RESEARCH REPORT

Final Report on the Hawaiʻi P-3 Evaluation

Gail L. Zellman and M. Rebecca Kilburn

RAND
CORPORATION

For more information on this publication, visit www.rand.org/t/rr1100

Support RAND
Make a tax-deductible charitable contribution at
www.rand.org/giving/contribute

www.rand.org

Preface

Since the beginning of this century, efforts to improve educational outcomes have increasingly focused "upstream" on improving the readiness of students entering kindergarten as well as increasing coordination and alignment of curriculum, assessments, and expectations between preschool providers and the kindergarten through 12th grade system. The growing "preschool to third grade," or "P–3," movement reflects this trend. This document reports the findings from RAND's five-year evaluation of the Hawai'i P–3 initiative, which began in 2009 and formally ended in 2014, although some activities continue at the time of this writing in May 2015 under a no-cost extension from the funder. These findings include both qualitative and quantitative assessments of the initiative's implementation and outcomes.

This report should be of interest to individuals who work in the P–3 area as well as those who work in the areas of school readiness, early education, elementary education, and school reform. Findings from the first year of the evaluation are reported in Zellman and Kilburn (2011), which is available for free on the RAND website.

The Hawai'i P–20 Partnerships for Education, housed at the University of Hawai'i, was the P–3 grantee and the sponsor of this research; the state's P–3 project was supported by a grant from the W. K. Kellogg Foundation. This research was conducted jointly in RAND Education and RAND Labor and Population, two units of the RAND Corporation. For inquiries related to RAND Education, please contact Darleen Opfer, director, RAND Education, at Darleen_Opfer@rand.org. For inquiries related to RAND Labor and Population, please contact Krishna Kumar, director, RAND Labor and Population, Krishna_Kumar@rand.org.

Contents

Figures

Tables

Executive Summary

Background

In response to a growing consensus that the U.S. education system needs to find ways to improve student achievement, a range of education reforms have emerged over the past decade. One of these is the "preschool to third grade," or "P–3," education reform movement (Graves, 2006; Rice and McLaughlin, 2007; Takanishi and Kauerz, 2008; Kaurez and Coffman, 2013). Proponents of the P–3 approach point out that many students enter kindergarten lacking basic language, social, and pre-mathematics skills needed to succeed in school. Deficits in these skills are apparent when children enter kindergarten, and children who start school with deficits in these skills tend to carry these deficits forward (Cannon and Karoly, 2007). Further bolstering the case for the P–3 approach are research and essays showing the importance of decreasing these skill deficits by third grade and that third-grade reading skills are a predictor of high school graduation and other life outcomes (Fiester, 2010; Hernandez, 2011).

P–3 education reform efforts focus on ways to improve the readiness of students entering the K–12 system and also encourage prekindergarten (PreK) and K–12 systems to better integrate their efforts to promote student learning. The logic behind this movement is that P–3 activities will better prepare students for kindergarten, which will promote third-grade performance and lead to better educational and career outcomes. There is some variation among the existing P–3 initiatives, but they typically include these components:

- Greater access to PreK for three- and four-year-olds
- Support for transitioning from PreK to kindergarten
- Alignment of curriculum, standards, and assessment from PreK through third grade
- Training for teachers of PreK to third grade that focuses on child development and child-centered learning approaches
- Instructional practices that support an individual child's learning and social and emotional development
- Parent engagement in learning from PreK through third grade
- Use of data for quality improvement and accountability

- Emphasis on the "whole child" concept of learning (Kauerz, 2008).

In 2007, with support from the W. K. Kellogg Foundation (WKKF), the Hawai'i P–20 Partnerships for Education (P–20) launched the state's P–3 initiative. The Hawai'i P–3 initiative's goal was for every child in Hawai'i to read at grade level by third grade.

The initiative included state-level work on broad policy and data issues and support for five local demonstration sites. The P–3 demonstration site *Request for Applications* asked potential sites to describe their planned work in terms of a framework defined by seven focus areas selected by P–20; these focus areas were supported by available research evidence pointing to their importance in furthering the P–3 goal. One focus area, Data, was later dropped from demonstration site requirements; P–20 assumed this focus area when it recognized that P–20 had more capacity to address this area than the demonstration sites. The six remaining focus areas included:

1. Leadership for literacy
2. Standards, curriculum, and assessment
3. Instruction
4. Teacher professional development
5. Comprehensive early learning services and access to services for children from birth to age five
6. Family–school transitions and partnerships.

P–20 required that applications include commitments from and the signatures of PreK and K–12 institutions in the demonstration sites. This set in place expectations for cooperation between PreK providers and Hawai'i Department of Education (HIDOE) personnel and resulted in substantial cooperation over the life of the initiative.

Objectives and Approach

The RAND Corporation was asked to evaluate Hawai'i's P–3 initiative, which began in 2009 and formally ended in 2014. This report presents our findings of that five-year evaluation, although some state-level initiative activities continue at the time of this writing in May 2015 under a no-cost extension from the funder. These findings were derived from both qualitative and quantitative evaluations of the initiative's implementation and outcomes.

The primary goals of RAND's *qualitative evaluation* were to:

- Identify the strategies that P–20 and the demonstration sites used and assess their alignment with best practices in P–3.
- Analyze the degree to which the five demonstration sites and P–20 executed their strategies and plans described in their logic models.
- Assess if P–20 and the demonstration sites developed plans likely to promote long-term effects or the sustainability of their P–3 efforts.
- Determine if P–20 and the demonstration sites engaged in activities that promote system change and assess how much system change occurred.

To achieve these goals, we collected information about the strategies that P–20 and the demonstration sites included in their logic models, which defined their work. We compared these strategies to what were considered best practices in P–3 at the time. We assessed the implementation of the plans made by the demonstration sites and P–20 and examined the likelihood that the work at the sites and the state would have long-term effects. Finally, we examined the work of P–20 and the sites from a system change perspective, described in Chapter 3.

The *quantitative evaluation* included two sets of analyses. The first one assessed responses by elementary school principals in the five demonstration sites to three waves of online School Information Forms (SIFs)—2011, 2013, and 2014—to collect information on the extent of principals' awareness of the sites' P–3 work and their attitudes toward it and whether the schools were implementing P–3 activities.

The second quantitative analysis assessed third-grade reading scores, comparing the scores of students in the five demonstration sites with the scores of students who were not exposed to P–3. The test-score data were drawn from the Hawai'i State Assessment (HSA) reading assessment. This analysis addressed the ultimate goal of the P–3 initiative: all students reading at grade level in third grade. Specifically, it assessed whether students exposed to P–3 in the five demonstration sites had a greater likelihood of reading at grade level in grade 3 than students whose complex area did not participate in P–3.[1]

---

[1]Hawai'i has just one, statewide school district, which is overseen by the Hawai'i Department of Education (HIDOE). Complex areas are roughly similar to school districts, although policy is promulgated at the state level. Complex areas are administered by complex area superintendents (CASs).

Evaluation Findings

Did P–20 and the demonstration sites develop plans that aligned with best practices in P–3 at the time?

When the Hawai'i P–3 initiative began in 2009, several features were commonly understood to be at the core of a P–3 approach. In general, the framework that P–20 developed to guide its own work and the work of the demonstration sites captured these features. The demonstration sites were required to identify and implement activities in each of the six focus areas that represented the components of the initiative's P–3 approach. The objectives-based contracts between P–20 and the demonstration sites specified the long-term outcome to be achieved: improved third-grade reading scores. Activities to achieve that goal were not specified by P–20 but were to be suggested by the sites. P–20's requirement that the demonstration sites engage in activities in all six focus areas appears to have played a large role in the demonstration sites' work covering all aspects of best practices in P–3 rather than selected components of P–3.

