randomization for understanding the causal impacts of the different tracks and producing clear, actionable findings the city could use to strengthen and refine its PL model. The DOE is interested in carrying out more research on PL tracks, and it recognizes the value of maximizing the rigor of that research. But it has also underscored a rationale for prioritizing program leaders' preferences in track assignment—they see this approach as critical both for meeting the sites' individual needs and interests and for building buy-in among leaders and teachers. Indeed, the DOE's early data on teachers' attendance at PL sessions indicated high attendance among programs that received their first or second choice, with a notable drop-off in attendance for programs who received their third-choice track. Keeping in mind the two goals of honoring program leader preference and conducting high-quality research as our starting points, our partnership has searched for ways to embed randomization in the DOE's process for assigning sites to PL tracks in a way that still prioritizes choice.

Figure 3 illustrates one approach we came up with, simplified for illustration purposes (for example, we haven't included various preceding steps or exclusion criteria). Briefly, we jointly developed a process in which subsets of sites could be stratified by similar first and second preferences among tracks and then randomized. For example, sites that listed Explore and Inspire as first and second choices (in either order) would be grouped together and then randomized to either Explore or Inspire. Through this process, all the sites would receive either their first or second choice. At the same time, site-level randomization would make it possible to evaluate the causal effects of one track relative to another among sites with similar preferences (and other characteristics). Thus we can embed rigor and opportunities for research while still honoring program leaders' preferences.

As in the Explore evaluation we described above, we wouldn't have thought of this

Figure 3. Randomizing Hypothetical Sites Based on Preference Subgroups

	Track Rank-Ordered Preference			
Site	**Inspire**	**Explore**	**Thrive**	**Create**
Site A	1	2	3	4
Site B	2	1	3	4
Site C	2	1	4	3
Site D	1	2	4	3
Site E	1	3	4	2
Site F	1	4	3	2
Site G	2	4	3	1
Site H	2	3	4	1

Randomly assign sites that ranked Explore & Inspire #1 & #2 (in either order) → **Explore** / **Inspire**

Randomly assign sites that ranked Create & Inspire #1 & #2 (in either order) → **Create** / **Inspire**

approach outside the context of an RPP built on a history of joint efforts and a foundation of trust and mutual respect for each other's perspectives, expertise, and commitment. Finding this potential solution required flexibility on the part of both researchers and policy partners as we worked together toward the shared goal of embedding research that generates the strongest evidence for understanding and strengthening the PL system.

Challenges

We've shown how the NYU-DOE partnership navigated a complex system to find or create opportunities to conduct rigorous and relevant research. But these examples also reveal challenges. Any research project involves tradeoffs, and the designs we've used in our work have limitations. In the case of the Explore natural experiment, we found that despite randomness in track assignment due to funding delays, the treatment and control groups weren't fully equivalent on all site-level characteristics. In our broader evaluation of the Pre-K for All PL system, we've faced other challenges. For example, although the lotteries that are created during the application process ensure that children who win lotteries (treatment participants) and lose lotteries (control participants) are equivalent in expectation, the sites they attend are not. This means that the treatment contrast we can estimate using this experimental approach isn't as precise as we might like, given that the sites that lottery winners attend differ from the sites that lottery losers attend with regard not only to PL tracks but to other characteristics as well. These differences may contribute to differences in children's experiences and outcomes, which required us to consider complementary methods and alternative designs to isolate the effect of PL.

The methodological issues we've encountered underscore a few important points. For one thing, conducting rigorous research in fully scaled systems is messy. We've found that the advantages of capitalizing on rigor that already exists in such a system outweigh the disadvantages. But it's not always a perfect solution. It's critical to conduct an extensive set of analyses to examine the validity and limitations of the intended research designs, and then to revisit and revise the designs or analytic approach

on the basis of these findings. Using multiple analytic approaches to answer the same research question may also be a useful way to allay some of the methodological challenges of evaluating fully scaled programs.

Another problem is that the research process may be slower than the evolution of the system, program, or population we seek to study, a challenge other partnerships have also experienced (for example, see the article in this issue by Christina Weiland and Jason Sachs). The delay between planning a study and executing it can challenge both researchers and their practice and policy partners. Because local education programs or initiatives are constantly evolving, research designs and measurement protocols that may be perfectly appropriate at a given time may not be appropriate a year later because components of the system have changed. Indeed, these sorts of changes are expected and desired by leaders who work to improve the system's implementation. From the policy partners' perspective, multiyear research projects may cause district leaders to feel locked in to certain decisions or restricted in the programmatic changes they could make to strengthen the system.

We've encountered some of these challenges of studying an evolving system in New York City, where the DOE invests significant time and resources to improve the quality of Pre-K for All—including efforts to modify and refine the PL model that is a focus of our joint research. On one hand, this means that the system is continually improving and innovating, which ideally leads to better outcomes for children. On the other hand, our research may be too slow at times to contribute to these changes, and the right time to freeze and evaluate a constantly evolving system isn't always clear. The communication structures in our RPP and a foundation of trust built over the years have made it possible for the DOE to raise these kinds of concerns with NYU early and work toward mutually agreeable solutions. Flexibility has been key in this process: thanks to flexibility in our funding and flexibility in our partnership, we've adjusted aspects of our research plans while staying true to the key aim of evaluating the impacts of different PL models on teachers and children.

Another challenge of partnership work at scale is navigating a large district central office and a large university. To ensure shared understanding and alignment across partners, we've relied on frequent communication through multiple channels with the many people who have a stake in our work. We've also thought about how and when to loop in researchers and/or policy partners during meetings and planning processes versus when to move the work forward internally in our respective organizations. For example, both NYU and DOE need to be involved when developing research questions or discussing possible research approaches. But examining the validity or limitations of particular methodological approaches likely doesn't require as much DOE involvement. Likewise, rather than involve NYU researchers in every planning meeting before a new initiative is rolled out, in some cases the DOE brings in researchers only at key decision points, toward the end of a planning process, or when there may be implications for research.

Successes

Our partnership's work has a number of implications for practice and policy in New York City, for early childhood education more broadly, and for the way we use and conduct research within and outside our

partnership. Most importantly, our work supports the DOE's efforts to refine and strengthen Pre-K for All. Over the course of our partnership, NYU has helped the DOE by sharing emerging trends in research, lending expertise about methods and measurement tools, and telling the DOE about other districts undertaking similar efforts or facing similar problems. The DOE has identified priorities, shaped our research agenda, provided opportunities for rigorous research, and navigated potential barriers to our joint work. By thinking and planning together, particularly around PL and program quality standards, we developed the framework for monitoring and supporting quality in prekindergarten sites across the city. As findings emerge from our evaluations of the PL system, these too will guide further adaptations of the Pre-K for All system—including both incremental improvements such as strengthening onsite implementation support and larger changes such as discontinuing tracks that aren't shown to be effective.[18] By identifying components of the system that are working well, as well as areas that could be improved, our co-conducted research provides concrete information for decision-making and helps the DOE secure additional funding to support effective program models. Indeed, promising findings from the first year of the Explore natural experiment provided initial support for continuing to invest in that track. Understanding whether certain approaches are more effective for different types of sites or children will be especially important as the DOE seeks to refine differentiation of PL within the system, including the ways that PL might be sequenced over multiple years.

Our partnership's work has also strengthened the quality and relevance of our research, and therefore enhanced our contributions to education science broadly. New York is unique in some important ways, but findings from our work are generalizable to other large urban districts. We expect our work to advance the scholarship on PL in early child education by focusing on implementation and impacts in fully scaled real-world systems. As we've said, most of the rigorous evidence about whether PL models or curricula are effective comes from smaller, tightly controlled trials where program developers typically had a hand in program implementation and/or evaluation. By identifying programs that are effective at scale under real-world conditions across a diverse population, our work will produce evidence that's more relevant to the realities faced by cities, states, and districts and thus more likely to support practice and policy.

Our work has also increased the capacity of NYU and the DOE to sustain research-practice partnerships. By developing structures for regular communication and feedback, engaging partners at different levels of the organization, and establishing a successful track record, we've developed a culture and shared language around rigorous research, evidence, and data. That's allowed us to jointly identify new research opportunities and new ways to use data or prior research to guide decision-making.

Our experiences in the partnership have influenced our work outside the partnership as well. As policy makers, we've increasingly emphasized data and research in other areas of decision-making. As researchers, we've become more flexible and creative with the research tools at our disposal and learned to find opportunities for rigor even when we don't have strict control over randomization. In addition, in the examples we've described here, we examined the

extent to which the research designs were appropriate for answering our questions of interest. We also identified limitations in our designs and came up with ways to overcome them. These issues are particularly salient in our work evaluating Pre-K for All, where we as researchers have less control over the assignment processes underlying our research designs. But carefully testing the assumptions and validity of a given approach is no less important in other research. The same can be said about whether a sample is representative of the population being studied and whether results can be generalized to other contexts. Given that the DOE needs to understand how PL is working across the system, we've had to carefully consider how findings from the Explore natural experiment sample and the lottery subsample might or might not be applicable to other Pre-K for All sites and children. Paying more attention to these issues has carried over into our work outside the partnership.

Recommendations for Researchers and Policymakers

RPPs are uniquely positioned to capitalize on existing opportunities and/or negotiate small tweaks in existing systems to strengthen the rigor of research designs. In our experience, this is made possible by a sense of trust built over the years that characterizes effective and sustained partnerships. Open communication channels and flexibility on the part of both researchers and policy partners have also been critical. Below we offer three recommendations for researchers, policymakers, and practitioners who are considering partnership work to advance the science and practice of early childhood education.

Build a Foundation of Trust and Openness to Research.

The current phase of our research is possible only because of the joint work that preceded it and the relationships we cultivated through years of close collaboration. In our partnership, NYU has remained committed to supporting the DOE's efforts across a range of initiatives, and the DOE has remained open to and invested in high-quality research as a means to strengthen Pre-K for All. From its earliest stages, NYU and the DOE made this partnership a high priority, followed through with timely and high-quality work, and dedicated staff and funding to support the effort. We've found that shared understanding of—and responsiveness to—each partner's goals, concerns, and expertise are critical to the partnership's health. A foundation of trust allows partners to be open and candid, which helps in building a relevant and rigorous research agenda. Identifying mutually agreeable research aims and designs comes more easily when policymakers are transparent about their goals and priorities beyond research and when researchers are honest about the advantages, limitations, and potential burdens on the system of different designs.

We encourage anyone looking to establish or sustain an RPP to engage members at multiple levels of each institution, communicate regularly, and find opportunities to meet in person. Engaging a range of people with varied roles and expertise helps to build a culture of using research to improve practice, helps sustain the partnership as people transition in and out, and brings new opportunities to light that might be missed if collaboration were confined to a select few. Establishing project-specific meetings has helped us focus on a particular task or

issue, while standing check-ins allow partners to share general updates, raise issues or questions, or tee up new ideas. This structure ensures that we make steady progress on planned work while also letting new research opportunities come to light as we keep each other in the loop about other efforts outside the partnership. These factors helped us uncover and understand the Explore natural experiment, for example, and identify ways to embed randomization in the process for assigning prekindergarten sites to tracks. Regular check-ins, with many members of our partnership at the table, have let us identify and capitalize on new opportunities for research or research-based decision-making as they emerge, or shape our research agenda in ways that better align with new priorities or upcoming initiatives.

Consider the Right Time for Causal versus Descriptive Research

Although rigorous evaluation is often critical to guide practice and advance ECE scholarship, descriptive research can also be useful for RPPs. As we said earlier, our partnership's initial research efforts were primarily descriptive, and we shifted to causal methods only when our jointly developed research agenda changed its focus to the PL system and its impacts. Our perspective is that trustworthy, well-sampled descriptive data are invaluable to senior leaders as a foundation to evaluate whether new policy initiatives are on the right track. Moreover, we expect that senior leaders benefit from gaining fluency with quantitative data about children's school readiness and program quality as a necessary first step toward more complex questions that might involve more sophisticated analytic approaches.

A few factors contributed to our shift from descriptive to causal research. First, the city's priorities changed from monitoring the system during a period of rapid expansion to strengthening quality within it. Second, after a few years of Pre-K for All, the DOE and NYU agreed that certain components of the system were ready for evaluation. We first focused on impacts on program quality and then expanded our focus to child outcomes once the system stabilized to some degree after a period of rapid expansion and development. Third, by this point, our partnership was strong enough—thanks to trust built during the first phase—to engage in evaluations with higher stakes than the descriptive work we first took on. Other factors might come into play in other settings, such as program scale, resource or other constraints, the speed with which results are needed to make decisions, the structure of the system or program under investigation, or the nature of the process being studied. It should be clear that we don't recommend maximizing rigor at all costs. Researchers and policy makers should carefully consider when and where causal versus quasi-experimental versus descriptive research is most appropriate given the research questions and context.

Be Flexible

With the examples in this article, our intention is to encourage research and policy partners to think creatively together about:

1. whether rigor that can be used for research purposes may already exist in a system and;

2. whether there are ways to add rigor while staying true to other priorities and operating within existing constraints.

Indeed, certain programmatic constraints such as limited funding or capacity present unique opportunities for rigor, and these can be used to the advantage of research and policy partners rather than being seen as barriers. In this article, we discussed three approaches to rigorous research that fit our needs and constraints in New York City—capitalizing on a natural experiment, using existing randomization of children to oversubscribed prekindergarten programs, and creating new opportunities for rigorous research by incorporating randomization into the way sites are assigned to PL tracks. We encourage research and policy partners to consider these types of designs, as well as other approaches. For example, waitlist designs (when an intervention is withheld from a control group for a certain period of time but is offered to that group once the research is complete) may be suitable when a new program or initiative can't be rolled out instantly; randomization may be possible when an intervention has limited capacity to serve prekindergarten teachers or students; regression discontinuity designs may work well when clear cutoffs are used to make decisions about how to allocate resources or assign children or sites to interventions. And as we've said, descriptive work may be preferable depending on the focus of the research question, the stage of the program being studied, or the stage and strength of the partnership itself. Finding the optimal design requires researchers to be willing to deviate when necessary from randomized controlled trials and explore other, more feasible approaches that still maintain high standards of rigor, including quasi-experimental designs when randomization isn't possible in any form.

It is also worth noting that our partnership was established to support the monitoring and strengthening of the Pre-K for All system broadly, rather than to answer a specific question or evaluate a specific component of the system. That has given us the flexibility to revisit and revise research aims and successfully adapt to evolving priorities and emerging questions, setting our approach apart from more traditional university-driven research. For example, we shifted from descriptive work to evaluation of the PL system when the city needed answers about how to improve quality, and we adapted our early plans to evaluate impacts on child outcomes to evaluate impacts on program quality after recognizing that program-level results would be more helpful to the DOE in the first few years of the PL system's implementation. The guiding principle for our multiyear collaboration thus far has been to put science to work, and our process has supported our broad goal of a sustained partnership that facilitates bidirectional influence of rigorous research and policy/practice in ways that are useful to the city first and to education science generally second.