Did P–20 and the demonstration sites execute their plans as intended?

P–20 and the demonstration sites executed the majority of the activities outlined in their plans; the exceptions tended to be activities modified or added to plans to respond to unforeseen opportunities or changes in the P–3 landscape. An example is the unanticipated HIDOE adoption of the Danielson assessment of instructional quality, which reduced the necessity for the P–3 initiative to promote the use of the alternative Classroom Assessment Scoring System (CLASS) assessment in K–3 classrooms. Participants in the P–3 initiative were not able to anticipate all the changes that might be encountered during its tenure, and sites balanced their commitments to stated plans with adaptations in the face of new opportunities and implementation challenges.

Did P–20 and the demonstration sites develop plans likely to promote long-term effects or sustainability of their P–3 efforts?

Much of the statewide and demonstration site work is likely to contribute to long-term change in Hawai'i's P–3 landscape. It has already increased the appreciation for the importance of early learning. Many stakeholders attribute at least some of this change to the work of the demonstration sites and to P–20's efforts. The initiative also

produced changes in knowledge and skills of many individuals engaged in P–3 services in the state, and many of these individuals will continue to work in the P–3 realm long after the initiative has formally ended. Also likely to endure are some of the state-level products of the initiative—for example, Hawaiʻi Early Learning and Development Standards (HELDS), the PK-3 Graduate Certificate program, and a number of Professional Development Experiences that Educate and Empower (PDE3) courses that focus on early childhood topics that HIDOE staff can now complete for salary credits.

Although the activities undertaken during the project period will have long-term effects, few specific initiative activities will continue past the end of the WKKF grant. There are instances where HIDOE or individual complex area superintendents have decided to continue individual elements of P–3 activities, but few activities have been institutionalized. At the same time, there has been no systematic effort to plan for a next phase of the P–3 initiative or identify funding to continue the activities. P–20 and the sites engaged in little strategic communication about their accomplishments; thus, stakeholders often attribute the P–3 initiative with having an important role in some of the changes that occurred during the project period but are vague on the specific achievements of the P–3 grant per se.

Did P–20 and the demonstration sites engage in activities that promoted system change?

Both P–20 and the demonstration sites implemented most of the elements of system change that the framework we applied (described in Chapter 3) posits to be necessary to effectively promote it. The requirement that HIDOE and PreK entities jointly commit to the demonstration site work created an expectation of cooperation even before the sites were funded. The expansion of P–20 expectations for the demonstration sites (i.e., that the sites engage in activities representing all six focus areas) helped to broaden the scope of demonstration site efforts. P–20 gave the sites considerable responsibility to define and organize the work, because site contracts were outcome-based rather than activity-based. Within sites, responsibilities were negotiated among team members. P–20 did not offer incentives or impose consequences for demonstration site performance; thus, incentives played little role in motivating work beyond paying for substitute teachers in some sites. Still, stakeholders noted that the availability of funds and time that P–3 afforded made an enormous difference. Some activities, such as the HELDS, would not have been possible without the P–3 initiative. Time was critical as well: As

many stakeholders noted, people's attitudes need time to change, and cross-entity cooperation efforts required the often slow building of trust through shared work. P–20 engaged in limited monitoring of the demonstration sites at first, but monitoring increased over time with P–20 staff restructuring. The P–3 work undertaken by P–20 was monitored lightly by funders and the three organizations that formed the P–20 partnership.

At the state level, P–20 engaged in a range of work that promoted alignment of PreK and K–12, such as the longitudinal database work, and a range of efforts that increased system capacity, such as the PK-3 Graduate Certificate program and support for the PDE3 courses and the establishment of the Executive Office of Early Learning (EOEL). HIDOE's recent involvement in delivering PreK to four-year-olds outside of early intervention preschools has brought it more actively into the early learning space, which might also increase alignment.

Did the P–3 initiative increase the fraction of third-grade children reading at grade level?

The ultimate goal of Hawai'i's P–3 initiative was to have all children reading at grade level by third grade. We were able to analyze reading-score data to assess how much the initiative reached this ambitious goal. Despite some study limitations discussed in Chapter 3 that would tend to reduce the likelihood of finding measurable effects from the demonstration site P–3 work, we found evidence that more years of participating in the P–3 initiative was associated with a modest but significant increase in student reading scores and increased the likelihood of scoring proficient on the state reading test. The results also show a reduction in the initial gap in reading scores between the demonstration site schools, which were chosen because of their low student performance levels, and other schools. Specifically, we found that being in a school that was exposed to the P–3 initiative for five years was associated with a 3.5-point increase in students' HSA reading score, which is an effect size of about 0.1. This is comparable to other estimates of the effects of nine additional weeks of schooling (Chingos, Whitehurst, and Gallaher, 2013) and is higher than an estimate of the average effect size for elementary school interventions for mainstream students (Hill et al., 2008). Although the P–3 demonstration sites undertook activities during this period other than P–3 work, which might have contributed to these outcomes, these findings

show encouraging evidence that the reading-score-test gap between demonstration site schools and other schools narrowed by the end of the P–3 initiative.

Implications for P–3 Initiatives and Evaluations

The following are implications for future P–3 initiatives that we derived as we tracked the implementation of the initiative, assessed its progress, and documented its accomplishments.

Determine in Advance an Appropriate Balance Between Standardization and Site-Specific Needs and Resources

It is widely understood that when working with local communities, one-size-fits-all approaches are less effective than approaches that are targeted toward specific community needs and that take advantage of community assets (Kilburn and Maloney, 2010). But it is difficult to determine the right balance of local autonomy and the standardization required to maximize the initiative's benefits when presented with widely varying sites in a multisite initiative. In Hawai'i, P–20 allowed individual sites to determine which activities they would pursue in each focus area and to specify appropriate measurable outcomes for each. This allowed more diversity in strategies and activities. This diversity was permissible because all demonstrations sites would ultimately be measured by whether they had achieved the same outcome—children reading at grade level by third grade.

Consider Contracts That Specify Outcomes Rather Than Activities

P–20 specified the long-term outcome to be achieved—improved third-grade reading scores. It did not specify the activities that sites would implement in each of the six focus areas of the initiative. Public administration has increasingly recognized the value of such an outcomes-based approach to service delivery, whereas education and social service contracts have generally continued to rely on activities-based contracts. This initiative suggests that multisite initiatives can harness the advantages of outcomes-based contracting, which include local decisionmaking processes and activities tailored to local needs and resources and central monitoring of activities and progress.

Establish Measurable Outcomes for the Work

It is important to identify measurable outcomes for all the activities to be undertaken. This way, monitoring is possible and changes can be made early on if the work is not producing expected outcomes. Standardized outcomes represent a desirable feature for multisite work because they enable cross-site comparisons and provide conceptually simple benchmarks. One of the strongest features of the P–3 initiative assessed here was that all stakeholders were able to clearly state the overarching objective and agree on its importance. This feature led to extremely strong buy-in for the initiative.

Explicitly Plan for Changes in Policy and Personnel Turnover

An initiative like P–3 required collaboration from a variety of institutions, many of which would alter policies during the initiative. Having the flexibility to adapt in response to policy changes rather than adhere to ex ante plans allows an initiative to take advantage of opportunity and avoid continuing activities that might no longer be needed. However, this flexibility must be balanced against the risk that permitting change might contribute to avoidance of accountability through frequently changing plans. Personnel turnover was also an important but uncontrollable challenge in several of the demonstration sites.