Our work shows that it's possible through partnerships to build studies that can provide strong causal evidence that's directly relevant to policymakers and meets the standards of rigor that are necessary for scholarship, even when constraints or competing priorities in large systems make it impractical or impossible to use randomized controlled trials. With flexibility and persistence, researchers and policy partners are well positioned to produce the kind of evidence that advances understanding of key issues in ECE, directly supports programming and decision-making on the ground, and maximizes benefits for prekindergarten leaders, teachers, and children.

Endnotes

1. John Q. Easton, "IES: Promises and Challenges" (lecture, Annual Meeting of the American Educational Research Association, Philadelphia, PA, April 5, 2014); C. Cybele Raver, "Targeting Self-Regulation through Intervention: Lessons from RCTs" (lecture, Society of Research on Educational Effectiveness Fall Conference, Washington, DC, September 26, 2013).

2. Vivian Tseng, John Q. Easton, and Lauren H. Supplee, "Research-Practice Partnerships: Building Two-Way Streets of Engagement," *Social Policy Report* 30, no. 4 (2017): 3–15, https://doi.org/10.1002/j.2379-3988.2017.tb00089.x.

3. Gregory Camilli et al., "Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development," *Teachers College Record* 112 (2010): 579–620; Deborah A. Phillips et al., *The Current State of Scientific Knowledge on Pre-Kindergarten Effects* (Washington, DC: Brookings Institution, 2017); Hirozaku Yoshikawa et al., *Investing in Our Future: The Evidence Base on Preschool Education* (Washington, DC: Society for Research in Child Development, 2013).

4. Eric A Hanushek, "The Economic Value of Higher Teacher Quality," *Economics of Education Review* 30 (2011): 466–79, https://doi.org/10.1016/j.econedurev.2010.12.006; Andrew J. Mashburn et al., "Measures of Classroom Quality in Prekindergarten and Children's Development of Academic, Language, and Social Skills," *Child Development* 79 (2008): 732–49, https://doi.org/10.1111/j.1467-8624.2008.01154.x; Christina Weiland et al., "Launching Preschool 2.0: A Road Map to High-Quality Public Programs at Scale," *Behavioral Science & Policy* 2, no. 1 (2016): 37–46; Steven W. Barnett, "Effectiveness of Early Educational Intervention," Science 333 (2011): 975–8, https://doi.org/10.1126/science.1204534; Margaret Burchinal et al., "Predicting Child Outcomes at the End of Kindergarten from the Quality of Pre-Kindergarten Teacher–Child Interactions and Instruction," *Applied Development Science* 12 (2008): 140–53, https://doi.org/10.1080/10888690802199418; Tran D. Keys et al., "Preschool Center Quality and School Readiness: Quality Effects and Variation by Demographic and Child Characteristics," Child Development 84 (2013): 1171–90, https://doi.org/10.1111/cdev.12048; Yoshikawa et al., *Evidence Base*.

5. Karen L. Bierman et al., "Executive Functions and School Readiness Intervention: Impact, Moderation, and Mediation in the Head Start REDI Program," *Development and Psychopathology* 20 (2008): 821–43, https://doi.org/10.1017/S0954579408000394; Douglas H. Clements et al., "Mathematics Learned by Young Children in an Intervention Based on Learning Trajectories: A Large-Scale Cluster Randomized Trial," *Journal for Research in Mathematics Education* 42 (2011): 127–66, https://doi.org/10.5951/jresematheduc.42.2.0127; Alice Klein et al., "Effects of a Pre-Kindergarten Matheматics Intervention: A Randomized Experiment," *Journal of Research on Educational Effectiveness* 1 (2008): 155–78, https://doi.org/10.1080/19345740802114533; Christopher J. Lonigan et al., "Promoting the Development of Preschool Children's Emergent Literacy Skills: A Randomized Evaluation of a Literacy-Focused Curriculum and Two Professional Development Models," *Reading and Writing* 24 (2011): 305–37, https://doi.org/10.1007/s11145-009-9214-6; Pamela Morris et al., *Impact Findings from the Head Start CARES Demonstration: National Evaluation of Three Approaches to Improving Preschoolers' Social and Emotional Competence* (New York: MDRC, 2013); C. Cybele Raver et al., "Improving Preschool Classroom Processes: Preliminary Findings from a Randomized Trial Implemented in Head Start Settings," *Early Childhood Research Quarterly* 23 (2008): 10–26, https://doi.org/10.1016/j.ecresq.2007.09.001; C. Cybele Raver et al., "Targeting Children's Behavior Problems in Preschool Classrooms: A Cluster-Randomized Controlled Trial," *Journal of Consulting and Clinical Psychology* 77 (2009): 302–16, https://doi.org/10.1037/a0015302; Barbara A. Wasik and Mary Alice Bond, "Beyond the Pages of a Book: Interactive Book Reading and Language Development in Preschool Classrooms," *Journal of Educational Psychology* 93 (2001): 243–50, https://doi.org/10.1037%2F0022-0663.93.2.243; Carolyn M. Webster-Stratton, Jamila Reid, and Mike Stoolmiller, "Preventing Conduct Problems and Improving School Readiness: Evaluation of the Incredible Years Teacher and Child Training Programs in High-Risk Schools," *Journal of Child Psychology and Psychiatry* 49 (2008): 471–88, https://doi.org/10.1111/j.1469-7610.2007.01861.x.

6. Greg J. Duncan and Katherine Magnuson, "Investing in Preschool Programs," *Journal of Economic Perspectives* 27 (2013): 109–32, https://doi.org/10.1257/jep.27.2.109; Robert C. Pianta et al., "The Effects of Preschool Education: What We Know, How Public Policy Is or Is Not Aligned with the Evidence Base, and What We Need to Know," *Psychological Science in the Public Interest* 10 (2009): 49–88, https://doi.org/10.1177/1529100610381908.

7. Bridget K. Hamre, Ann Partee, and Christina Mulcahy, "Enhancing the Impact of Professional Development in the Context of Preschool Expansion," *AERA Open* 3, no. 4 (2017): https://doi.org/10.1177/2332858417733686.

8. Raj Chetty et al., "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR," *Quarterly Journal of Economics* 126 (2011): 1593–1660 https://doi.org/10.1093/qje/qjr041; Kenneth A. Dodge, "Community Intervention and Public Policy in the Prevention of Antisocial Behavior," *Journal of Child Psychology and Psychiatry* 50 (2009): 194–200, https://doi.org/10.1111/j.1469-7610.2008.01985.x; Alan B. Krueger, "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics* 114 (1999): 497–532, https://doi.org/10.1162/003355399556052.

9. Annalisa Mastri et al., *Evaluating National Ongoing Programs: Implementing the WIA Adult and Dislocated Worker Programs Gold Standard Evaluation* (Washington, DC: Mathematica Policy Research, 2015).

10. Hamre, Partee, and Mulcahy, "Enhancing the Impact."

11. Tseng, Easton, and Supplee, "Research-Practice Partnerships."

12. Douglas H. Clements and Julie Sarama, "Experimental Evaluation of the Effects of a Research-Based Preschool Mathematics Curriculum," *American Educational Research Journal* 45 (2008): 443–94, https://doi.org/10.3102/0002831207312908; Julie Sarama et al., "Scaling Up the Implementation of a Pre-Kindergarten Mathematics Curriculum: Teaching for Understanding with Trajectories and Technologies," *Journal of Research on Educational Effectiveness* 1 (2008): 89–119, https://doi.org/10.1080/19345740801941332.

13. Laurie Miller Brotman et al., "Promoting Effective Parenting Practices and Preventing Child Behavior Problems in School among Ethnically Diverse Families from Underserved, Urban Communities," *Child Development* 82 (2011): 258–76, https://doi.org/10.1111/j.1467-8624.2010.01554.x.

14. Howard S. Bloom and Rebecca Unterman, *Sustained Progress: New Findings about the Effectiveness and Operation of Small Public High Schools of Choice in New York City* (New York: MDRC, 2013); Howard S. Bloom and Rebecca Unterman, "Can Small High Schools of Choice Improve Educational Prospects for Disadvantaged Students?," *Journal of Policy Analysis & Management* 33 (2014): 290–319, https://doi.org/10.1002/pam.21748.

15. Bloom and Unterman, *Sustained Progress*.

16. Bloom and Unterman, *Sustained Progress*.

17. Karl E. Weick and Robert E. Quinn, "Organizational Change and Development," *Annual Review of Psychology* 50 (1999): 361–86, https://doi.org/10.1146/annurev.psych.50.1.361.

18. Weick and Quinn, "Organizational Change and Development."

Building Capacity for Research and Practice: A Partnership Approach

Maia C. Connors, Debra M. Pacchiano, Amanda G. Stein, and Mallary I. Swartz

Summary

Building a research-practice partnership that spans two or more organizations can be challenging. University researchers, for instance, may find themselves in schools or departments where the incentives and culture favor more individualistic, single-discipline research projects. Practitioners and policy makers may not see the value of investing in research and evaluation, or they may hesitate to prioritize this work above more pressing operational and strategic aims.

In this article, Maia Connors, Debra M. Pacchiano, Amanda G. Stein, and Mallary I. Swartz show how an "embedded partnership" between the program implementation and research teams within a single nonprofit organization (Start Early, formerly the Ounce of Prevention Fund) can help overcome such challenges and build important capacities that research-practice partnerships need in order to be effective: an organizational culture that values research evidence, sound measurement, and continuous learning; interdisciplinary expertise; and sustainable infrastructure, including administrative support, technology, and financial resources.

At Start Early, Maia C. Connors is director of research and policy initiatives, Debra M. Pacchiano is vice president for translational research, Amanda G. Stein is director of research and evaluation. Mallary I. Swartz is a vice president at Start Early and a co-principal investigator for the National Center on Parent, Family and Community Engagement.

The authors thank Rebecca Berlin, Brenda Eiland-Williford, Rebecca Klein, Eleni Manos, and Abigail Sylvester for their input on previous versions of this article. They are also immensely grateful to all of their practice partners, without whom none of this work would be possible.

Increasingly, people who work in early childhood care and education (ECCE) are emphasizing that research evidence and data should guide the development, implementation, and improvement of programs, as well as decision-making at multiple levels. Yet the capacities needed to effectively achieve this ideal are substantial, and they're rarely achieved by researchers or practitioners alone.[1] In recent years, research-practice partnerships (RPPs) have garnered support as a promising solution to improving ECCE programs, outcomes, services, policies, and systems.[2] In a seminal report, Cynthia Coburn, William Penuel, and Kimberly Geil define five key aspects of successful RPPs: mutualism, commitment to long-term collaboration, a focus on problems of practice, the use of intentional partnership strategies, and trusting relationships.[3] In this article, we argue that RPPs can't reliably develop these qualities without intentionally building and fully integrating partnership capacities—including organizational structure and culture, interdisciplinary expertise, and sustainable infrastructure—among both researchers and practitioners.

Yet it's a substantial challenge to build and integrate these capacities for both partners following a traditional RPP model, in which researchers and practitioners are in separate organizations, each with its own discipline-specific capacities and focus. An alternative approach is to embed researchers and practitioners in a single organization. Such an embedded RPP may help partners more successfully develop organizational capacities that support strong partnerships and make use of them in sustainable ways. Here we discuss the unique advantages that embedded RPPs confer as well as the challenges they pose. We also offer recommendations for how to build capacity to support any successful RPP, including how to establish and foster research capacity in practice organizations to create embedded RPPs, as well as how nonembedded, cross-organizational RPPs can cultivate similar capacities when an embedded structure isn't desirable or feasible.

The lessons we share here derive from our experience conducting research and evaluation as part of embedded RPPs at Start Early (formerly the Ounce of Prevention Fund), a large early childhood nonprofit organization. Throughout the article, we draw examples from two RPPs: one focused on improving and scaling the Essential Fellowship (formerly Lead Learn Excel), a professional development program for ECCE leaders, and a second focused on implementing and improving outcomes of Educare Chicago, Start Early's local Early Head Start/Head Start school for children from birth through age five and their families.

Moving toward an Embedded RPP

Recently, practitioners, policy makers, and researchers have been calling for changes in the way intervention research is conducted. RPPs may be, in part, an answer to this call. Conducting research as part of an RPP is fundamentally different from conducting research with practitioners. In an RPP, instead of first defining a research question and then recruiting practitioners who are willing to participate in that research, researchers enter a partnership with practitioners to define the research questions that they'll pursue together.[4] This approach requires researchers and practitioners to seek a balance among their different needs and perspectives. For example, the need for causal evidence at key

points must be balanced against the need for a wealth of descriptive information about implementation to guide development and improvement of the program or practice (see the article in in this issue by Rachel Abenavoli and colleagues). Yet many research and practice organizations lack the capacity to effectively engage in this kind of partnership. For RPPs to succeed as a new way to conduct intervention research, both research and practice organizations must strengthen their capacity to do this work. In our experience, an embedded RPP structure that fully integrates research within practice and practice within research, so that the capacities of the partners are inextricably entwined, shows great promise as a sustainable and effective approach.

Embedding Researchers in Practice Organizations

The very existence of our roles as researchers in a practice-oriented nonprofit organization is a testament to the capacity our organization has built to support the work of RPPs. As an interdisciplinary organization, Start Early has always valued research and inquiry, but it has taken many years to realize the structures and capability needed to support these values internally. Our embedded RPP approach first took shape decades ago when the organization needed to report to various public and private funders on the implementation of our ECCE programs and grant activities. As funder groups began asking for more rigorous evidence of outcomes and quality, our role as Start Early researchers expanded to include selecting assessments, advising on measurement concepts (for example, reliability, validity, and standard scores), synthesizing program data, translating research findings, and generating questions of interest for grant applications. As a result, practitioners now had more information about their programs, and they started seeking assistance not only with collecting data but also with interpreting and using it to understand how to improve practice and outcomes and to translate those discoveries into policy. In response, Start Early researchers and program staff instituted new RPP routines (such as monthly meetings) to advance these shared aims.

Working within embedded RPPs in this way has helped us integrate rigorous research and evaluation studies into continuous quality improvement processes. As a result, practice as usual can support our research activities, and our ongoing research activities in turn bolster implementation, improvement, and decision-making among practitioners. This means, for example, that building sustainable data collection capacity and developing data use routines have become integral parts of both the RPP and the program model, rather than add-ons or tasks that take away from the "real" work of implementing the program or conducting research studies. This structure also helps us build long-term, trusting relationships that allow us, as research partners, to stick with our practice partners throughout the life of the program, reminding them of important data sources or relevant research findings as the program evolves, translating and reinterpreting findings as practitioners continue to draw from them in new ways, and tailoring future research to questions that arise when implementation challenges come up. We believe that as embedded researchers who don't work, for example, as tenure-track faculty in a university and thus aren't subject to the pressure to produce peer-reviewed publications, or at a research firm in which we'd be required to bill our time to contracted projects, we're particularly

well positioned to invest time and effort in this work.