Consider Sustainability from Inception

Now that demonstration site funding has ended, some demonstration sites will continue with aspects of the P–3 work with funding from other grants, but most of the work will cease. This was a source of disappointment to several of the demonstration sites. However, it was not clear that P–20 emphasized planning or efforts related to sustainability. One way to encourage and support sustainability is to publicly share a project's accomplishments with policymakers and potential funders. At the end of the project, there did not appear to be an effort to "tell the story" of the accomplishments of the P–3 initiative. While many stakeholders have a positive view of the initiative and believe it was an important contributor to the improvement of early childhood education and early elementary education, they are often not able to articulate specific accomplishments attributable to P–3.

Require Explicit Agreements Between Parties Required for Collaborative Work

One area of uniform agreement among all stakeholders participating in interviews was the value inherent in requiring HIDOE and other partners to be part of the proposal and award-agreement processes, to make their commitments explicit, and to ensure that any commitments made come from a high-level authority within the organizations. Stakeholders agree that one of the hallmarks of this initiative was that it contributed to unprecedentedly strong coalitions between the early education and elementary education sectors.

Acknowledgments

Abbreviations

AYP	adequate yearly progress
CAS	complex area superintendent
CLASS	Classroom Assessment Scoring System
CPC	Chicago Child-Parent Centers
DAP	developmentally appropriate practice
ECE	Early Childhood Education
EOEL	Executive Office on Early Learning
FFT	Framework for Teaching
FRPL	free or reduced-price lunch
GBA	Good Beginnings Alliance
GOLD®	*Teaching Strategies GOLD®*
HAEYC	Hawaiʻi Association for Education of Young Children
HCPS	Hawaiʻi Content and Performance Standards
HELDS	Hawaiʻi Early Learning and Development Standards
HIDOE	Hawaiʻi Department of Education
HIECC	Hawaiʻi Island Early Childhood Conference
HSA	Hawaiʻi State Assessment
HSSRA	Hawaiʻi State School Readiness Assessment
INPEACE	Institute for Native Pacific Education and Culture
K–3	kindergarten through third grade
K–12	kindergarten through 12th grade
KEA	kindergarten entry assessment
KKPCA	Kaʻu-Keaʻau-Pahoa Complex Area
KRT	Kindergarten Readiness Test
NAEP	National Assessment of Educational Progress
N-W	Nānākuli-Waiʻanae
OECD	Organisation for Economic Co-operation and Development
P–3	preschool through third grade (initiative)
P–20	preschool through higher education
PCNC	Parent Community Networking Center

PDE3	Professional Development Experiences that Educate and Empower
PLC	Professional Learning Community
PPVT	Peabody Picture Vocabulary Test
PreK	prekindergarten
PreK–3	prekindergarten through third grade
RFA	request for application
RFP	request for proposal
RTTT	Race to the Top
RWJF	Robert Wood Johnson Foundation
SIF	School Information Form
SPARK	Supporting Partnerships to Assure Ready Kids™
STS	Scholastic Testing Service
WKKF	W. K. Kellogg Foundation

1. Introduction

A range of statistics supports the idea that our nation's public schools are failing to prepare students for success in work and life. Reports of the National Assessment of Educational Progress (NAEP) repeatedly indicate that the majority of students are not proficient in core subjects (NAEP, 2011), and reports of the Organisation for Economic Co-operation and Development (OECD) document that the United States is no longer one of the top 20 countries in terms of high school graduation rates and mathematics and science performance (OECD, 2014). In response to a growing consensus that the U.S. education system needs to find ways to reverse these trends, a number of education reform movements have emerged over the past decade. One of these is the "preschool to third grade," or "P–3," education reform movement (Graves, 2006; Rice and McLaughlin, 2007; Takanishi and Kauerz, 2008; Kaurez and Coffman, 2013). Proponents of the P–3 approach point out that many students enter kindergarten lacking basic skills necessary to succeed in school. Deficits in language, social, and pre-mathematics skills are apparent when children enter kindergarten, and children who start school behind tend to stay behind (Cannon and Karoly, 2007). Further bolstering the case for the P–3 approach are research and essays that show that third-grade reading skills are a predictor of high school graduation and other life outcomes (Fiester, 2010; Hernandez, 2011).

To address these challenges, P–3 education reform efforts focus on ways to improve the readiness of students entering the K–12 system and to encourage prekindergarten (PreK) and K–12 systems to better integrate their efforts to promote student learning. This logic is based on the premise that P–3 activities will better prepare students for kindergarten, which will in turn improve third-grade performance and lead to better long-term educational and career outcomes. There is some variation among existing P–3 initiatives, but they typically include these components:

- Greater access to PreK for three- and four-year-olds
- Support for transitioning from PreK to kindergarten
- Alignment of curriculum, standards, and assessment from PreK through third grade

- Training for PreK-3 teachers that focuses on child development and child-centered learning approaches
- Instructional practices that support an individual child's learning and social and emotional development
- Parent engagement in learning from PreK through third grade
- Use of data for quality improvement and accountability
- Emphasis on the "whole child" concept of learning (Kauerz, 2008).

In 2007, with support from the W. K. Kellogg Foundation (WKKF), the Hawai'i P–20 Partnerships for Education (P–20) launched the state's P–3 initiative. P–20's goal is for 55 percent of Hawai'i's working-age adults to have a two- or four-year college degree by the year 2025. The primary strategies that Hawai'i P–20 employs to achieve this goal include having all children reading at grade level by third grade; strengthening the rigor of the high school curriculum; increasing student access and success in college; and facilitating program and policy development based on research and data (Hawai'i P–20 Partnerships for Education, 2014).

Hawai'i P–20 is a collaborative that is jointly led by the primary state leaders in public early education, K–12 education, and higher education, respectively, the Executive Office of Early Learning (EOEL), Hawai'i Department of Education (HIDOE), and the University of Hawai'i system. The P–3 initiative is the primary early childhood component of P–20's work. The goal of the Hawai'i P–3 initiative is for every child in Hawai'i to read at grade level by third grade.

Hawai'i P–20 Sought a Comprehensive Evaluation of Its P–3 Initiative

Evaluations of education reform initiatives often recognize the value of assessing both the implementation and outcomes of these initiatives (O'Donnell, 2008). The former focuses on what activities the initiative proposed and whether these activities were carried out as intended. The latter assesses whether the initiative achieved its intended effects. P–20 requested that this evaluation of its P–3 initiative provide a comprehensive assessment of implementation, including system change and organizational and student outcomes. Furthermore, P–20 requested that the evaluation include not only the initiative activities executed by P–20, but also the P–3 activities that a small number of demonstration sites undertook with funding from P–20.

Research Questions

This evaluation examined Hawaiʻi P–3 activities over the past five years, which includes implementation stages ranging from the program planning stage to initial implementation to mature implementation (Chen, 2005). Reflecting the fact that the evaluation covered the entire life cycle of the initiative, our evaluation incorporates research questions that span planning to full implementation:

- Planning: Which strategies did the P–3 initiative employ, and did these align with best practices in P–3 at the time?
- Plan implementation: Did P–20 and the sites implement the initiative according to their plans?
- System change: Did P–20 and the sites engage in the activities needed to successfully accomplish system change?
- Full implementation: Did P–20 and the demonstration sites develop plans likely to promote long-term effects or sustainability of their P–3 efforts?

One of the notable features of Hawaiʻi's P–3 initiative was the clear focus on one ultimate outcome: having all children reading at grade level by third grade. Reflecting this, the research question for the outcome evaluation is:

- Were third-grade reading scores higher for children in schools that were part of P–3 demonstration sites?

Multilevel Analysis and Multiple Data Sources

P–20 designed the Hawaiʻi P–3 initiative to include activities at two levels: the state and the five local demonstration sites. Hence, this evaluation examines plans, plan implementation, and system change at both these levels. Data utilized for the evaluation capture state-level and site-level information and derive from the following sources:

- Documents
- Individual or group interviews
- School information form completed by elementary principals of schools in the five demonstration sites
- Administrative data from HIDOE reporting students' third-grade reading scores and other background information.

Chapter 2 presents P–3 theory and previous research and describes the Hawai'i P–3 initiative in detail. Chapter 3 presents the methods we used. Chapter 4 describes findings at the site level, and Chapter 5 provides state-level findings. Chapter 6 concludes the report and outlines implications of this evaluation for future P–3 initiatives and evaluations.

2. P–3: Theory, Research, and Hawaiʻi's Initiative

The P–3 Approach

References to the "P–3" approach began appearing only a handful of years before P–20 received its grant from the WKKF to implement the P–3 initiative described here. The earliest references to P–3 appear around 2003, and many of the early documents that lay out the vision for the P–3 approach were supported by the Foundation for Child Development (Bogard, 2003; Bogard and Takanishi, 2005). By this time, most states had begun to initiate efforts to increase PreK accessibility and quality; in 2004, 38 states were devoting some funding to PreK education (Bogard and Takanishi, 2005). These efforts were bolstered by evidence that children who entered kindergarten better prepared were better able to succeed in elementary school (Karoly et al., 1998).

The P–3 approach built on the PreK momentum, arguing that a seamless integrated system of early education serving children from PreK through third grade would maintain gains from early education. Proponents of P–3 pointed toward well-known evidence that effects of early interventions often "fade out" over time and that applying the strategies of effective early education through third grade as part of an aligned system would help reduce fadeout (Bogard and Takanishi, 2005). The pillars of P–3 advocated by early proponents were primarily alignment and coordination of PreK through third-grade (PreK–3) curricula, quality learning opportunities across PreK–3, and universal access to early learning opportunities (Bogard and Takanishi, 2005; Bogard, 2003). Later descriptions of the P–3 approach elaborated on these pillars by defining quality both as instruction by well-qualified and trained teachers and as learning environments that include child-centered instruction and the use of assessments to improve instruction (Graves, 2006; Reynolds, Magnuson, and Ou, 2006). A second wave of P–3 position papers added parent involvement and accountability elements. More recent elements of P–3 frameworks (Kauerz and Coffman, 2013) include such concepts as cultural inclusiveness and data-driven improvement. However, these later additions are often considered components of good education practice more generally and are not necessarily hallmarks of the P–3 approach (What

Works Clearinghouse, 2015). For a current overview of P–3 components, see Kauerz and Coffman (2013) or the Foundation for Child Development (undated).

Hawai'i P–20 developed its P–3 proposal and initiative work plan in 2006 at a time when the theory of P–3 had been articulated and disseminated, but the strategies that were central to a P–3 initiative were still being debated and refined. Several early-learning initiatives were under way, but none adhered closely to the P–3 model that was gaining consensus. For example, Florida's approach, which included community supports, was broader than P–3 (Golan et al., 2008). Other initiatives focused on selected P–3 components, for example, PreK access and teacher certification in New Jersey (Rice and McLaughlin, 2007). Washington state efforts emphasized building collaboration among key P–3 players, including providers of private and public PreK and kindergarten through third grade (K–3) ("Starting Strong in Washington State," undated).

The closest effort to a P–3 initiative at the time that the Hawai'i P–3 project commenced in 2007 came from the Chicago Child-Parent Centers (CPC), which had begun in 1967. The CPC included P–3 practices such as PreK–elementary school continuity, early-learning quality, and parent involvement. A major CPC strategy involved colocating preschools at elementary schools. Consequently, the CPC has often been characterized as an "extended early intervention," rather than an education reform strategy like P–3 (Reynolds, Magnuson, and Ou, 2010). The Hawai'i P–3 initiative was one of the first initiatives that specifically drew on contemporary guidance concerning the critical components of P–3. As such, it provides important insights about the implementation of P–3 efforts. It also provides data about the student outcomes of a comprehensive P–3 initiative.

Previous P–3 Research

Early advocates of the P–3 approach cited its value in terms of being grounded in research. Although research was lacking on the value of the entire bundle of P–3 components implemented together, research did support the developmental importance of ages three through eight and the value of individual components of the P–3 approach. When Hawai'i P–20 was crafting its proposal for the P–3 initiative, research evidence supporting the foundational importance of ages three to eight for educational success and the potential for high-quality PreK to improve educational

outcomes for at-risk students was well-known and well-documented (for reviews of this literature, see Karoly, Kilburn, and Cannon, 2005; and Bogard and Takanishi, 2005). Advocates for a P–3 approach argued in research reviews that if the individual P–3 components improved children's education outcomes, then a bundle of these components would perform as well or even better (Bogard and Takanishi, 2005; Reynolds, Magnuson, and Ou, 2006).

Longitudinal studies of what were referred to as "extended early childhood programs," which incorporated many P–3 components, were also cited as support for the P–3 approach even though only a few studies demonstrated lasting effects on student performance (Bogard, 2003; Graves, 2006; Bogard and Takanishi, 2005; Reynolds et al., 2010). The strongest results came from studies of the CPC, which found that participation was associated with higher academic achievement, greater likelihood of graduating from high school, and lower rates of grade retention and special education placement (Fuerst and Fuerst, 1993; Reynolds, 1995; Reynolds et al., 2002). Furthermore, students who participated in more years of CPC beyond PreK realized even greater gains for each additional year of elementary school services, and unlike many studies of early interventions, the benefits from CPC did not fade out; they persisted to age 15 (Reynolds, Magnuson, and Ou, 2006). Similarly, studies of the Carolina Abecedarian Project, which provided enriched early childhood education for children starting at four months of age, found that students who also got three years of additional elementary school services had the most improved academic performance (Campbell and Ramey, 1995).

The Foundation for Child Development, one of the leading proponents of the P–3 approach, has created a map of P–3 efforts occurring throughout the United States (Foundation for Child Development, undated). This map indicates that 16 states have P–3-related activities under way. Most of these P–3 efforts include some P–3 components, but do not include the comprehensive set of P–3 components found in the Hawai'i P–3 initiative. Some P–3 components, such as Oklahoma's universal PreK program (Gormley and Phillips, 2005), have been evaluated rigorously, but no one has undertaken rigorous implementation or outcome evaluations of P–3 initiatives as a whole. Some P–3 efforts have conducted process evaluations and documented lessons learned. Examples of these include a report that describes school districts' implementation of P–3 components in their communities ("Starting Strong in

Washington State," undated), and a report that describes P–3 efforts in New Jersey and identifies lessons for other states that aspire to develop P–3 systems (Mead, 2009). Our evaluation of the Hawai'i P–3 initiative is one of the first studies of a comprehensive P–3 initiative that includes evaluation of both implementation and outcomes.

Background on the Hawai'i P–3 Initiative

Hawai'i's P–3 initiative represented the primary early childhood component of a state partnership called Hawai'i P–20 Partnerships for Education, or P–20. This partnership, which initially included the University of Hawai'i system, HIDOE, and the Good Beginnings Alliance (GBA) (as the early childhood partner),[1] works to strengthen the educational pipeline from preschool through higher education (P–20) so that Hawai'i's citizens achieve college and career success. The primary goal of Hawai'i P–20 is for 55 percent of Hawai'i's working-age adults to have a two- or four-year college degree by the year 2025, and the main strategies Hawai'i P–20 employs to promote this goal are:[2]

- Having all children reading at grade level by third grade
- Strengthening the rigor of the high school curriculum
- Increasing student access and success in college
- Facilitating program and policy development based on research and data.