Roles in an Embedded RPP

The hallmark of any RPP, and of our embedded RPPs in particular, is the joint management of the partnership by the research and practice partners, who make equal yet role-specific contributions to the work. In the sections that follow, we describe how partner roles are defined and differentiated in our embedded RPPs and also consider the important contributions that external researchers can make when added to embedded partnerships—an approach we have frequently taken.

The Research Partner

In our RPPs, the research partner has four functions:

- to design and build the infrastructure to support regular collection of reliable data on implementation and outcomes as an integral part of program implementation;
- to conduct evaluations to guide program implementation, improvements, and expansion;
- to conduct applied research to generate novel and generalizable findings; and
- to help program implementers interpret findings to guide their decisions.

Traditional RPPs primarily fill the second and third functions. But it's the first and fourth functions that create opportunities for research to be most useful to practitioners. For example, researchers can support strong and sustainable data collection infrastructure by leading survey development, helping to integrate survey data into existing program databases and reporting, and creating tools for program staff that remind them of the data collection timeline, measures, and purpose. Similarly, researchers can develop protocols to support practice partners' interpretation and use of the data. These functions require that research partners work closely with practice partners from the very beginning of a new program and continue doing so beyond the conclusion of any one study. To a degree that is relatively unusual in our field, this approach to sharing the journey from the start makes it easier to collect data, use that data to improve programs, and produce novel, actionable evidence.

The Practice Partner

In addition to developing and implementing an intervention or program, practice partners focus on identifying goals for data collection and research and on using findings to improve their practice. Specifically, practice partners

- specify key inputs, activities, and expected short- and long-term program outcomes and help determine what information needs to be collected about each;
- independently monitor data regarding basic program implementation (for example, number of participants or attendance);
- collaborate with research partners frequently to review, interpret, and reflect on data and use that information to identify small changes that can be made to increase fidelity

of implementation and effectiveness of model core components; and

- collaborate with research partners to use research findings to determine and design content additions or refinements to the program model itself.

For example, our practice partners often take the lead on documenting the supports they provide to families or ECCE professionals. Insights from their qualitative data are often especially useful in providing context for interpreting program evaluation findings in ways that can help programs improve.

External Research Partners in Embedded RPPs

Even when embedded research and practice partners both successfully fill their roles, expanding an embedded RPP to include external research partners can help overcome some key obstacles. For example, although embedded researchers may be best positioned to foresee and plan for the key points at which conducting a summative impact study is necessary and appropriate to assess the program's effectiveness in achieving its intended outcomes, they may lack the expertise or capacity to conduct a rigorous evaluation alone. In these cases, partnerships with external evaluators can be invaluable.

Specifically, because practice organizations typically have small research teams, embedded researchers can't possibly have expertise in all the topics and methods of inquiry that a long-term RPP may eventually require. Partnering with other researchers can bring distinct perspectives and specialized, complementary expertise to an embedded RPP just when it's needed most. External research partners are also valuable simply for what they represent or have access to. Researchers working in practice-oriented organizations may, for example, face challenges in meeting research funding requirements, building institutional review boards, developing information technology infrastructure, gaining academic library access, and obtaining staffing support from student research assistants or data analysts. Typically, these critical research supports are most well developed in universities and large research firms. Partnering with researchers who work in more traditional research institutions can help embedded RPPs access such resources. And embedded RPPs would also benefit from future work focused on understanding how to better provide these resources to researchers operating outside of traditional research institutions.

Working with outside researchers has one more advantage. Because embedded researchers are so deeply integrated with the practice partners who develop and implement the program they're studying, funders, for example, may express concerns about real or perceived biases. Outside researcher partners may alleviate these concerns by serving as neutral third-party evaluators. Indeed, each of the embedded RPPs we discuss in this article either grew out of or was supported by external evaluators who were instrumental in helping to build their initial capacity.

Our Embedded RPPs

Two of our embedded RPPs illustrate the capacities and partnership processes that help researchers work side by side with practitioners at Start Early. The primary

aim of both RPPs is to undertake formative evaluation of program implementation and short-term outcomes to guide improvement as the program develops and evolves. Through this work, we hope not only to provide our current partners with information they need to improve their practice in the short term but also to generate research findings in the longer term that are both relevant and generalizable to practitioners elsewhere.

The Essential Fellowship RPP

The Essential Fellowship is a professional development program designed to help ECCE leaders improve the quality of classroom instruction through instructional leadership and job-embedded professional learning routines for teachers in both schools and community-based ECCE programs. The program offers training, coaching, and peer learning, as well as practical tools and resources for local professional development providers and site-level ECCE leaders. The Essential Fellowship RPP has been at the core of this initiative since its inception: our research and evaluation team codesigned and co-constructed a robust data collection and research infrastructure that is embedded in the program model and that enables implementation, improvements, sustainability, and expansion.

In phase 1 of the project (2011–14), a team of researchers and practitioners at Start Early, supported by a federal Investing in Innovation development grant, developed and implemented the model, which was piloted with 15 administrators and 60 teachers serving a diverse group of approximately 600 infants, toddlers, and preschoolers in four community-based ECCE programs in Chicago serving low-income families. It also funded external researchers from the University of Illinois at Chicago (UIC) to study the program's effects. In phase 2 (2014–17), with Race to the Top-Early Learning Challenge funds from the state of Illinois, Start Early implementation and research teams refined key professional learning supports, tools, and strategies and implemented the Essential Fellowship statewide to help schools and community-based centers improve preschool instruction; UIC researchers also conducted external evaluations of the refined model. In this phase, the evaluators focused on implementation and outcomes (for example, changes in participants' mindset, knowledge, and practices over the course of the Essential Fellowship) rather than impact (causal estimates of differences in practice change between Essential Fellowship participants and a control group who didn't participate in the program).

The UIC partners were initially brought in because of funder-defined evaluation requirements. But as the first two phases of the work progressed, the long-term, cross-institution RPP that emerged among the Essential Fellowship program implementation and research and evaluation teams at Start Early and UIC became invaluable. In this partnership, Start Early researchers at the master's and doctoral level functioned as liaisons between the program implementers and UIC evaluators by promoting mutual understanding and improvement. The processes, methods, and measures used by UIC served as an important model for the Start Early research team because once the formal relationship with the external evaluation partners ended, we continued to formatively evaluate the program internally via our embedded RPP. Currently, the Essential Fellowship RPP

is working to answer program evaluation questions such as "How frequently do ECCE leaders participating in the Essential Fellowship use data when coaching their teachers?" as well as broader research questions such as "How do community-based versus school-based ECCE leaders use data in ways that are similar or different?"

The Educare Chicago RPP

Educare Chicago, directly operated by Start Early, is a high-quality ECCE center on the South Side of Chicago serving over 140 low-income or otherwise disadvantaged children from six weeks to five years old, as well as their families. Educare Chicago is the flagship school in a network of over 20 similar schools across the nation. Each full-day, full-year Educare school is a public-private partnership between a Head Start or Early Head Start provider, a local school district, and one or more philanthropic organizations. The schools subscribe to the Educare model, which includes four core practice domains: data use, embedded professional development, high-quality teaching practices, and intensive family engagement. To guide continuous practice improvement, Educare school leaders form RPPs with local evaluation partners (LEPs).

Start Early's research and evaluation division serves as an embedded LEP for Educare Chicago to support the school's staff and leaders. The division gathers direct assessment, survey, and observational data each year, providing considerably more information than Head Start centers typically collect. These data are then integrated with data gathered by program staff and used as part of existing, embedded routines (for example, management and grade-level meetings, lesson planning, and coaching) that focus on setting goals, planning for implementation, evaluating progress, and improving practice, staff professional development, and student and family outcomes. Together, Educare Chicago leaders and the LEP team identify priority areas of inquiry and generate questions that can be studied to help improve program quality.

The Educare Chicago RPP dates back to 2005, when there were only three Educare schools. Together, the schools developed a research design and approach to gathering common data that would help them better understand children's and families' progress in the program and illuminate the links between child and family outcomes and implementation of the Educare model. Since then, though both the design and implementation of the Educare Chicago RPP have evolved, the group has held steadfast in its commitment to using priority problems of practice identified by program leaders and staff to drive data analysis and to focus on the multilayered routines of continuous quality improvement.

Within the network, Educare Chicago has always been unique in that the school staff and LEPs are employed by the same organization; at most Educare schools, LEPs are university based. We believe that researchers who are embedded rather than based in a university have greater flexibility to spend time and effort building a sustainable infrastructure for research. For example, because the Educare Chicago LEPs and school staff work for the same organization, we don't have to negotiate a contract (and be limited to its terms) in order to work together. LEPs access the school's data directly by logging in to their various

data systems (such as ChildPlus or Teaching Strategies GOLD), and are acutely aware of practice requirements and constraints because we can see the terms of the school's funding agreements and contracts.

Building Capacity to Support the Work of RPPs

We've found that three types of capacity are essential to organizing and operating our embedded RPPs effectively and sustainably:

- an organizational structure and culture that values research evidence and sound measurement and that encourages professionals to engage in interdisciplinary collaboration and continuous learning;
- an adequate number of interdisciplinary staff who possess key knowledge, skills, and dispositions, and sufficient time to engage in these partnerships; and
- an infrastructure that enables this work to be done successfully and sustainably.

These three capacities may have made possible the lasting, dynamic relationship between our research and practice partners.[5] Cultivating these capacities to support both research and practice partners is necessary for any RPP to become the standard "way of working," whether the RPP is embedded or not. But we've found that building these capacities to support both partners in a single organization has made traditional boundaries between research and practice professionals and activities more permeable than they are in RPPs that are organized across two or more organizations.[6]

Organizational Structure and Culture

RPPs may thrive most readily in organizations that cultivate shared values and commitments and that facilitate processes and methods conducive to translating new knowledge into new practices and behaviors.[7] In particular, we expect that RPPs are likely to be most successful when partners work in one or more organizations that embody certain *learning-oriented* characteristics.

For example, strong organizational leadership that is committed to teamwork and continuous learning has been essential to our RPPs. Such leadership crafts the organizational structure and allocates staff time in ways that prioritize collaboration among diverse professionals. Specifically, we believe research and practice partners work together most efficiently when leaders structure their organizations to be inherently interdisciplinary by making it easier to work across teams and with external partners. At Start Early, teams of professionals with different kinds of expertise—including in policy, research, and practice related to both home visiting and center-based ECCE and in the development of professionals and systems—routinely work together on projects that cut across these categories. Executive leaders at Start Early have supported this way of working by making investments in strengthening both research and practice teams' ability to collaborate with one another. For example, the research team has grown to include more doctoral- and master's-level researchers whose jobs focus on working directly with Start Early program teams, and program teams now include staff who concentrate on strengthening program implementation

and operations through collecting, managing, using, and reporting data with support from their research partners.

We also expect that RPPs will be more successful when the organizations that employ both practice and research partners strongly value answering critical research questions or solving problems of practice or policy and are intentionally structured to do so. This focus is more likely to yield novel findings that are relevant to a broader audience and to help partners work together to translate and apply this new knowledge to ECCE policy and practice. Although a focus on inquiry is likely common in the organizations where research partners work, it may not be as common in practice partner organizations, and it takes time to develop.

RPP's may thrive most readily in organizations that cultivate shared values and commitments and that facilitate processes and methods conducive to translating new knowledge into new practices and behaviors.

Finally, both research and practice partners need to work in organizations that value their shared pragmatic goals: to support effective design and implementation of high-quality programming, practices, and policies; continuous quality improvement; and strong child and family outcomes. These organizational characteristics are likely to be central to the organizations in which practice partners work, but the same may not be true of most research organizations. In fact, one of the primary advantages of embedding research partners in the same organization as the practice partners is that it allows researchers to center so much of their attention on these pragmatic goals. For example, Start Early's research and evaluation division is explicitly charged with partnering with practice and policy colleagues to identify and study critical unanswered questions and problems of practice and then to translate findings about those questions into actionable interventions, policies, or improvements. In contrast, most research organizations tend to have narrower objectives—concentrating primarily on original research, program evaluation, and testing interventions, and focusing less on dissemination to broad audiences, translation of evidence, and program improvement.

Interdisciplinary Human Capital

Whether researchers and practitioners are embedded in a single organization or work across multiple organizations, a successful RPP functions as an interdisciplinary group that makes better decisions together because of the members' diverse knowledge, expertise, and perspectives. At Start Early, staff members are hired both for their discipline-specific competencies and their interdisciplinary experience. This approach has helped not only to build the organization's human capital in multiple disciplines (including child development, social work, public policy, psychology, ECCE, K–12 education, knowledge sharing and technology, program evaluation, and applied research) but also to support collaborative work across these areas of expertise.

RPPs must also cultivate social capital; that is, they must establish trusting relationships among partners, engage in active listening, and demonstrate respect for others' discipline-specific standards and priorities. Partners' different priorities (for example, the research partner's need for high response rates to implementation surveys versus the practitioners' concerns about burdening their staff with documentation requests) inevitably lead to tensions. Placing both partners in a single organization doesn't eliminate this problem. But when researchers and practitioners are equally committed to the organizational mission and to each other as colleagues, the resulting climate can help bolster motivation, flexibility, and even creativity in navigating these inevitable tensions.

Finally, organizations must also think about human capital in terms of the composition, depth, and breadth of capabilities across their entire workforce. At a basic level, having adequate staff with dedicated time is critical to RPPs' success.

Sustainable Infrastructure

A fully integrated RPP requires solid infrastructure that bridges disciplinary gaps, thereby encouraging sustainable researcher-practitioner collaboration, program implementation and improvement, and research activity.[8] Four primary components of our RPPs' infrastructure have been key to accomplishing these goals.

First, our embedded RPPs rely on systems for planning, scheduling, and agenda setting that help coordinate and advance the work across the entire cycle of program development, piloting, refinement, and evaluation. With these systems in place, researchers and practitioners come together routinely to collaborate on activities such as determining research questions; developing or selecting appropriate measures; collecting, managing, and analyzing data, and preparing data for interpretation; and using data for reporting, knowledge building, and program improvement. Without this kind of infrastructure, these activities either don't happen or aren't accomplished collaboratively in ways that support relationship building, reflection, and planning. A typical RPP at Start Early meets formally once a month, but informal discussions and quick check-ins occur weekly or even daily.

Tools that structure the work of the RPP and supply a scaffold for members as they collaboratively plan, reflect, and decide together are also critical. These include protocols, learning agendas, and logic models. For example, collaboration protocols, or step-by-step instructions that often include guiding questions, are used to structure RPP activities such as generating research questions, reviewing and interpreting data, determining the implications of findings, and prioritizing among options.