The P–20 partnership supports a diverse set of activities, each funded by a different outside funder, designed to promote student performance at different grade levels as well as overall capacity building.[3]

The P–20 partnership is not part of state government and has no authority to mandate cooperation among the actors such as elementary schools, preschools, or teachers who are essential to achieving P–20 goals. In the absence of authority, P–20 works to engage key actors through a process of consensus building and community collaboration and relies on relationships and nonbinding agreements to maintain

---

[1] GBA was replaced in the partnership by EOEL when GBA turned its efforts toward advocacy work.

[2] See Hawai'i P–20 Partnerships for Education, undated.

[3] Examples of projects oriented toward older students have included development of common core standards for K–12, with an emphasis on quantitative literacy; an effort designed to encourage low-income students to go to college; and a project to encourage eighth- and ninth-grade students to seek a Board of Education Recognition Diploma, which is an honors high school diploma.

involvement and cooperation. Indeed, the only exception to this lack of authority might be found in P–20's relationship with the demonstration sites. P–20 was the demonstration site funder; it issued contracts that specified that the demonstration sites were to execute particular tasks in exchange for funding. We return to this issue later.

In 2007, the WKKF awarded P–20 a grant to support its P–3 initiative, then called "Capturing the Momentum." The grant funding eventually totaled $11.5 million. The P–20 executive director was responsible for overseeing the P–3 initiative, which was the primary early childhood component of P–20. Staff who work on the statewide P–3 initiative are P–20 staff. The University of Hawaiʻi is the official grantee for the Kellogg grant that supports the P–3 initiative, and all P–20 staff are employees of the University. This report focuses on P–20's work related to the P–3 initiative funded by WKKF as well as on related PreK–3 activities that P–20 undertook during the five years covered by this evaluation.

When P–20 applied for funds from the WKKF, the state had limited infrastructure for state-supported early childhood education. Most of the state's efforts on behalf of young children's school readiness focused on the provision of child care oriented toward supporting parental employment. This was overseen by the Office of Human Services and, given its funding and goals, focused largely on regulations and subsidy rules.

At that time, leadership for early childhood education came from the private sector. The GBA was an active early childhood advocacy group that was working for increased government support for early childhood education, including the establishment of a department of early learning. Although P–20 was not an advocacy group, WKKF support gave P–20 credibility and enabled it to bring advocates to the table and support a range of activities that were viewed as being important contributors to a more robust early childhood infrastructure and greater investments, as discussed below.

Like many projects that run for a substantial period, the Hawaiʻi P–3 initiative evolved in terms of its goals and strategies; a timeline for the P–3 initiative is presented in Figure 2.1. In 2007, the first year of the WKKF grant for the P–3 work, the focus was on building connections between PreK providers and HIDOE staff, particularly

complex area superintendents (CASs)[4], elementary school principals, and teachers in grades K–3. Such connections did not exist then in most communities. Small grants of approximately $25,000 were awarded to 17 different communities across Hawai'i to bring together the key players who would form a local P–3 collaborative: PreK teachers, kindergarten and grade 1 teachers, elementary school principals, and other stakeholders, including parents. These efforts, according to those involved, were important and groundbreaking. They introduced people to others in their community who had been working, in some cases for many years, on similar issues. Sitting down together, they clarified their similar goals and set in place a mechanism for shared work. P–20 staff also engaged in some activities early on to enhance the capacity of the state in the P–3 area. For example, they hosted conferences where attendees could learn about key components of P–3 models, such as the importance of school readiness for elementary school success.

---

[4] Hawai'i has just one school district, which includes all schools in the state. Complex areas are similar to school districts and are overseen by CASs.

Figure 2.1—Timeline of Hawaiʻi P–3 Initiative

State awarded Race to the Top grant, which includes early childhood components in two communities—one already a P–3 demonstration site, and the other to become the 5th demonstration site.

2007	2008	2009	2010	2011	2012	2013	2014	2015

P–3 initiative funding from Kellogg Foundation commences.

17 sites around state funded to form community-based P–3 collaboratives. Statewide activities, such as conferences.

Decision to focus work in fewer demonstration sites. RFA issued and first two sites selected.

Farrington and N-W demonstration sites begin their P–3 work.

Two additional demonstration sites selected.

Windward and Honokaʻa demonstration sites begin work.

KKPCA begins P–3 project.

RAND continues to monitor implementation in P–3 sites for remainder of P–3 grant.

RAND collects and analyzes student outcome data.

Original end of P–3 grant from Kellogg Foundation.

Demonstration site funding ends.

P–20 awarded three-year no-cost extension for the state-level work.

RAND completes P–3 evaluation

However, by the end of the second year of the initiative, P–20 staff began to worry that the P–3 work in the 17 communities was not gaining sufficient traction and might not be able to produce significant observable effects on student outcomes by the end of the WKKF grant period. This early work (prior to the commencement of this evaluation) reportedly produced few long-term effects, in part because the projects focused heavily on building connections and left little time for implementing activities.

While there was agreement that promoting collaboration was valuable, a growing consensus among P–20 leadership suggested that more intensive work in fewer sites might "move the needle" more effectively toward the ultimate P–3 goal—improved third-grade reading scores. Fewer sites would enable P–20 to focus its resources on a small number of local demonstration sites, where locally developed ideas for reaching the ultimate goal could be attempted with enough funds to have a measurable impact.

11

Ideally, the successful strategies from the demonstration sites could be sustained, scaled up, and disseminated.

In 2008, P–20 imposed some standardization on the P–3 work across all the demonstration sites. This refocused the initiative from being entirely community-driven in terms of goals and activities to balancing community needs with some standardization that was driven by empirically based best practices in the field. To bring about this change, P–20 developed a set of seven focus areas that P–20 leadership believed could improve third-grade reading scores (one focus area, Data, was assumed by P–20). Local communities would continue to be able to design and implement activities that matched local strengths and addressed local needs, but these activities would have to fit within the six focus areas devised by P–20. The sites would be expected to implement activities in all six focus areas. The smaller number of funded sites would be subject to far more oversight of their plans, activities, outputs, and outcomes.

These ideas led to a decision to produce a request for application (RFA) for the continuing P–3 work in 2009. The RFA model was objectives-based rather than activities-based; interested sites were asked to explain how they would go about implementing the basic concepts and approaches established by the P–20 staff to reach the goal of improved reading at grade 3. These approaches and activities are described in more detail below. The P–20 partnership would issue contracts to the successful demonstration sites to implement their plans and activities as described in their RFA. It was also in 2009 that the RAND Corporation was selected through a competitive process to undertake the evaluation of the Hawai'i P–3 initiative.

Two sites were selected to receive funds under the 2009 RFA—Farrington and Nānākuli-Wai'anae (N-W), both of which had received funding in the previous small-grants phase of the P–3 initiative. Each site was awarded $1 million over five years. The two winning sites understood that they would be evaluated and expected to work with the outside evaluator chosen to design and oversee the evaluation work. They were also expected to present their work at various forums both within the state and in other venues, including Harvard's Preschool Institute and meetings of P–3 initiative staff convened by the WKKF. The demonstration sites had already accepted the overarching P–3 goal of improving literacy skills at grade 3; they understood that within the framework of the initiative's focus areas, described below, they had considerable

discretion to design and implement activities that they believed promised the greatest payoff.

In 2010, the P–20 staff issued a second RFA and selected two additional sites—the Windward site, on the eastern side of Oahu, and the Honokaʻa site on the Big Island. The Windward site was awarded $1 million, and the far smaller Honokaʻa site received approximately half that amount. These sites began to implement their P–3 work at the beginning of 2011. Also in 2010, the Hawaiʻi Department of Education (HIDOE) learned that Hawaiʻi was one of 12 states to be awarded a Race to the Top (RTTT) grant from the U.S. Department of Education. The RTTT grants incentivized states to adopt ambitious reform plans for raising student achievement, promoting high school graduation, and reducing achievement gaps; states were asked to demonstrate the feasibility of those plans. The executive director of P–20 was one of five members of Hawaiʻi's RTTT core team, and the application included numerous references to P–20 activities, including the P–3 initiative. The Hawaiʻi RTTT plan included early childhood activities in one of the two original P–3 demonstration sites—Nānākuli-Waiʻanae—as well as a new site in Kaʻu-Keaʻau-Pahoa Complex Area (KKPCA) on the Big Island. Recognizing the opportunity to leverage RTTT funds, P–20 successfully applied to the WKKF for additional funding to add the KKPCA site as a fifth P–3 demonstration site. This site received $650,000 for work during the remaining years of demonstration site funding.

RAND evaluation staff visited the first two demonstration sites in 2010 to learn about site context, meet P–3 team members, and gather information about the activities that were planned with P–3 support. In subsequent years, RAND staff visited all of the demonstration sites, working with each (and with P–20) to design logic models within the focus area framework. The logic models served as monitoring tools for P–20 as well; sites reported on each of the outputs and outcomes of their activities on a regular basis.

From the beginning, there has been a significant amount of turnover in P–20 staffing for P–3 work. At a number of points, this turnover occasioned a restructuring of the staff, although the number of staff has not changed. The most recent change, in 2013, occurred because a senior P–20 staff member was tapped to be the second director of the EOEL. This departure was viewed as salutary; a person who knew P–3 inside and out would now lead the office. The staff restructuring that followed led to the hiring of

a considerably more junior person. This restructuring left P–20 without a strong P–3 policy leader; this seemed to signal P–20's decision to move to a more supportive position on state-level activities as the P–3 grant was approaching its end.

A supplement that WKKF provided to the P–3 grant was the inclusion of Hawai'i in state "Learning Lab" meetings, which brought together stakeholders from four states that were undertaking P–3 initiatives. WKKF convened these meetings at least once a year over half a decade to enable those working on P–3 activities to exchange information and experiences and obtain training on topics that were related to their work, such as communicating the importance of early childhood education to other state stakeholders. Opinions about the value of the Learning Lab meetings were decidedly mixed. Some found the Learning Labs to be very important in changing attitudes and promoting early childhood efforts in the state. The Learning Labs presented an opportunity to bring in new stakeholders and hopefully convert them to early childhood advocates. This worked effectively in Hawai'i: A number of stakeholders reported that the Learning Labs or their time at Harvard's PreK–3 Institute helped them to understand the importance of early learning.[5] Others, particularly those who were already committed to early learning, reported that the purpose of the Labs remained obscure. WKKF hosted the last Learning Lab gathering in 2012.

WKKF oversight and direct engagement with the work of the initiative varied over the course of the grant. The WKKF also added new expectations at several points. For example, at one national Learning Lab meeting, WKKF staff announced that states were to begin incorporating racial equity into the work.

Early Childhood Context

As the demonstration sites launched their activities and designed their logic models, the early childhood context in the state continued to change in ways that were generally supportive of the P–3 work. A signal event was the election of Neil Abercrombie as the governor of Hawai'i in 2010. According to many stakeholders, he made early

---

[5] Harvard's PreK–3 Institute, sponsored by its Graduate School of Education, brought together educational practitioners, administrators, and policymakers in a multiday program that provided an overarching framework and practical strategies to develop and sustain effective PreK–3 programs. The WKKF-sponsored Learning Labs, a national project to radically improve early learning (birth to age five) for all children in the United States, sponsored annual meetings through 2012 that brought together partners located in Florida, Hawai'i, Mississippi, and Washington state.

childhood a policy priority. He established the EOEL as a Cabinet position, delivering the department that activists had been wanting for years. The establishment of the EOEL was hailed as a major gain for the early childhood community as it provided a new, highly visible leadership role for state government in early childhood. P–20 began to work closely with the EOEL on a number of activities. There was a perception that the EOEL was likely to carry forward much of the P–3 work that P–20 had been undertaking after the P–3 grant ended.

According to stakeholders, Governor Abercrombie also actively supported efforts to increase the number of PreK slots in both the public and the private sectors. The ability to transfer public funds to private PreKs was considered important by most early learning advocates, as the vast majority of children in Hawai'i are served in private PreK settings. Furthermore, HIDOE lacks capacity to accommodate a substantial number of additional public PreK classrooms. Because the state constitution prohibits the transfer of public funds to private providers, it was deemed necessary to change the state constitution to enable private preschools to receive general state funds that would allow them to offer state-funded preschool programs. A ballot initiative was drafted and put on the ballot in 2014 to allow public funds to flow to private PreKs; this initiative would require a two-thirds approval from voters.

The governor supported the amendment, but its approval was not assured. Besides needing two-thirds support, the structure of the initiative was confusing: not voting in favor of the initiative would be recorded as a "no" vote even if the voter simply chose not to vote on it at all. Perhaps even more important, the education community did not uniformly support the amendment; teacher unions opposed sending public funds to private institutions—particularly ones where staff were not unionized. The amendment was defeated.

In a surprise upset, Governor Abercrombie lost a primary contest so could not run for governor in 2014. The new governor, David Ige, does not appear to provide the same level of support for early childhood that the previous governor did. For example, as of six months after his election, he had not appointed a new director to EOEL to replace the previous director, who resigned when the new governor was elected. At the time of this writing in May 2015, the office is slated to move to HIDOE in July 2015; at that point most assume that a director will be appointed, but in the interim, the momentum that EOEL had built has reportedly come to a halt.

Over the course of the P–3 initiative, HIDOE has also been moving forward with a set of education reforms, many of which had implications for early childhood education. Funds were allocated to HIDOE from the legislature to open a small number of PreK classrooms on elementary school campuses. This funding was in part a response to changes in school entry age: effective July 1, 2014, a child must be five years of age by July 31 of the calendar year to enter kindergarten. It was considered necessary to provide parents of four-year-olds who were excluded from kindergarten entry under this policy an opportunity to learn. HIDOE opened 21 PreK classrooms in 18 elementary schools. Given the impetus for this funding, it is unclear what will happen to these classrooms when all late-born five-year-olds enter kindergarten in school year 2015–16. But to many in the early childhood community, these classrooms are significant because they represent the first HIDOE delivery of PreK outside of early intervention preschools.