A third component of our RPP infrastructure is technology. Effective RPPs need specialized platforms and software for developing measures, collecting data, and managing, analyzing, and visualizing. Such specialized platforms are generally available to researchers (at least those in traditional research organizations), but they're not commonly used by practitioners or by most organizations that employ them. Thus, these platforms often need to be integrated with other systems being used to manage and implement the ECCE program itself, such as a management information system for programs directly serving children and families (such as ChildPlus, for example) or a

learning management system for professional development programs (such as CornerStone on Demand). Integrating these technologies so that they can be used by staff to support the work of an RPP typically requires a skilled information technology team.

Finally, funding has also been critical to building capacity for our RPPs. Without sufficient funding for the required infrastructure, the time needed to carry out the work, and an interdisciplinary work force, fully integrated RPPs wouldn't be possible. Unfortunately, grant requirements and funding structures don't always make it easy to use funds for this sort of capacity building and partnership. At Start Early, we're still working to determine the best approach to fundraising and fiscal management for conducting research and evaluation through embedded RPPs. But an advantage of having research and practice partners embedded in a single organization is that we can create efficiencies across our multiple RPPs. For example, research and evaluation functions and team members can be positioned as a shared service, akin to information technology or communications in some organizations. In our RPPs, this structure helps us access funding, such as philanthropic and private donations, to support our RPPs' general operating costs. These types of funds may be harder to get in an RPP made up of partners working across two or more organizations.

Overall, these four infrastructure components—and the ways they're integrated with one another—give RPPs the flexibility to use the data they generate for multiple purposes and to answer a variety of questions, many of which have not been defined in advance. At the same time, the infrastructure should also be durable enough to sustain the partnership and RPP activities over time and across individual projects. Start Early aims to achieve this goal by building every component of our RPP infrastructure in such a way that they first and foremost support the design, implementation, and improvement of programs and interventions; the fact that doing so ensures robust data collection that meets research standards is a bonus of program implementation as usual.

Challenges and Successes of Our Embedded Partnerships

Our embedded RPPs have benefited greatly from these three capacities; they have allowed both practice and research partners in our RPPs to achieve meaningful successes. Yet building these capacities in a practice-oriented organization has also involved several challenges along the way. Next, we discuss some of the capacity-building challenges and successes that we've encountered in our Essential Fellowship and Educare Chicago RPPs.

Demand and Efficiencies

Overall, our embedded RPPs have inspired positive attitudes (and less fear) about evaluation, data, and translating research evidence among our practice partners, as well as new mindsets of continuous learning and improvement that guide their day-to-day work. As a result, we are seeing more demand among Start Early program implementers for evaluation evidence and robust data and a stronger appetite for participating in RPPs. This is a success, but it has also presented a significant capacity challenge for our research and evaluation team, which is now engaged in multiple embedded RPPs simultaneously with programs in different phases of development, implementation, or scaling. To address this

challenge, we're working to standardize our approach to research and evaluation across RPPs by systematizing core evaluation procedures, methods, definitions, data use processes, and even some metrics and measures. In this way, embedded researchers may be in some ways better positioned to learn across programs than they would be in a more traditional structure that involves only one program.

It isn't always obvious who should take the lead.

Roles and Responsibilities

In our embedded RPPs, we can work in an integrated, interdisciplinary way. However, this means we need to be especially intentional and transparent about the division of labor and ensure a balance in assigning partners to lead various phases and functions of the work. It isn't always obvious who should take the lead on tasks such as setting the research or learning agenda, developing data collection tools, monitoring data collection and response rates, establishing and implementing data use routines, and writing reports to oversight agencies and funders. For example, Educare Chicago leaders and staff wanted to explore the relationships between students' attendance and daily arrival pickup times with learning and development outcomes. To answer these questions, their LEPs conducted analyses, found positive associations of attendance and student outcomes, and presented those findings to the school leadership team, who helped them to contextualize and interpret the findings. From that point, the research partners took the lead on integrating these findings into annual program reflection and goal-setting routines and developing graphics and brief data-based messages for families. Program partners took on the responsibility to communicate this information to staff and families as part of their larger efforts to improve attendance rates.

These roles are further complicated by the fact that Start Early's organizational structure includes multiple shared services, including marketing, information technology and knowledge sharing, instructional design, project management, and development. Because we're all employed by the same organization, all members of our embedded RPPs frequently interact closely with colleagues in these other disciplines. Although our ability to engage these various colleagues has undoubtedly improved the quality of our work, it can be challenging to assign roles, make decisions, and come to consensus on goals, priorities, and approach.

Communications and Dissemination

Effectively summarizing, communicating, and disseminating research findings to diverse audiences is a longtime challenge for our embedded RPPs. Many groups are invested in understanding and using information generated from direct assessments of Educare Chicago children, including program leaders, teachers and staff who work with families, parents and family members, our governing board and policy council, and program funders. Each group has distinct requirements for using the same data and findings as well as different interests and capacities. In response, LEPs have generated reports and routines that present and explore child assessment data separately for each group. Although this approach has been largely successful, it is also extremely

time consuming and has often tested the limits of our communication capacities.

Other RPPs likely face similar challenges: Successfully communicating research and evaluation findings requires skills such as writing for different audiences, data visualization, graphic design, and marketing. Yet researchers' academic training often prepares them mostly for communicating research results to other researchers. And practitioners may not have the skills, interest, or time to take on this role. The embedded nature of our RPPs has let us address this challenge head on: our organization has worked to build communications capacity—including both human capital and funding—to better support RPPs in dissemination activities.

Data Systems

Many RPPs face challenges in establishing an infrastructure—which includes selecting and implementing technology platforms—that will enable them to collect, analyze, and report data. This is particularly hard when the RPP intends to use data for more than one purpose. In the Essential Fellowship, for example, we've struggled to link individual participant records from their initial registration to their responses to surveys that generate data for evaluation and continuous improvement—a data infrastructure capacity that we need to support both program implementation and research.

Despite these challenges, we've partnered with our information technology team to create sustainable, integrated systems of data collection and sharing that our practice partners now rely on for robust information about their programs. For example, our Essential Fellowship practice partners can now access real-time online data dashboards, automated emails that send completed online surveys back to participants immediately, and job aids for coaches that support in-the-moment reflection and data-informed practice in a way that also captures reliable progress and interim outcomes. As a result, they can now use these data independently, as well as in monthly "research-to-practice" meetings with their research partners.

These data systems have also given research partners access to large, often longitudinal administrative data sets with relevant and rigorous implementation and outcome data. Importantly, because research partners in our embedded RPPs are so deeply integrated with the program teams and collaborate on designing the infrastructure for administrative data, they are able to contribute to decisions about that data, including what is measured and when and how it is collected. For example, Educare Chicago participated in an initiative that addressed alignment between early childhood and K–12 schools. From this initiative, it concluded that a common data source about students' learning and outcomes could be an important alignment support. To help create this support, LEPs trained preschool teachers to administer, score, and use data from a formative assessment of early literacy knowledge and skill that was also being used in K–5 schools; as a result, we were then able to confidently use that data as an indicator of children's learning.

Recommendations for Researchers and Practitioners

Embedded RPPs aren't typical in the early childhood field. Most often, researchers and practitioners in an RPP are employed by separate organizations and operate relatively independently, even as they form trusting

and mutually beneficial collaborations in the context of research projects and grants. Organizations, researchers, and practitioners interested in organizing an RPP for the purpose of long-term mutual learning, continually improving implementation, and generating innovative evidence and practice must work to build and integrate the capacities to support their work together. But doing so requires all parties to operate differently than they otherwise would. Here we offer reflections and recommendations for building the capacities to support these ways of working—for RPPs embedded in the same organization and for nonembedded partnerships as well—that are based on our own experiences. We recognize that we also need more research to investigate the benefits, drawbacks, and ideal mechanisms for building these capacities in research and practice organizations.

First, in any successful RPP, researchers and practitioners will need to spend a large portion of their time functioning at the boundaries of their discipline. In embedded RPPs, this expectation is clear to partners by virtue of the fact that they work for an interdisciplinary organization; indeed, both research and practice partners in embedded RPPs have likely knowingly self-selected into such boundary-spanning organizations and roles. Such an expectation may be less clear to partners in nonembedded RPPs. Key strategies for success in any RPP include aligning around a diverse set of aims and goals, respecting all perspectives, showing understanding for the discipline-specific issues partners face, and allowing partners time to assimilate and accommodate new understandings, methods, and skills. But researchers in particular aren't typically trained to work in these ways, and they're often employed by institutions with incentive structures that privilege scholarly publications over such partnership-building activities. Thus even as RPPs gain popularity, working as part of an RPP remains, in many ways, peripheral. To support the success of RPPs—and perhaps encourage more researchers to choose to work in practice organizations—more graduate education programs will need to include training and experiences that expose junior scholars to embedded partnership work and applied research. This is already standard practice in several doctoral training programs as well as the Institute for Education Sciences' Predoctoral Interdisciplinary Research Training Programs.

In any successful RPP, researchers and practitioners will need to spend a large portion of their time functioning at the boundaries of their discipline.

Second, we have found the most success and sustainability when research and practice partners construct the RPP together in ways that produce co-ownership of the processes and results. In an embedded RPP, this kind of co-ownership may occur more readily, because partners working in the same organization may be primed to think of themselves as on the same team. In nonembedded RPPs, achieving true co-ownership of processes and results may take more effort. Indeed, in many traditional RPPs, the research partners drive or own much of the data collection and monitoring. We recommend instead that researchers—whether within or outside of practice

organizations—help practitioners develop the capacity to take over some of the data gathering and processing responsibilities so that they can use this capacity independently or with minimal support from researchers to monitor their implementation and use data to make decisions. Data gathering and processing may represent a new skill set for program leaders and staff and may not be how practitioners are accustomed to working with research partners. Clearly defining and revisiting capacity, roles, and processes within and between the research and practice partners can help support RPPs with a division of labor that draws on each partner's strengths and areas of expertise while also preserving partners' time and independent decision-making authority.

Third, although there are strong examples of RPPs with research and practice partners in different organizations and locations (for example, see the article in this issue by Christina Weiland and Jason Sachs), we've found that having partners who share organizational routines, systems, and culture—indeed, often work in the same office—can facilitate an RPP's success. In embedded RPPs, with research and practice partners working side by side in a single organization, impromptu conversations occur frequently. At Start Early, these informal conversations have led to new insights, raised questions that might not otherwise have been asked, shaped our research and evaluation design, and helped with interpretation of research findings and data. Moreover, conversations about day-to-day program implementation have often revealed program aspects that need to be codified or better specified or ways that metrics, measurement tools, and data sourcing could be improved. For example, Lead Learn Excel's implementation is measured and tracked through a data collection tool that program implementers complete regularly. The extent and magnitude of adaptations made to the program model in practice were only revealed to the embedded research partners through informal, impromptu conversations that took place around the office. We were then able to take this "found" information back to both our practice and to UIC evaluation partners and use it to revise the tool so that it more effectively captured critical implementation details. We thus recommend that research and practice partners make every effort to engage with one another as frequently as possible outside of formal RPP meetings and routines.

Finally, establishing a fully integrated RPP is likely to require a long startup period in which each partner and the partnership together must focus on capacity -building if either partner is to realize many of the benefits. In this way, researchers and practitioners who take this approach must be both patient and generous. But in the long run, taking this approach simultaneously builds an integrated capacity and infrastructure that meets the needs of both practitioners and researchers (rather than just one or the other). Therefore, we recommend that both research and practice leaders take a more active role in obtaining funding that can provide access to the organizational structures, human capital, and infrastructural resources that enable this work. Researchers and practitioners working in an RPP must then plan how to enact RPP processes and collaboratively design, build, implement, and refine the infrastructure to make the RPP successful and sustainable. This may be easier to achieve with an embedded RPP, in which all resources are going to the same organization.

Conclusions

In this article, we've argued that fully integrating research within practice and practice within research shows great promise as a sustainable and effective approach to organizing the capacities needed to support RPPs. Based on our experience, RPPs can more readily develop the qualities Coburn and colleagues argue define the successful RPP (mutualism, commitment to long-term collaboration, a focus on problems of practice, the use of intentional partnership strategies, and trusting relationships) when research and practice capacities are inextricably entwined. We have accomplished this by embedding both partners in a single organization; this structure offers distinct benefits that make embedded partnerships an especially promising approach to generating knowledge and improving the quality of early childhood interventions. We hope that the necessary capacities will become increasingly more common in both research- and practice-oriented organizations, paving the way for new embedded RPPs in years to come.

Endnotes

1. Vivian Tseng, "The Uses of Research in Policy and Practice," *Social Policy Report* 26, no. 2 (2012): 3–16.
2. Vivian Tseng, John Q. Easton, and Lauren H. Supplee, "Research-Practice Partnerships: Building Two-Way Streets of Engagement," *Social Policy Report* 30, no. 4 (2017): 3–16.
3. Cynthia E. Coburn, William R. Penuel, and Kimberly E. Geil, *Practice Partnerships: A Strategy for Leveraging Research for Educational Improvement in School Districts* (New York: William T. Grant Foundation, 2013).
4. Barbara Means and Christopher J. Harris, "Towards an Evidence Framework for Design-Based Implementation Research," *Yearbook of the National Society for the Study of Education* 112 (2013): 350–71.
5. Robert C. Granger, Vivian Tseng, and Brian L. Wilcox, "Connecting Research and Practice," in *Societal Contexts of Child Development: Pathways of Influence and Implications for Practice and Policy*, ed. Elizabeth T. Gershoff, Rashmita S. Mistry, and Danielle A. Crosby (New York: Oxford University Press, 2014), 205–19.
6. William R. Penuel et al., "Conceptualizing Research–Practice Partnerships as Joint Work at Boundaries," *Journal of Education for Students Placed at Risk* 20, nos. 1–2 (2015): 182–97.
7. David A. Garvin, Amy C. Edmondson, and Francesca Gino, "Is Yours a Learning Organization?," *Harvard Business Review* 86, no. 3 (2008): 109–17; Peter M. Senge, *The Fifth Discipline: The Art and Practice of the Learning Organization*, rev. ed. (New York: Doubleday, 2006).
8. Vivian Tseng and Sandra Nutley, "Building the Infrastructure to Improve the Use and Usefulness of Research in Education," in *Using Research Evidence in Education: From the Schoolhouse Door to Capitol Hill*, ed. Kara S. Finnigan and Alan J. Daly (Cham, Switzerland: Springer International, 2014), 163–75.

Philanthropic Support for a Research-Practice Partnership

Jacqueline Jones

Summary

Many if not most research-practice partnerships depend to at least some degree on funding from philanthropic organizations. In this article, Jacqueline Jones discusses how and why the Foundation for Child Development decided to invest in such a partnership, the New York City Early Childhood Research Network, as the city was building its universal prekindergarten program. She also explains why the foundation chose the type of partnership known as a *research alliance*—a long-term, mutually beneficial collaboration that promotes the production and use of rigorous research about problems of practice.