In the two RTTT sites, use of the Danielson Framework for Teaching (FFT) tool (Danielson, 2007)[6], which is often referred to as the "Danielson," was required as part of the RTTT work. HIDOE mandated its use in all HIDOE schools as of school year 2014–15. HIDOE also supported P–20's work on the Hawaiʻi Early Learning and Development Standards (HELDS), research-based standards that identify expectations about what children should know and be able to do at different ages, from birth to kindergarten entry. The standards were to align with HIDOE's K–12 standards. The federal early childhood context has changed as well over the five years covered by this evaluation. Of particular relevance, Head Start mandated the use of the Classroom Assessment Scoring System (CLASS) in all Head Start classrooms in 2012 as a basis for renewal of contracts. All four of the Head Start grantees in Hawaiʻi are using CLASS; the two Early Head Start grantees are not. In addition, the private school system Kamehameha Schools instituted the use of CLASS as well as the Danielson in their PreK program; CLASS was implemented prior to the P–3 initiative. The use of CLASS

---

[6] Created by Charlotte Danielson, the FFT is a research-based tool to evaluate teacher effectiveness that is used as the basis for teacher evaluation systems in thousands of schools. It was adopted by the state of New York, for example, in 2011, and is used statewide in Arkansas, Delaware, Idaho, and South Dakota. Supporters of the FFT argue that its use ensures a consistent process for evaluating teacher effectiveness based on a solid foundation of research and is demonstrated to be strongly correlated to student growth.

established a precedent that supported the P–20 policy that demonstration sites would promote the use of CLASS for coaching-based professional development.

Site-Level Background

In examining the work of the demonstration sites, it is important to remember that the sites participated in P–3 work for widely varying amounts of time. The first two sites, Farrington and Nānākuli-Waiʻanae, began P–3 work well before the other three sites. If one "counts' the early work before the 2009 RFA, which was funded at a low level, the two original sites received P–3 funding for close to seven years. In contrast, the newest site, KKPCA, first began funded P–3 work after Hawaiʻi was awarded its RTTT grant in 2010. Moreover, the five sites came to P–3 with very different experiences in related work. The N-W site, for example, is the home of the Institute for Native Pacific Education and Culture (INPEACE), a long-established nonprofit community organization. INPEACE was a partner in leading the P–3 work and was the grantee for the state of Hawaiʻi's SPARK site. SPARK, which stands for "Supporting Partnerships to Assure Ready Kids™," was also a W. K. Kellogg Foundation initiative that aspired to create seamless transitions to school for vulnerable three- to six-year-olds. N-W partners who had been involved with early childhood work in the area for some time uniformly viewed the majority of the N-W P–3 work as the next iteration of the SPARK initiative; they saw P–3 as providing the resources to continue the programs that were most successful under SPARK.

In the two RTTT sites, significant amounts of resources began to flow into K–12 schools, and some of those funds supported early childhood work. For example, preschool subsidies allowed children who had not previously attended PreK to do so at no cost to their families. RTTT rules also presented some challenges to P–3 efforts, particularly those targeted to K–3. Elementary schools had to comply with RTTT around assessment and other requirements. These requirements led elementary teachers to resist assessment requests from P–3.

These different site characteristics complicate the assessment of P–3 outcomes. The sites vary in length of exposure to P–3, in experience with P–3 concepts and activities, and in available resources. They also differ in more standard ways; for example, some sites are very rural with very limited PreK infrastructure, some have high percentages

of Native Hawaiʻians or immigrants, and some face the challenges associated with being located on an island other than Oahu.

3. Methods

As noted above, the evaluation employed multiple methods, sources of data, and informants over time (see Figure 2.1 in Chapter 2 for a timeline of the initiative). Qualitative data were collected in the course of individual and group interviews with P–20 staff, demonstration site team members, and other key stakeholders, and through review of documents. Quantitative data derived from two sources: School Information Forms (SIFs), which were completed by elementary school principals in the five demonstration sites, and third-grade reading scores collected by HIDOE. The methods employed to collect and analyze these data are described in detail below.

Qualitative Data

The primary goals of RAND's qualitative evaluation were to
- Identify the strategies that P–20 and the demonstration sites employed and assess their alignment with best practices in P–3
- Analyze the degree to which the five demonstration sites and P–20 executed their strategies and plans described in their logic models
- Assess the likely long-term effects of the site- and state-level work
- Determine whether P–20 and the demonstration sites engaged in activities that promote system change and assess the degree to which system change occurred.

To achieve these goals, RAND collected information about the strategies that P–20 and the demonstration sites included in their logic models, which defined their work. We compared these strategies to what were considered best practices in P–3 at the time. We then assessed the implementation of the plans made by the demonstration sites and P–20 and examined the likelihood that the work at the site and state levels would have long-term effects. Finally, we examined the work of P–20 and the sites from a system change perspective, which is described below.

Identify Strategies

To identify the strategies that demonstration sites and P–20 followed, we reviewed P–20 and demonstration site documents. We also collected information about the

strategies in interviews with P–20 staff, the demonstration site teams, and other key early childhood stakeholders at the demonstration site and state levels. P–20 staff and the demonstration sites each created a logic model based on a template P–20 developed. We compared these logic models to both P–20 goals and best practices for P–3 work as they were understood when the logic model work commenced in 2009.

Analyze Plan Implementation

Once the logic model template was developed in the first evaluation year, the demonstration sites were required to submit a progress report and plan update in each subsequent year that described completed work as well as any updates to activities to be undertaken, expected outputs, and expected outcomes. We reviewed logic models and conducted interviews with demonstration site staff and P–20 staff about their plans and progress. These interviews were conducted on the phone and in-person yearly until the final evaluation year, when interviews were conducted exclusively on the phone. In the final year, two rounds of interviews were conducted to capture both final achievements and assessments about the sustainability and long-term effects of the work. Our analysis of plan implementation compared the planned activities outlined in the logic models to the completed activities in the annual progress reports.

Assess Long-Term Effects and Sustainability of the Work

In the course of logic model reviews and interviews with P–20 staff and demonstration site teams, we explored how sustainability and long-term effects were being considered and planned for. In later years, we specifically focused on what P–20 and the sites had done to continue the work beyond WKKF support. We also examined respondents' views about the likelihood that the work would be continued and the mechanisms that would facilitate its continuation.

Conduct Systems Analysis

The request for proposal (RFP) for the P–3 initiative evaluation specifically requested assessment of the initiative from a systems change perspective. To enable this analysis, we relied on a framework developed at RAND that draws from previous work on accountability systems in public agencies (e.g., Stecher et al., 2010; Gormley and Weimer, 1999) and in the private sector (e.g., Welch, 2001; Pande and Neuman, 2000). The framework also draws from work on standards-based accountability in

education (e.g., Armstrong, 2002; Hill and Bonan, 1991; Adams and Kirst, 1999; McLaughlin and Shepard, 1995) and education reform work more generally (e.g., Lieberman, 2005). It directs attention to the system components and processes that together define social systems focused on producing defined outcomes, which generally include expectations, responsibilities, rewards, and outcomes, and specify who is accountable, to whom, for what, as well as the consequences for meeting or failing to meet specified responsibilities (e.g., Hill and Bonan, 1991; O'Day, 2002; Rothman, 1995). A key advantage of this framework is that: (1) it can be used to describe systems and track differences over time and space (as we are doing here beginning in year 1 of our P–3 evaluation); (2) it can be applied to all levels of a system including all the players in a hierarchy; (3) and it can be used to not only monitor the implementation of innovations but identify changes that might bring the system components into better alignment and thus help to improve system functioning. Using the framework, we analyzed the work of each demonstration site as well as P–20 as a system, then examined the interrelationships of these individual systems, particularly as they support or mitigate change at each level of the Hawaiʻi P–3 initiative.

The framework allows us to answer key questions, such as, what incentives are in place to promote the activities believed to be important to producing site outcomes? Are performance standards clear? And, how well is the system working to achieve its goals? The framework includes five components:

1. Setting explicit goals, expectations, and standards for the system
2. Clarifying the responsibilities of key system actors
3. Establishing incentives for participation and appropriate consequences for meeting (or failing to meet) expectations and standards
4. Monitoring and evaluating the performance of key system actors and entities and reporting on progress in a transparent way
5. Ensuring that key actors have the capacity, including training, resources, and authority, to carry out their respective responsibilities.