Funders, Jones writes, are primarily concerned with the impact of the work they support. Yet traditional research activities may take years to be complete, and it may take even longer to determine whether the research had any impact on policy or practice. In a place-based research-practice partnership, collaborative construction of research questions ensures that the work is relevant, and rapid response research models mean that policy makers and practitioners begin to get answers to their questions—and funders begin to see the impact of their investment—much sooner. In this way, Jones writes, research-practice partnerships provide context-relevant data that can lead to quick policy changes, making them rewarding investments for funders.

Jacqueline Jones is president and CEO of the Foundation for Child Development.

Research-practice partnerships (RPPs) can be compelling experiments in connecting research and practice. But RPPs need financial support, which often comes from philanthropy. How can the philanthropic community help RPPs initiate, expand, and sustain their work?

Every philanthropic entity hopes that its investments will have a significant impact on the populations important to the organization, and local events can serve as the catalyst for funding opportunities. The Foundation for Child Development's mission is to build better lives for young children by bringing research to bear on matters related to improving policy and practice. We're based in New York City, where a push to make preschool available to all the city's four-year-olds inspired the development of an RPP, the New York City Early Childhood Research Network (ECRN).

The Context

September 2014 marked the beginning of the city's Pre-K for All initiative, spearheaded by the mayor. At about the same time, the Foundation hired me as its new president and CEO, and we began to rethink our programmatic goals and strategies. Fulfilling the Foundation's mission was guided by three overarching responsibilities:

- filling gaps in knowledge by identifying important areas of policy and practice that need more empirical evidence and by building new knowledge in those areas,
- influencing policy by providing policy makers with the research evidence necessary to make informed decisions about young children's learning and development, and
- exercising good stewardship of the Foundation's financial, human and reputational assets.

In this context, the board and I were also aware of growing national attention to the training and compensation of and support for the early childhood workforce: teachers, center directors, elementary school principals, and other early childhood professionals who work directly with young children and their families to provide high-quality early learning experiences.[1] Although the number of early childhood programs had significantly increased over several decades and a robust body of knowledge had developed, child outcomes remained modest and inconsistent.[2] The Foundation wished to examine the growing suspicion among researchers that the wide variation in early childhood teacher preparation programs and requirements for teacher certification and licensing were primary factors in the continuing disparities in early childhood program quality and child outcomes across a variety of settings. Our board and staff believed that if early childhood educators are to offer young children the kind of high-quality learning experiences that can make a significant contribution to closing opportunity and achievement gaps, those educators must be well prepared, appropriately compensated, and regularly exposed to meaningful professional learning experiences.[3]

The theory of change was straightforward: Enhancing the knowledge, improving the skills, and increasing the compensation of early childhood professionals had significant potential to improve program quality and lead to stronger outcomes for young children. The Foundation divided its programmatic focus on the early childhood

workforce into three areas. One was improving the quality of early childhood practice through implementation research, which isn't just a matter of studying a program's impact but also, as the research firm MDRC puts it, "investigat[ing] *how* the program produces those impacts," using "both quantitative and qualitative data to assess the programs and policies that are the subject of an evaluation."[4] The Foundation's goal was to promote implementation research to provide a deeper understanding of what works, for whom, and under what conditions in early childhood education programs. New York City's implementation of Pre-K for All offered the Foundation an opportunity to pursue this goal.

The Foundation's involvement in Pre-K for All began after the Obama administration launched its Invest in US initiative in December 2014. Invest in US was intended to catalyze an increase in private sector funding for early learning initiatives across the country. Though the Foundation was founded and based in New York City, it had become national in scope and hadn't made significant investments in its home city for some time. Several board members felt strongly that Pre-K for All presented an opportunity to support a New York City initiative that was aligned with our emerging program focus on the early childhood workforce. Through Invest in US, the Foundation committed $2 million to support the city's universal preschool effort.

Evaluating the program in the standard way—through a randomized controlled trial—could highlight differences between children who attended Pre-K for All and those who didn't. But we thought that the city might benefit from a more nuanced understanding: what components of Pre-K for All were being received by and having an impact on specific subgroups of children and under what conditions? This type of information might better guide the city's efforts at continuous quality improvement. Since the Foundation is also interested in enhancing the early childhood workforce, we were particularly interested in the role that lead teachers, assistant teachers, coaches, and others play in providing high-quality experiences for children.

Why Invest in an RPP?

Beginning in the winter of 2014–15, our challenge was to develop spending guidelines for producing results that, from a research perspective, were aligned with the Foundation's mission and programmatic areas and, from a public policy perspective, were useful to the city in executing Pre-K for All.

From a research perspective, four factors prompted our interest in RPPs. The first factor was the potential of RPPs to build strong and meaningful connections between research and policy, a need that's been recognized for some time. The second factor was the potential to foster scholarship, especially to support early career young scholars. By 2014, the Foundation's Young Scholars Program, in operation since 2003, had already funded over 50 early career scholars to conduct empirical studies related to young children's wellbeing, and continuing that tradition was essential to the Foundation's board of directors. The third factor was the Foundation's emerging interest in implementation research. Few studies had been conducted on using implementation research in early childhood education, and our board and staff believed that supporting

an RPP could advance the Foundation's long-term goal of using implementation research to deepen the body of knowledge on early childhood programs.[5] The fourth factor was the role that other philanthropic organizations, such as the William T. Grant Foundation and the Spencer Foundation, had taken in supporting RPPs.

Thus we began to seriously contemplate supporting some type of RPP that would resonate with the Foundation's history of funding research that connects to policy and practice. Historically, most philanthropic funding for such partnerships had been directed either to existing RPPs or to organizations that had decided to collaborate on their own. With these considerations in mind, we decided to proceed.

The Foundation approached this initiative in a learning mode. We hoped that developing and supporting a rather complex RPP could produce useful research information and offer the philanthropic community insight into the roles that funders might assume in connecting research to policy and practice. By the summer of 2015, a plan was emerging.[6]

Building an RPP Network

Though there are many types of RPPs, a *research alliance*—typically defined as a long-term, mutually beneficial collaboration that promotes the production and use of rigorous research about problems of practice—seemed the best fit for the Foundation's support.[7] Such long-term partnerships between school districts and research organizations seek to solve specific problems of practice or policy. In these place-based RPPs, scholars and school districts construct research questions together and continue to collaborate as the partnership matures. The RPP generally conducts the research and communicates its findings back to the district and other interested parties with the intent to guide policy making and improve practice in the district.

Engaging Multiple City Agencies

Working within a complex multiagency system, the New York City Department of Education (DOE) took the lead in developing Pre-K for All, but three other agencies were also involved: the Administration for Children's Services, the Department of Health and Mental Hygiene, and the Center for Economic Opportunity. To expand existing services to make full-day preschool programs available to all four-year-olds in the city, Pre-K for All was delivered not only in schools but also in private child care and federally funded Head Start centers.

As the federal grantee for several Head Start programs and as the administrative entity for the federal child care subsidy system, the Administration for Children's Services had programmatic authority over many community-based Head Start and child care centers. The Department of Health and Mental Hygiene held licensing authority over such sites. In addition, the mayor had been the chief advocate for Pre-K for All, and the Center for Economic Opportunity, within the mayor's office, was deeply involved in the program's planning and implementation, which occurred in a highly political context. Though it would have been easier if only the DOE had partnered with researchers, the Foundation believed that the RPP it was funding needed representatives from

all four city agencies. This would allow a more complete picture by including voices that represented multiple aspects of the program's implementation. The Foundation was also interested in learning whether the RPP's research findings could be useful for other city agencies beyond DOE.

Engaging Multiple Research Partners

The New York City school district—the largest in the country—is no stranger to research studies. The city houses almost 120 higher education institutions and a school district of approximately one million children and thus offers many opportunities to explore the nature of teaching and learning. The Foundation knew that researchers from multiple institutions of higher education had established long-standing relationships with city agencies—for example, the collaboration between the DOE and New York University described by Rachel Abenavoli and colleagues elsewhere in this issue. However, we wondered if it was possible to enhance the district's research capacity by coordinating a group of researchers from multiple universities, all of whom had an interest in studying the implementation of early childhood programs in the city. We were particularly intrigued by the idea of a research alliance RPP with a network of university researchers from across the New York metropolitan area who could work together to collect and share data on the early implementation of Pre-K for All through a series of small-scale studies. Not only would this give the city access to richer data and facilitate cross-institutional collaborations, but it might also lead to the development of a larger birth-to-third-grade research agenda.

An Organizing Entity

It became clear early on that a backbone organization would be needed to develop the group's infrastructure and manage logistical tasks such as organizing convenings and setting agendas for meetings. The Foundation engaged the New York Professional Development Institute (PDI), an established and respected organization that understood the city's early childhood landscape and was capable of working with both the research community and with city agency personnel.

Research Parameters

The Foundation needed to be very clear about the purposes and potential consequences of the research to be funded. Before convening researchers and city agency representatives, we set important parameters for the types of research we would fund and the manner in which the work was to be conducted. These were communicated clearly and directly to researchers and policy makers alike.

Implementation, not Evaluation

The project didn't aim to evaluate Pre-K for All's effectiveness. We explicitly assured representatives of city agencies and researchers that studies funded by the Foundation wouldn't constitute a program evaluation. Rather, the project's primary intent was to give city agencies useful information to continuously improve the quality of Pre-K for All as it was being implemented. Researchers also understood that their studies were intended to give policy makers deeper information about how specific components of the program were operating. The Foundation worked to build an atmosphere of trust by strengthening

prior relationships with researchers and city officials as well as establishing new ones.

Constructing Research Questions Collaboratively

Because research findings needed to be meaningful to city agency representatives, it was critical that the research questions were important to policy makers and could be examined through empirical studies. Thus we aimed to form an RPP in which researchers and policy makers would work together to create the research questions.

Early Childhood Workforce

We made clear to both researchers and city agency representatives that a focal point of the studies was to be the role of various members of the early childhood workforce. Considering the broad nature of the early care and education workforce, the Foundation requested that the studies focus on the workforce from three broad perspectives: instructional practice, workforce status, and professional development.

Place-Based Strategy

Given the size and diversity of New York City, it was important to define the geographic areas in which the research was to be conducted. The Foundation was interested in understanding how Pre-K for All was implemented in high-, medium-, and low-income areas of the city. With assistance from New York University and the Center for Economic Opportunity, three community districts were identified at each income level, for a total of nine. All of the studies funded by the Foundation would be conducted in these nine districts.

Once the research parameters were set, representatives from the four city agencies were told of the nature and intent of the work and asked to commit to attending regularly scheduled meetings with researchers. Everyone agreed, and PDI staff began the work.

Convening and Proposals

During the summer of 2015, 17 researchers representing eight universities in the New York/New Jersey metropolitan area met with Foundation staff and the city officials representing the four agencies responsible for implementation of Pre-K for All in New York City. The Foundation's goals were explained, and the scope of work expected from the group was outlined. In August, we released a request for proposals from researchers.

After reviewing all the proposals but before making any awards, the Foundation's staff met again with the city agency representatives to review the proposals that we had targeted for funding and make sure the work was aligned with policy makers' needs and goals. The first three grants were awarded in November 2015. Another five were awarded over the next year, for a total of eight.

As proposals were being reviewed and the first set of grants were awarded, it also became clear that conducting multiple studies in nine community districts would require significant coordination to ensure that researchers didn't overwhelm educators with requests for access and jeopardize data collection. The Foundation decided to engage an external research firm, MDRC, to develop a coordinated sampling strategy for the eight research projects across the nine districts. MDRC researchers met with each research team, designed a sampling protocol for all the studies, and assigned sites

to each research project. City agencies then wrote letters of introduction and support to the assigned sites for the research teams. This approach ensured that the researchers were able to gain access to the number of participants they needed and that the burden to programs was minimized.

In line with the Foundation's goals, the studies concentrated on taking a close look at the initial aspects of the program's implementation. The nature of the questions developed in the very early stages of Pre-K for All reflected what policy makers needed to understand as the program took shape. From the Foundation's perspective, the initial studies set the stage for a more complete story that could be told about the development of New York City's universal preschool initiative. Study topics included practices to support dual language learners, men in early childhood education, and the use of formative assessments to guide instruction. After funding the eight studies and engaging PDI and MDRC for logistical and sampling tasks, the Foundation ended up committing approximately $2 million more.

What Would Success Look Like?

The Foundation's decision to develop an RPP with multiple universities and multiple public agencies meant that the definition of success would comprise many elements and that understanding the RPP's successes and challenges would take some time. We turned to the five dimensions of effectiveness for RPPs developed by Erin Hendrick and colleagues for the William T. Grant Foundation to help philanthropic entities gauge the success of their investment in such partnerships.[8] They include:

1. Building trust and cultivating partnership relationships.
2. Conducting rigorous research to guide action.
3. Supporting the partner practice organization in achieving its goals.
4. Producing knowledge that can guide educational improvement efforts more broadly.
5. Building the capacity of participating researchers, practitioners, practice organizations, and research organizations to engage in partnership work.

Table 1 shows progress to date in each of these dimensions.

The Foundation has also gauged progress with respect to important programmatic considerations. One of these was the intent to foster scholarship among its research members and support early career young scholars, which so far appears to be succeeding. It was important to the Foundation to support research by faculty in New York City's own higher education system, the City University of New York. These faculty typically have higher teaching loads than their colleagues in private institutions, limiting their ability to conduct primary research. So far, at least two tenure decisions and one promotion have likely been enabled by the Foundation's research funding. ECRN has also been able to launch an early career scholars program, which has supported one doctoral candidate and two postdoctoral researchers in their work in applied research that has implications for improving early childhood policy and practice.[9]

Another measure of success is the ability to attract support from other funders.

Table 1. Understanding the Foundation's Investment: Five Dimensions

Dimension 1: Building trust and cultivating partnership relationships	After five years of work, researchers and policy makers continue to meet regularly. Each New York City agency has seen key personnel changes, and the policy makers who helped create and signed on to the original research questions in 2015 often weren't the same ones who received the initial study results in 2018. Incoming policy makers had to be brought up to speed on the history of the RPP and the rationale for the research questions. Thus building trusting relationships has been an ongoing process. DOE policy staff report that attending the research network meetings, hearing about the studies before results were publicly released, and collaborating on press releases reassured them that their work would be based on the latest research.
Dimension 2: Conducting rigorous research to guide action	Initial studies were primarily qualitative in nature, in line with the questions developed with city agencies. Though all studies were of good quality, more experienced researchers brought a higher degree of rigor to their work. However, junior researchers were able to hone their research skills and learn from exchanges with their more experienced colleagues.
Dimension 3: Supporting the partner practice organization in achieving its goals	The DOE used ECRN findings from Beverly Falk and Marianna Souto-Manning's study *Quality UPK Teaching in Diverse Settings* and input from the researchers to help create their *Early Childhood Framework for Quality*. The descriptive findings of the collaboration between NYU and the DOE (see Rachel Abenavoli and colleagues in this issue) on the key role of leaders in the early childhood education centers' advice networks (for example, in fostering professional development and informally modeling good practices) helped persuade the DOE that its leadership training initiatives were on target and should be continued and expanded. Descriptive findings regarding social and professional connections and teacher wellness confirmed the importance of the department's interest in understanding and supporting teachers' wellbeing, a critical focus during the COVID-19 pandemic.
Dimension 4: Producing knowledge that can guide educational improvement efforts more broadly	DOE's director of data and analytics for early childhood has consulted with ECRN members on research and resources related to costs associated with implementing high-quality prekindergarten programs in community-based sites. The work of the network is being disseminated to three audiences: New York City policy makers, the broader research community, and the philanthropic community. For example, presentations on the network's development and progress have been made at meetings of the American Educational Research Association and the Association for Public Policy and Management
Dimension 5: Building the capacity of participating researchers, practitioners, practice organizations, and research organizations to engage in partnership work	The work of ECRN is still in its infancy, and whether its members value engagement in long-term collaborative inquiry enough to develop the capacity to support such engagement is still to be determined. The foundation is eager to learn whether city agencies will develop a solid culture around the use of research and evidence and whether they will continue to document any impact on public policy resulting from the partnership.

Source: Erin Henrick et al., *Assessing Research-Practice Partnerships: Five Dimensions of Effectiveness* (New York: William T. Grant Foundation, 2017).

We understood from the beginning that sustaining and expanding ECRN would require investments from other members of the philanthropic community. After the Foundation alerted other early childhood education funders to the RPP's work, the Heising-Simons Foundation made a $400,000 grant to support three research studies on infant and toddler development and to continue PDI's infrastructure work, and Early Childhood Partners NYC provided $100,000 for PDI. The W. Clement and Jessie V. Stone Foundation and the Spencer Foundation have also provided funding.

Implications for Philanthropy

Although the Foundation used a somewhat unorthodox strategy for developing an RPP, our process has implications for other philanthropic organizations. What follows are several insights we've gained in developing ECRN that may be useful to others in the philanthropic community.[10]

Whether they're providing resources for direct services or for research studies, funders are primarily concerned with the impact of the work they support. Funders who support research activities must often wait several years for the research to be completed—and even longer to determine whether the research had any impact on policy or practice. The place-based nature of RPPs offers the potential to fund research that investigates real-world problems of policy and practice and is focused on a specific context. The collaborative construction of research questions ensures that the work is relevant. In an RPP, researchers can work to design studies that are aligned with policy makers' timelines. The use of rapid response research models and continuous dialogue among researchers, policy makers, and practitioners as the research is being conducted can reduce the time needed for policy makers to begin to get answers to their questions. Funders are exploring strategies that can shorten the time to bring relevant empirical evidence to policy makers and practitioners. During the COVID-19 pandemic, policy makers couldn't wait several years to develop research-based policies to slow the spread. As researchers worked on rapid response models, many funders expedited their proposal review protocols and were able to distribute funds faster than usual, thereby allowing researchers to tackle policy and practice questions as quickly as possible.

Conclusions

Based on the ECRN experience, we at the Foundation have come to understand that RPPs require cultural shifts for both researchers and policy makers. The researchers appear to be most successful when they view the primary goal of these partnerships as the search for empirical evidence that will assist policy makers as they struggle with difficult program- and policy-related decisions. Thus researchers would do well to gain an understanding of the issues that their key policy partners face and the scope and limitations of their authority. Those with the ability to implement researchers' recommendations may be several rungs up the hierarchy from the policy partners with whom the researchers meet. Recommendations that fall outside the parameters of a policy maker's authority can be more frustrating than helpful. However, the research results may provide powerful arguments that lower-level policy makers can draw on to

advocate to those at higher decision-making levels.

In addition, policy makers can support the effectiveness of an RPP when they are able to see the utility of using evidence as a basis for their decisions and when their participation is based on the conviction that the researchers are genuinely trying to help improve the quality of programs. These partnerships take time and require long-term commitment, perseverance in the face of changing political landscapes, and ongoing external support. Yet RPPs clearly have potential to be rewarding investments for funders for the context-relevant data they can provide, the potential for quick policy changes, and the opportunity to build trusting relationships between researchers and agency representatives. They are well worth the investment.

Endnotes

1. Institute of Medicine and National Research Council, *Transforming the Workforce for Children Birth through 8: A Unifying Foundation* (Washington, DC: National Academies Press, 2015).
2. Center on the Developing Child at Harvard University, *From Best Practices to Breakthrough Impacts: A Science-Based Approach to Building a More Promising Future for Young Children and Families* (Cambridge, MA: Center on the Developing Child, 2016), https://developingchild.harvard.edu/resources/from-best-practices-to-breakthrough-impacts.
3. Institute of Medicine and National Research Council, *Transforming the Workforce*.
4. MDRC, "What Is Implementation Research?," n.d., https://www.mdrc.org/what-implementation-research.
5. Foundation for Child Development, *Getting it Right: Using Implementation Research to Improve Outcomes in Early Care and Education* (New York: Foundation for Child Development, 2020), https://www.fcd-us.org/assets/2020/06/GettingitRight_UsingImplementationResearchtoImproveOutcomesinECE_2020.pdf.
6. Foundation for Child Development, *The New York City Early Childhood Research Network: A Model for Integrating Research, Policy, and Practice* (New York: Foundation for Child Development, 2018); Ruth N. López Turley, "Connecting Research and Policy to Reduce Inequality," *Russell Sage Foundation Journal of the Social Sciences* 2, no. 5 (2016): 272–85, https://doi.org/10.7758/rsf.2016.2.5.13.
7. Cynthia E. Coburn, Willian R. Penuel, and Kimberly E. Geil, *Research-Practice Partnerships: A Strategy for Leveraging Research for Educational Improvement in School Districts* (New York: William T. Grant Foundation, 2013), https://wtgrantfoundation.org/library/uploads/2015/10/Research-Practice-Partnerships-at-the-District-Level.pdf.
8. Erin Henrick et al., *Assessing Research-Practice Partnerships: Five Dimensions of Effectiveness* (New York: William T. Grant Foundation, 2017).
9. New York City Early Childhood Research Network, "Early Career Scholars," n.d., https://www.earlychildhoodresearchny.org/about/our-early-career-scholars.
10. Foundation for Child Development, *Early Childhood Research Network*.

COMMENTARIES

A Dean's Perspective

Robert C. Pianta

I write this commentary from two perspectives: that of an academic researcher who for several decades has relied on partnerships with practitioners and agencies in the field, and that of a dean of a school of education. From both perspectives, it's clear that research-practice partnerships are essential if research in the social and educational sciences is to be both strong and relevant—that is, if research is to have impact. I come to this belief from experience in a wide array of research-practice partnerships, most of which, fortunately, have been constructive. Here I reflect on the benefits and challenges of these experiences and the types of support that higher education institutions must provide to sustain these essential components of research infrastructure and career development. Research-practice partnerships are also critical to a university's service mission and to advancing practice and professional education. Three themes frame my comments: that partnerships are strategically important for strong and relevant social science, that effective partnerships require explicit investment and development strategies, and that partnerships are developmental contexts for scholars.

An Essential Tool

The value of scholarship in social and educational science is most often described in terms of the metrics employed in academia—journal impact factors, citations, peer review. Although these factors clearly help determine the value or strength of research from the perspective of the scientific community, they don't fully indicate the value of scientific research for parents, community leaders, policy makers, and the public at large. At times, people outside academia criticize social and educational science research for not yielding results that can help them in their work and for the ways that scholars communicate their results. Thus it's not surprising that we're paying closer attention to increasing our research's relevance.

Scholarship is relevant to the extent that it can answer questions that are important to interested people outside academia, can be understood and used by those people, or can influence the decisions and actions of practitioners and policy makers. Because partnerships across the research-practice boundary enable the exchange of experience, information, and perspectives among academics and stakeholders, they are an

Robert C. Pianta is dean of the School of Education and Human Development, the Novartis US Foundation Professor of Education, a professor of psychology, and the founding director of the Center for Advanced Study of Teaching and Learning at the University of Virginia.

indispensable resource for individual scholars and institutions under pressure to produce social science that is both strong and relevant.

Stronger Science

From the point of view of the traditional metrics of academia, effective partnerships with practitioners (either individuals or organizations) add value in a number of ways. First, partnerships provide insights into questions, processes, mechanisms, and variables of interest to investigators—insights we can't get without the information and perspectives of field partners. These insights and this information help scholars create more refined measurement tools, develop data collection plans and protocols that are more likely to capture the phenomena we want to study, and advance a much more informed and comprehensive interpretation of the results we obtain. Second, partnerships can help researchers take advantage of scale. For many research questions in the social sciences and particularly in education, sample size is a critical factor in designing and interpreting research. Many of our research questions involve malleable factors in education systems that fall into the framework of "what works for whom under what conditions." By their very nature, such questions require a sample that draws from populations larger and more diverse than, say, the set of local fourth-grade math classrooms. To make the science strong enough to yield interpretable results, it's essential to investigate processes and mechanisms across cultural, racial, economic, or other background characteristics of students or across assorted school or classroom features. The power of our statistical analyses, our ability to understand mediating and moderating processes, and our ability to follow a sample over time are all enhanced by increased scale. Scientific work that takes place at the level of school divisions (or districts, as they're called in many states), or that is representative of regions or populations of interest, is crucial to building knowledge that reflects reality. Research-practice partnerships that allow access to scale and to variation can thus strengthen a program of research and dramatically enhance the scientific value of a single study.

Relevant Science

Partnerships also enhance the relevance of scientific work. Again, relevance is important for both the research program of an individual investigator and for academic institutions. For example, in my present role in a public institution of higher education, I am reminded often of our mission to serve the public. To live up to that responsibility, we are obliged to produce scholarship that benefits practice. Our university regularly reports to state legislators and agencies on our work with state and local partners in the Commonwealth of Virginia, and research in education science is particularly meaningful to those policy makers and to the university's public mission. In this way, research-practice partnerships can build social and political capital for the value of education science. Partnerships also make research more relevant to practitioners and decision makers in their function as a feedback loop for scholars; working with partners in the field (for example, working together to design programs of research and data collection) helps investigators synchronize their research aims and methods with the needs and perspectives of the organizations and individuals they intend to serve. For individual investigators, partnerships in which we collect and use data from the field and pose questions of

interest to our practitioner partners make our work useful to those professionals and settings. For faculty, such partnerships are important because they expose them to the complicated nature of the questions they study, underscore the need for interdisciplinary teams, and guard against oversimplified recommendations from researchers. And interactions with practitioners sharpen scholars' engagement and presentation skills—fostering perspective taking, leadership and team building, and the use of clear, succinct, and jargon-free language—in ways that can't be replicated elsewhere. These skills also have downstream benefits for faculty in their roles as academics—they make them better teachers, enhance their scholarly presentations, and help them engage with the media and with funders.

Partnerships Need Investment and Development

Partnerships that link scholarship and practice often feature a faculty member or research team and an organization such as a school division or state agency who together focus on a question of mutual strategic interest. In most cases, these partnerships are initiated when one of the parties seeks the other out and begins a process of articulating interests and aims and establishing an operational working relationship. Partnerships always take time and effort and are often costly to initiate, develop, and maintain. My experience as a researcher over many years in dozens of such partnerships—and now more than a decade as an academic administrator interested in building effective partnerships that benefit both researchers and their partners in the field—has taught me that we need to be far more strategic about developing and maintaining a set of strong partnerships at both the individual and institutional level.

Partnerships between organizations can support the work of individual faculty members or programmatic research by teams of faculty, and they can help a university contribute to the public good. By fostering such partnerships, leaders can greatly enhance the value of their institutions' scholarly work and increase the likelihood of receiving sponsored support. Indeed, for the social, behavioral, and educational scientists who make up a considerable proportion of the faculty at a research university, the field is their laboratory. Thus effective research-practice partnerships are as important to the university's goals and mission as wet labs in the biosciences, imaging equipment in neuroscience, or servers for computer and data science. Investments for research-practice partnerships in the social and educational sciences conceived of as research infrastructure might include:

- dedicated funds in faculty start-up packages for activities and assets that support partnering, such as stipends for teachers, travel to sites, and so forth;
- funding for staff at the university and unit levels dedicated to identifying, managing, and sustaining partnerships, and;
- funds for communications and materials that help practitioners both participate in research and consume and use the research they help produce.

Budgeting for these forms of support should be a focus of an institutional research infrastructure plan that specifies the

parameters of effective research-practice partnerships.

But most universities' research infrastructure budgets have no dedicated funding stream for partnerships. Most of the cost of developing and maintaining effective partnerships is borne by schools and faculty, and thus it's not easily captured in financial accounting. Yet these assets are every bit as essential as the on-campus buildings and equipment that house and support the biomedical sciences and engineering and material sciences, and the work of scholars in related fields. Robust and sustained partnerships require dedicated personnel to maintain communication and relationships; they require subsidies for travel and meetings; and they require systems and personnel to manage data and interpretation. Universities would do well to recognize research-practice partnerships as core elements of research infrastructure and find dedicated financial support to develop and sustain them.

Such investments offer financial returns. Strong partnerships enhance the likelihood of securing sponsored support; for example, grant programs at the National Science Foundation and Institute of Education Sciences require documentation of partnerships. Sponsored grants supported by such partnerships often require fewer university resources and less overhead, and thus they may provide a greater return on investment than sponsored awards that rely on the expensive infrastructure needed in other areas of scientific work.

At the University of Virginia School of Education and Human Development, we've supported research-practice partnerships by investing in a long-term relationship with a group of school division superintendents whose schools represent 75 percent of the school-age population in the state. This relationship gives our academic and research programs and faculty access to a very large and diverse group of school leaders and the populations they serve, and it has led to numerous grants, pilot research studies, and dissemination activities aligned with faculty members' and research teams' interests. We've also invested in several research centers that have formed partnerships with practitioner organizations, state agencies, and school divisions in Virginia and nationally (indeed, most of these partnerships are with groups outside of the state). One of our research centers, working closely with school divisions and the Virginia Department of Education, has secured funding from the university to further enhance the partnership's capability to focus research on questions of mutual interest. This has required investments in personnel whose job it is to communicate across the university and with agency partners and to manage partnership processes and tasks. In another instance, a research center focused on adolescent development is partnering with dozens of practitioner organizations nationally to focus on redesigning middle school. Much of that partnership involves translating developmental and education science so that it's accessible to nonspecialists and building opportunities for experiments and replication.

Still other partnerships focus on the work of individual faculty and their research teams. Examples include one that works with local school divisions with which we share very close connections as training and research sites, one that works with a dozen large urban school divisions in and outside

of the state, one that works with government agencies in almost a half dozen states, one that works with large clusters of community colleges, and one that works with several multistate organizations that serve state agencies or practitioner groups (for example, the Southern Regional Education Board), to name a few. The need to keep track of these partnerships has prompted us to consider opening an office of partnerships through which we might manage these essential resources more strategically. In moving to a more strategic approach, one challenge we face is simply accounting for the hidden costs of this work and the tradeoffs between central supports and individual faculty-driven needs.

Partnerships Help Scholars Grow

As a young scholar, my primary aim was to establish a research program to examine how various indicators of school readiness, as well as children's relationships with parents and teachers, might forecast their later success in school. I was fortunate to be in a community where the school division leaders were interested in related questions and open to working with me. I started down the path of partnership as we planned the work, although I suspect I was more focused on preserving a functional relationship and a data collection protocol that would yield publishable results than I was on forming a partnership per se. The planning process was long. During more than a year of meetings with school division leaders, we made plans to undertake a large-scale assessment of children as they entered kindergarten, the results of which we hoped to quickly supply in a form that teachers could use; to identify questions that interested school leaders with regard to policies; and to write articles for peer-reviewed journals.

A team of doctoral students assisted me, but by and large I led these efforts without much institutional support beyond introductions. Our work was successful in many ways. The school division found the information useful. Teachers were eager to use the assessment time to interview parents, and they used the protocols we designed to assess students entering the prekindergarten program that was introduced a couple of years later. Familiarity with this information also helped division leaders make the case for the importance of preschool to the wider community. Doctoral students completed dissertations, I was awarded tenure, and our team contributed to research about the role of adult-child relationships in school success. And I learned a lot about forming and maintaining a research-practice partnership that served me well as my research program developed and expanded.

Partnerships with practitioners support scholars' own professional and personal development. Effective partnerships help faculty develop their perspective-taking and communication skills, which are extraordinarily valuable should they go on to assume leadership roles in their own institutions or be called on to consult with government or industry. Partnerships help faculty build networks that further enhance their careers. For example, a state agency that's had a positive experience with a faculty member may refer an agency in another state to that faculty member for consultation. In a very real sense, partnerships can multiply a scholar's capacity. The fact that partnerships enhance the connections between career development and research is one reason that in our school's Office of the Associate Dean for Research and

Faculty Development, we link research administration and support infrastructure with faculty development such as mentoring and seed investments.

Still, these benefits come with costs or present risks that are important to understand and mitigate. Partnerships are especially demanding for junior faculty, even more so when they're responsible for leading or initiating them. When such opportunities arise, then, junior faculty require strong mentoring and explicit support. Partnerships involve political, sociocultural, historical, and other contextual knowledge that most junior faculty haven't yet acquired. In fact, the lack of such knowledge can pose a risk to faculty and institutions if junior faculty are thrust or enticed into these roles. Thus, to the extent that junior faculty need to build and develop partnerships as a core component of a strong and relevant research program, the institution must ensure that they're supported. One form of support, as I have noted, is access to partners such as school divisions. Another may be seed funds for travel or stipends. But the most important one is mentoring.

Most partnership work simply requires a mentor who has experience in successful partnerships. Let me quickly offer an example of what such mentoring might look like. I received a call from a former postdoctoral student, now a productive tenure-track assistant professor working in a state university, who was approached by an organization interested in forming a partnership. She was awash in questions: "I have no idea how to navigate this relationship, . . . whether it makes sense for tenure, . . . [or] how to go about this." This very promising early-career investigator was facing a consequential decision, and she was able to draw on my experience to help guide her choice. But in many cases too few mentors are available.

If research-practice partnerships are to function effectively to develop junior scholars' careers (and support their strong and relevant scholarship), institutions need to take a programmatic approach to building their partnership skills. First, institutions need to provide information and resources, such as sessions that identify available partners and explain how to get access to them, or workshops on communicating with partners in a way that promotes engagement and research participation. Second, institutional leaders (such as associate deans) need to model engagement and facilitate introductions by bringing in leaders from partnership groups (such as school or division leaders or agency staff members) who can present and discuss questions of interest with early-career faculty. Third, early career faculty should receive programmatic mentoring in which they are assigned to mentors with specific partnership capabilities and relationships that can advance that faculty member's research program. These mentoring relationships should be reviewed annually to assess their quality and their support for early-career faculty's needs.

Research-practice partnerships are essential infrastructure for research in the educational and social sciences. Rigorous science that's relevant to an array of people, including both practitioners and scholars, is simply not possible without them. Research universities would do well to strategically develop and sustain the infrastructure such partnerships require and to train scientists to work in them in constructive and effective ways.

A Unique Opportunity for Education Policy Makers

Jenna Conway

When it comes to solving problems and measuring impact, public education organizations primarily rely on two approaches—using civil service to hire personnel directly or using competitive procurement to hire outside experts on contract. These two approaches can limit such organizations' efforts to gain insights, to incorporate cutting-edge research into policy and practice, and to develop innovative solutions quickly, nimbly and affordably. In this commentary, I discuss how public education organizations can use research-practice partnerships (RPPs), especially partnerships with public higher education institutions, to tackle new, unique, and complicated education problems and thereby help children.

Consider how traditional approaches can inadvertently limit public education organizations. In civil service, most jobs are full-time, classified positions with very specific descriptions that are subject to a range of restrictions to protect the rights of employees. It can be hard to gain new insights without procuring outside experts whose rates are significantly higher than most government pay scales. Similarly, public education organizations may have a research director or team, but they usually don't have the capacity to analyze all the available data or measure the impact of everything the organization does. For new initiatives or pilot programs, public education organizations rarely have the internal research or analytical capacity to conduct rigorous, real-time research on interventions as they are implementing them. Instead, they typically rely on outside research. Doing so entails a lengthy grant application or procurement process and requires an extended time period to collect and analyze data and produce findings. Finally, with procurement, a public organization must specify both the problem and proposed solution up front, establish constraints, and require bidders to define precisely what they will do and provide and how much it will cost. This can inhibit an iterative design process where prototypes can be tested and improved. To top it off, the procurement process itself can take a year or more. Procurement may work well for well-established projects and services, but it proves more challenging when a state is

Jenna Conway is chief school readiness officer at the Virginia Department of Education. She previously served as assistant superintendent for early childhood at the Louisiana Department of Education.

designing a new service, intervention, or tool, especially one that will be used by a large and diverse set of users.

In contrast, RPPs, particularly those involving public higher education institutions, offer a unique opportunity for policy makers to expediently and cost-effectively gain expertise, integrate real-time research into their policies and practices, and design and build innovative solutions in an iterative manner that better meets their needs.

First, RPPs allow public education organizations to gain cutting-edge expertise without having to hire full-time staff or procure expensive consultants, which may not be economically feasible. When I was assistant superintendent of early childhood for the Louisiana Department of Education, we established a multi-year partnership with Daphna Bassok at the University of Virginia School of Education. In an ambitious transformation of its birth-to-five early childhood system, Louisiana unified its child care, Head Start, and school-based prekindergarten systems. Specifically, over five years, Louisiana established uniform expectations across programs, supported local networks in every community, measured interactions in every classroom, offered incentives for quality improvement at the classroom level, and coordinated enrollment locally to make it easier for families to choose the best option for their children. Louisiana's approach, which was different from that of most other states, entailed significant change for early childhood educators and for the department. Operating from the hypothesis that teacher-child interactions are what matters most for child outcomes, Louisiana rapidly gathered classroom data from thousands of child care, Head Start, and school sites across the state.

It would have been hard to hire staff or craft a contract with specific targets when we were unsure what the data would reveal and when no other state or entity had done what we were doing. In exchange for our collaborating on research projects (and the grant proposals that preceded them), helping them safely and securely access appropriate data, and offering their graduate students valuable internship opportunities, UVA helped Louisiana make sense of large volumes of new classroom data, offered insights on how to encourage improvement, designed new approaches to engage families through enrollment, and helped the state produce better policy. As a result, since 2015 Louisiana has measured quality, surveyed families, and seen continual improvements in classroom quality across all settings.

Second, public education organizations can use RPPs to evaluate impact in real time, thus strengthening the effectiveness of interventions while building credibility with stakeholders. As the article "Fast-Response Research to Answer Practice and Policy Questions" shows, the Boston Public Schools (BPS) have developed a long-lived RPP with Christina Weiland (and the universities with which she's been associated) that allows BPS early childhood education leaders to quickly identify and study research problems. Whether they needed to determine whether to sustain a summer program, pursue outside accreditation for early childhood programs, or figure out how to best strengthen professional development, the BPS leaders have been able to use meaningful research on their own data rather than generic national research to guide policy making and practice. Again, it would have been much more difficult for Boston to hire staff or procure consultants to perform all of this work. And if the work were done only by employees or

consultants hired by BPS, there would be a risk that even though Boston data were used the public would dismiss the findings, believing that the experts had been "bought." Through its RPP, BPS could expediently make strategic decisions in a variety of areas, including several that were potentially controversial, while also building credibility with stakeholders by demonstrating that the decisions were based on thoughtful analysis of Boston data.

Third, RPPs allow public education organizations to engage in a more nimble and iterative design and build process. These partnerships attract top academic talent, often a set of innovative thinkers who are curious, who are willing to gather and use data differently, and who take a hands-on, iterative approach to problem solving. Hiring this type of talent full-time and creating the organizational climate that promotes experimentation (with the risk of failure) can be difficult for public education organizations. Similarly, it is harder to procure services in which specific targets aren't known because they're to be determined through the process itself. For example, in Virginia, lawmakers wanted to better understand children's level of preparedness at kindergarten entry. Yet school divisions were supplied only with a statewide literacy tool. And while policymakers, educators, and parents all acknowledged that school readiness was more than knowing ABCs and letter sounds, there was no agreed upon readiness measure. Rather than procure an existing product that would not be likely to meet all Virginia's needs, the state established an RPP between the Virginia Department of Education and the University of Virginia. As Amanda Williford and colleagues show elsewhere in this issue, this partnership allowed Virginia to design an innovative and individualized solution, the Virginia Kindergarten Readiness Program (VKRP), that built on existing tools and practices, provided a comprehensive baseline without overwhelming the system, and was gradually implemented over five years, allowing for iterations as needed to meet the needs of the field. The VKRP partnership positioned the Virginia Department of Education to support all school divisions in using the tool and its aligned supports, to collaborate with UVA to make sense of the data, and to use the results to better inform its board and lawmakers in order to shape policies that can help increase kindergarten readiness in the state.

Though RPPs can benefit public education organizations, they are not without challenges. First, as the name *research-practice partnership* indicates, the perception may arise that research is the top priority. Accordingly, public education leaders may think that an RPP is just one more thing taking attention away from their real jobs. It's imperative that education leaders understand the value of an RPP and believe that it's a strategic priority for the public education organization, not just the university. If leaders believe they're getting top talent and cutting-edge insights (and not just being asked to "participate in another research project"), they can help enable buy-in at every level.

Still, partners must get permissions, obtain secure data, ensure that decision-makers are well positioned to use findings, and coordinate implementation in the field. Successful partnerships require strong leaders on both sides who are adept at managing their internal systems so that they can fulfill all the obligations of the partnership.

Public organization leaders must also build systems of support behind the scenes so that they're prepared for negative research findings or serious implementation challenges in the field. This is especially important because partners are subject to different social and political contexts and timelines. Public education organizations typically operate on faster timelines, meaning they have to make policy or practice decisions that align with school calendars, or with political events such as the election of a new mayor or governor. Although public education leaders appreciate rigorous research, they often have to make decisions without it. Rather than let the perfect become the enemy of the good, RPPs can run more limited, rapid response research projects while candidly acknowledging the limitations of the findings.

Finally, different internal and external contexts can produce differences of opinion between partners. In Louisiana, some early childhood scholars, including our research partners, initially disagreed with aspects of the state's approach to classroom observations. Without any damage to the relationship, the partners acknowledged the disagreement, and Louisiana policy makers selected an approach that best met the needs of the state but also required measurement of results along the way. The partners then used this ongoing research to strengthen policy and practice as well as to measure the impact on children. Ultimately, Louisiana was able to establish an approach to statewide early childhood accountability that has helped drive continual improvement in classroom quality in child care, Head Start, and prekindergarten since 2015.

This issue of the *Future of Children* provides multiple insights about the opportunities and challenges of RPPs. It highlights how these relationships can benefit both public education organizations and higher education institutions. Though these partnerships take time and resources to form and sustain, require committed leadership, and necessitate acknowledgment of differences, they offer an array of benefits that outweigh the challenges. Research partners get to do meaningful work, see the impact of that work, and make a mark that transcends academia. Public education partners gain insights, incorporate timely and relevant research into their policy and practice work, and implement innovative solutions. Working together, practitioners and academics can maximize their impact, producing positive outcomes for those who matter most: children.

Building toward Effectiveness

JoAnn Hsueh, Barbara Condliffe, William Corrin, Samantha Steimle, Rekha Balu, Margaret Hennessy, Sharon Huang, Vianny Lugo-Aracena, Michelle Maier, Shira Mattera, Shay O'Brien, Amena Sengal, Jed Teres, Samantha Wulfsohn, and Jennifer Yeaton

Researchers, as well as school districts and other providers, all share the goal of ensuring that early childhood education positively affects young children's social, emotional, and academic development. Because they align practice with the best knowledge available, research-practice partnerships (RPPs) are one of the strongest hopes for bringing the varied perspectives and expertise of these players together and thus maximizing potential benefits for children.[1]

Many RPPs bring together talented university-based researchers with school districts and providers. Yet they can be challenging to sustain, and some last longer than others. RPPs, as the articles in this issue demonstrate, can vary widely in terms of their initial lines of inquiry, goals, and the kinds of researchers, practitioners, and policy makers who constitute them. Indeed, to meet the needs, interests, and circumstances of districts and providers, relationships in an RPP need to be flexible. Researchers can find it quite challenging to adjust as real-world contexts and circumstances change. An RPP in which an independent, nonpartisan, nonprofit organization like MDRC is the research partner serves as an example of one RPP model, and in this commentary, we describe our experiences and how we've been able to use our infrastructure and diverse expertise to meet the changing scope and needs of different phases of long-standing partnerships. In doing so, we aim to illuminate key principles and offer potential strategies that other researchers can adopt to establish long-standing, collaborative partnerships of their own.

Partnering with a Research Organization

What does a partner like MDRC bring to RPPs? MDRC is a mission-driven social policy research organization with a 47-

The authors are researchers at MDRC, a nonprofit, nonpartisan education and social policy research organization based in New York, NY, and Oakland, CA.

year history of developing and evaluating policies and programs that seek to improve the lives of low-income families, children, and communities. Supported by a blend of philanthropic grants and public funding, our work is informed by a deep institutional history of integrating impact and implementation research and of providing technical assistance and feedback to our partners, including school districts, government agencies, and nonprofits. We focus on supporting the use of data in decision making to strengthen program operations and implementation and to guide continuous program improvement.

Historically, MDRC has been best known for large-scale demonstrations and program evaluations that often employ random assignment. Usually, our projects involve years of research, during which we work closely with government agencies, schools, and program providers. In these partnerships, we've not only conducted research but also advised partners on their programming or provided technical assistance to them, thus both generating new knowledge and helping to apply it in practice. In fact, a significant portion of MDRC researchers have experience as program administrators, teachers, and frontline staff, meaning we have a strong understanding of how programs operate in the real world. We're also trained in multiple disciplines, including developmental and implementation science, applied behavioral sciences, and predictive analytics.

School districts and providers join RPPs because they want to be sure their programs are making a difference. As an organization known for assessing the impacts that programs achieve, MDRC focuses on understanding the value that an intervention produces compared with the business-as-usual approach. Thus we start from a philosophy of wanting to create value beyond the status quo.

To establish mutually beneficial RPPs in early childhood education, the members of each group that makes up an RPP must play a variety of research roles to supply information needed by their partners. Our work as an intermediary, in which we draw on evidence-based insights to ground and support program implementation with practice partners, has spanned multiple partnerships with public school districts, Head Start, and community-based childcare centers. One sustained partnership has been with the New York City's Department of Education (DOE), which has been eager to learn through interactive evidence-building and program improvement. Relying on both public and private funding raised collaboratively, this partnership has sought to generate insights that can guide the DOE's early childhood and early elementary school programming in a variety of ways. In concert with the DOE, a consortium of philanthropies, and several academic partners, we started this partnership by testing enhanced math instruction with aligned curricula in preschool classrooms and in supplemental small-group math enrichment clubs outside of regular classroom instruction in kindergarten. Building on lessons learned from that demonstration, we're currently working with the DOE to develop an in-classroom model of the small-group kindergarten math clubs.

As part of a group of organizations participating in the city's continuing expansion of universal prekindergarten, we and a network of researchers have partnered with the district to conduct

a series of qualitative and quantitative studies so that we can better understand the effects and implementation of new program components and supports, such as professional development, that are being rolled out for teachers. Talented academic partners also contribute to MDRC's RPP with the DOE, including Pamela Morris at New York University, Robin Jacob at the University of Michigan, Douglas Clements and Julie Sarama at the University of Denver, Katherine Baldwin at the Bank Street Education Center, and the Foundation for Child Development's New York City Early Childhood Research Network. Academic partners like these add targeted expertise to the partnership beyond our own, enabling the partnership to better respond to practitioners' changing needs.

In our work with the DOE (as with most of our partnerships), we aim to meet our partners' needs by looking both inward and outward. When we look inward, we examine a program's theory of change, the strength of the components of its model, and its implementation processes. When we look outward, we study the external systems in which a program is embedded, the organizations that support it, and the external conditions that affect it, and we compare the program with alternatives that might already be available to potential participants. As a result, the research activities are varied.

Early on, a partnership may focus on the strength of the program model, the coherence of its critical components, and the clarity of its design as represented in a theory of change and a logic model. Are the program model's critical components theoretically sound, logically connected to desired outcomes, and evidence based? The first stage of our work with partners is often refining or even developing the theory of change. In this way, we help connect the expertise of program developers with the operational realities of programs as understood and experienced by the practitioners with whom we work. To build enhanced math instruction in New York City's pre-K and kindergarten classrooms, for example, we outlined a theory of change whose goal was to foster alignment of the curricular models across grades (Building Blocks and High 5s, respectively) where none had previously existed.

In the early stages of our RPPs, it's often important to ensure that the intended program model is put into practice correctly and that the targeted participation level is achieved. How are critical components or activities of the program model implemented? Are they delivered as designed in terms of duration, intensity, and quality? Do critical components reach the targeted population? To improve the implementation process, we help programs define how they'll answer these questions. The criteria they develop make it possible to regularly measure implementation progress and program performance. For instance, in the case of the enhanced pre-K math instruction, we helped develop ways to characterize how often curricular components were delivered, the quality of teachers' delivery of the components, and the overall quality of implementation by lead and assistant teachers. These tools were shared with the DOE and served as the basis for monitoring how well the program was being implemented when the district began scaling the program as one of its professional development and curricular tracks in prekindergarten classrooms across the city.

To support program implementation, we focus on *service contrast*, which refers to how the opportunities and experiences that the program offers contrast with those that might otherwise be available to the targeted population of participants. This value-added perspective is often not well understood. If a program doesn't offer an experience that's substantially better than what's already available, it won't benefit its participants: that is, it won't have an impact. Even an informal understanding of service contrast can help identify features of a program that can be strengthened to improve outcomes for young children. For example, findings from the early phases of our partnership with the DOE suggested that high-quality learning opportunities consisting of small-group, hands-on instruction—among the practices that are most promising for supporting early learning—weren't common in kindergarten classrooms, particularly in under-resourced schools. Observations we conducted showed that math instruction in schools was primarily didactic in nature and that it was delivered almost exclusively in a large group or with students working individually at their seats. Math instruction in these classrooms lacked both the differentiation and the hands-on activities that support children's development of knowledge and skills—activities that are emphasized in the High 5s math clubs we were testing. These features of High 5s method may be a key reason for its positive impacts; when stacked on top of enhanced math instruction in pre-K, High 5s effectively closed 29 percent of the math achievement gap between low-income and higher-income children.[2]

In these partnerships, we often help programs by assessing their need to build rigorous evidence of their impacts, and their readiness to do so; then we work with them to build it. A program may need time to mature to the point that its components are sufficiently implemented before it's ready for an evaluation assessing its effectiveness. Likewise, programs may need to experiment with adaptations and innovations to assess feasibility and strengthen implementation, before it's possible to determine whether such initiatives are producing the desired effects. The current phase of our work with the DOE on math enhancements in pre-K and kindergarten illustrates such adaptation and innovation.

The kindergarten intervention, High 5s, was meant to test whether aligning enriched math instruction in kindergarten with enriched preschool math instruction could produce sustained benefits for young children. The High 5s math clubs offered small-group instruction of key mathematical learning milestones in a fun, engaging, and developmentally appropriate format. Clubs met outside instructional time for three 30-minute sessions per week, and each club typically included four children and one facilitator. Outside facilitators administered the program to more than 300 children in 24 public elementary schools during the 2015–16 school year. The original High 5s math clubs were designed as an add-on activity offered outside of the typical school day and not explicitly as a model that could be readily scaled.

Excited by the positive impacts of the combined pre-K intervention and High 5s, some superintendents and principals in the city expressed interest in strengthening the quality of math instruction in kindergarten in scalable, cost-conscious ways—including by transforming the original High 5s club format into a classroom version that could be integrated into the kindergarten daily

schedule without having to eliminate other academic or enrichment content. To this end, we adapted materials so that teachers could rotate through High 5s while the rest of the class participated in center-type, small-group activities. During the 2019–20 school year, we piloted the in-class version of High 5s in a small number of schools in New York City, and we collected feedback from teachers to assess the feasibility of implementation. In response to the COVID-19 public health crisis, we also adapted these materials to support High 5s activities for remote classroom instruction in collaboration with the district.

We also seek to help our partners build their own capacity for data-driven decision making so they can strengthen their programming in the future through the collection, monitoring, and interpretation of measures that track the delivery and implementation of critical components and practices. The DOE already has a strong data collection system in place, so we have helped it use the research data and findings to continue to identify areas for improvement of the existing pre-K to third grade pipeline.

A program's reach is a key factor in its effectiveness, particularly as the program is scaled up. For example, because the DOE recognized that some families face obstacles in applying to and enrolling in school, we explored barriers families face when they select, apply to, and enroll in kindergarten in New York City's choice-based system. We found that about 28 percent of families who enrolled in the city's kindergarten didn't participate in the formal selection and application process, and that this rate was higher among families living in temporary housing and those who didn't speak English at home.[3] Interviews and observations found that families were confused and frustrated by the multistep application process. MDRC and the DOE used these insights in our collaborative effort to design and test a set of behaviorally centered support strategies to improve families' experiences when applying for kindergarten. We are currently building on this project, extending the partnership by applying insights from behavioral science to design and test interventions for other grade levels.[4] For example, with funding from the Michael and Susan Dell Foundation, we've been helping the DOE make use of behavioral insights to support equitable access to school application and enrollment during the COVID-19 pandemic.

A Need to Be Flexible

Building long-lasting RPPs requires taking on different activities over time. As our partnership with the DOE illustrates, long-standing partnerships can build directly from the on-the-ground expertise and insights of programs and practitioners, draw on different funding sources, and engage other researchers over time. Yet the range of activities we've undertaken with the DOE is coherent. Tying these activities together and drawing on knowledge accumulated across different lines of inquiry over time can be quite powerful. By doing so, we are able to develop a set of complementary evidence-building activities that integrate impact and implementation research with program adaptation and improvement and that can be used to guide continued improvement and strengthening of programming and practice at scale. This broad range of activities is critical to realizing the full potential of mutually beneficial partnerships that align research and practice to support programming that can yield benefits for the populations served.

Long-term partnerships can face challenges, particularly in securing the time and resources to cover a broad scope of collaborative research endeavors, as well as ensuring that they have the flexibility to make midcourse corrections. We offer several observations from our experience that we hope will be useful to others.

First, a strong understanding of operational realities is important for making research relevant to practitioners. To provide useful technical assistance and support, we need on-the-ground knowledge of the challenges that our partners face in complex social and education service systems. As we offered teachers technical assistance and monitored their implementation of Building Blocks in pre-K classrooms, for example, we saw that center directors and school principals played an instrumental role in translating and providing implementation support of district-wide instructional initiatives in the classrooms. This is true in most state and local pre-K programs, but the scale of New York City's school district makes ensuring that an intended program model is implemented correctly a unique challenge. The sheer size of the district can lead to substantial variability in the messaging and climate around new initiatives, which in turn contributes to variation in teachers' buy-in and implementation. Our team helped the developers and the DOE identify opportunities for strengthening teachers' understanding of the intent of key instructional practices supported by a curricular model and of how these practices fit with other district-wide requirements. This work included making recommendations about how to reinforce teacher training and coaching as the district began to scale Building Blocks more widely. Close collaboration with our district program office partners also plays a critical role in ensuring that our work responds to operational realities. For example, in our design and evaluation of an intervention to support families in the kindergarten application process, we work closely with the DOE Office of Student Enrollment, taking advantage of our partner's on-the-ground operational insights to design interventions they will be able sustain over time and to set a learning agenda that can guide broader continuous improvement efforts.

Second, bringing new resources to RPPs helps ensure their adaptability and longevity. An essential aspect of our success in establishing long-term RPPs is our organization's diversity of expertise and infrastructure, which is designed to foster interdisciplinary collaboration and information sharing across teams. This infrastructure complements our efforts to generate resources across multiple research engagements with our partners. We're often able to braid funding streams in different applied areas of research to support the needs and priorities of our partners. In New York City, for example, our testing of aligned math curricular models in pre-K and kindergarten began as a collaboration with the anti-poverty foundation Robin Hood that was later expanded with additional funding from the Heising-Simons Foundation, the Overdeck Family Foundation, and the Richard W. Goldman Family Foundation. Funding was identified and secured after the research agenda and activities for a particular phase of the partnership were carved out. Elsewhere, in other RPPs, we've secured funding for particular types of research, such as applying predictive analytics to the problems that face social and educational policy and program systems, before a specific district and program

partner for the research had been identified. We're thus able to bring a set of funds to our partners to build their capacity and to engage in activities that align with their interests by taking advantage of related research in our organization. In these cases, we refine a set of research activities with our partners after funding is in place. Capitalizing on existing resources or integrating different philanthropic and public funding sources is only possible when we deeply understand the needs of a district or program partner. Thus MDRC's interdisciplinary structure allows us to think flexibly about how to apply a range of philanthropic and public resources to support sustained RPPs over time.

Third, balancing the complementary goals of the partners regarding research, practice, and policy is best achieved with both varied expertise and an array of services and capabilities. Our interdisciplinary areas of expertise and investments in infrastructure allow us to respond to the changing needs and priorities of our partners, and in so doing, we learn from our partners how to generate evidence for the field more broadly, particularly around scaling in real-world contexts. As more university-based researchers try to build long-standing RPPs, critical investments in infrastructure and institutional change may similarly be needed to bring together researchers with different but complementary disciplinary expertise.

Over the years, MDRC has invested in staff development to ensure that technical assistance and operational insights are integrated with implementation science, predictive analytics, and applied behavioral research so that operational challenges across complex social and education policy systems can be addressed. In our partnership with the DOE, these investments have helped us link implementation and impact research with technical assistance and feedback in a way that can guide program improvement; in this way, we can better realize the long-term potential of RPPs and support the innovations our practice partners have developed to benefit the populations they serve.

In sum, RPPs inevitably see shifts in the breadth and scope of inquiry over time as real-world contexts and circumstances change. Successfully meeting these demands requires carefully integrating a range of expertise and building an infrastructure that can evolve with the needs, interests, and capacities of districts, providers, and researchers into mutually beneficial relationships that productively support cycles of evidence building for all.

Endnotes

1. Cynthia E. Coburn, William R. Penuel, and Kimberly E. Geil, *Research-Practice Partnerships: A Strategy for Leveraging Research for Educational Improvement in School Districts* (New York: William T. Grant Foundation, 2013), https://wtgrantfoundation.org/library/uploads/2015/10/Research-Practice-Partnerships-at-the-District-Level.pdf.
2. Shira Kolnik Mattera, Robin Tepper Jacob, and Pamela Morris, *Strengthening Children's Math Skills with Enhanced Instruction The Impacts of Making Pre-K Count and High 5s on Kindergarten Outcomes* (New York: MDRC, 2018), https://www.mdrc.org/sites/default/files/MPC-High_5s_Impact_FR_0.pdf.
3. Barbara Condliffe and Rekha Balu, *Missing from the Start: Engagement in New York City's Kindergarten Application* (New York: MDRC, 2019), https://www.mdrc.org/sites/default/files/KEBS_Practitioner_Brief1.0.pdf.
4. Barbara Condliffe et al., "The Promise of a Collaborative Approach to Problem-Solving and Innovation-Testing: Reflections from a New Research-Practice Partnership in New York City," *NNERPP Extra* 2, no. 2 (2020): 14–18, https://nnerppextra.rice.edu/the-promise-of-a-collaborative-approach.

The views expressed in this publication are those of the authors and do not necessarily represent the views of the Princeton School of Public and International Affairs or the Brookings Institution.

To download a free electronic copy or sign up for our e-newsletter, go to our website, www.futureofchildren.org. If you would like additional information about the journal, please send questions to foc@princeton.edu.

www.futureofchildren.org

ISBN 978-0-9814705-2-8
90000>
9 780981 470528