Our analyses examined the interview data and documents that we collected to determine the extent and the manner in which each of these system components operated and to understand the degree to which each component was addressed in policies and activities. Then we analyzed the ways in which the different components

were aligned with one another, focusing particularly on the extent to which these components appeared to be working together to promote system goals and identifying gaps in alignment that are likely to interfere with system functioning. This assessment of the strength of individual system components and their alignment is a useful indicator of successful systems (Zellman et al., 2009; Ryan and Martinez, 2008).

Data Sources for Qualitative Analyses

The qualitative analyses described in this report draw from several sources of data, including in-person and telephone interviews, review of initiative-related documents, and reviews of relevant literature. These data were collected from multiple sources, including P–20 staff, demonstration site teams, and other P–3 stakeholders.

Documents. Documents that were reviewed included proposals submitted by P–20 to the W. K. Kellogg Foundation, the proposals that the demonstration sites submitted to P–20, strategic plans submitted by the demonstration site teams, copies of presentations made at various stakeholder meetings, state legislation related to early childhood issues, HIDOE guidance on such issues as the definition of "highly qualified teachers," narratives about the Hawai'i P–3 project in WKKF publications, project budgets, and other written materials. We also examined the P–3 website and the literature on P–3.

Interviews. RAND staff conducted interviews with P–20 staff and demonstration site team members, as well as HIDOE Complex Area Superintendents. We spoke on the phone with P–20 staff approximately twice per month to discuss P–3 initiative activities and to learn about changes in the context of early learning in the state.

At each of the five demonstration sites, we interviewed the site teams as a group; teams generally included the site's P–3 project coordinator and at least four other team members. These other team members represented local public schools (employees of HIDOE), early childhood education providers, and other organizations directly involved in the P–3 project.

We also interviewed other P–3 stakeholders to obtain a broader view of P–3 efforts. These included officials from the HIDOE and Kamehameha Schools (a private school system that serves Native Hawai'ian students); individuals from local child and family advocacy organizations and foundations, including Good Beginnings Alliance; and members of the original P–3 advisory group convened when the state received the

initial grant from the WKKF. We spoke more than once to the vice president for policy and planning at the University of Hawaiʻi, which is the official grantee for the WKKF grant that supports the Hawaiʻi P–3 work. We met with these stakeholders individually and in groups ranging in size from three to 12. (See Table A.1 in Appendix A for a listing of the interviews we conducted over the course of this evaluation.)

Prior to each interview, RAND staff provided information about the purpose of the interview and the confidentiality of responses. Interviewees provided oral and later written consent to proceed with the interview. The RAND Corporation Human Subjects Protection Committee and the University of Hawaiʻi Institutional Review Board reviewed and approved all consent statements and processes for the site visits and interviews. Additionally, HIDOE also reviewed study procedures that involved HIDOE data or staff. Prior to the interviews, RAND generated a set of open-ended questions to guide our discussions with each group and ensure that we obtained the information needed to address the key objectives of the project. For example, we asked explicitly about how the site plan addressed each of the P–3 focus areas, and we asked site staff to describe how they would know whether their objectives were being met. Interviewees were also encouraged to provide information on other topics that they wished to discuss. (See Appendix A for a copy of the 2014 interview guide for the demonstration sites. A guide for stakeholders asked different but related questions. Questions varied slightly over time.)

For all meetings, two interviewers were present. While both interviewers took notes, one interviewer had primary responsibility for note taking. These notes were then reviewed by the other interviewer, and any discrepancies were resolved.

Logic Models. As described above, the evaluation RFP explicitly required that RAND work with each site and the P–20 team to develop a logic model for their P–3 work. Development of a consensual logic model for each site was particularly important given the objectives-based contracts the sites had signed with P–20. The site's logic model would make explicit the activities to be carried out and the expectations concerning what outputs and outcomes these tasks would produce. These logic models were a key source of data for documenting the plans and progress of the demonstration sites and the P–20 work.

Quantitative Data

Two types of quantitative data were collected. An online SIF was sent to all elementary school principals in each of the demonstration sites. We also conducted an analysis of third-grade reading scores that compared children in the demonstration sites with other third-graders.

School Information Form

The purpose of the SIF, developed at RAND for this evaluation, was to elicit data about the degree to which the P–3 work was known to and valued by elementary school principals in the demonstration sites. A substantial body of literature indicates that principals are the instructional leaders of their schools. If they are aware of P–3 and see value in the P–3 work, this suggests that the P–3 team has been effective in getting its message out to an important constituency. SIF data would also help to clarify the extent to which local elementary schools were conducting activities consistent with P–3 strategies and goals.

The approximately 30 SIF questions (the number varied slightly across administrations, as a few additional questions were asked in later rounds) focused on principals' awareness of P–3 and involvement in P–3 activities, such as meetings. Several questions explored the degree to which teachers in the school were active in early childhood activities, including conferences, professional development, and early childhood certification. Principals were asked about activities that occurred in the school that were consistent with P–3 practices, such as coaching, CLASS assessments, and release time for cross-grade and cross-sector (PreK–kindergarten) meetings, and PreK–kindergarten transition activities. A final set of questions asked principals to assess the local P–3 work. (See Chapter 4 for examples of several SIF questions.)

The initial round of SIFs was completed in 2011 by principals in the first two demonstration sites. We were not able to distribute the form in 2012 because a change in HIDOE research review procedures slowed approval beyond a reasonable data collection window. In 2013, SIFs were completed by principals in all five demonstration sites. In 2014, another round of SIFs was fielded. At this time, a serious lava flow forced the evacuation of homes and schools in one of the demonstration sites. Students and faculty were reassigned to temporary locations, and the 2014 SIF was not fielded in this site.

With active support from the complex area superintendents in each site, the response rates were very high. In 2011, 93 percent returned a form (n = 14/15). In 2013, with all five demonstration sites responding, 93 percent of principals returned a form (n = 43/46). In 2014, in the four sites that were surveyed, 83 percent of principals returned a form (n = 33/40). Table 3.1 summarizes the dates of administration of the SIF by demonstration site.

Table 3.1

SIF Administration by Demonstration Site and Year

	Year		
Demonstration Site	**2011**	**2013**	**2014**
Farrington	X	X	X
Nānākuli-Waiʻanae	X	X	X
Windward		X	X
Honokaʻa		X	X
Kaʻu-Keaʻau-Pahoa		X	

For this report, we pooled the information from all sites in order to have a large enough sample to compare responses over time. We tested for differences across time using t-tests or Chi-squared tests, and we also tested for differences by principal tenure, which was suggested by descriptive analysis of the data. In general, as reported below, we find very few significant changes over time. Hence our reporting of SIF results focuses on the degree to which elementary school principals reported knowledge of P–3, use of P–3 in the final year of the initiative, and principals' attitudes about P–3.

Third-Grade Reading Scores

The goal of the Hawaiʻi P–3 initiative was for every child in Hawaiʻi to read at grade level by third grade. This analysis examines whether children who might have been exposed to P–3 in demonstration sites appear to have made progress toward this goal. We examine student test-score data collected by HIDOE. This analysis is unusual in the context of a study of a P–3 initiative: Most P–3 evaluations do not include a

standardized outcome measure, despite the fact that many strive to improve such student outcomes.

The test-score data we analyze are drawn from the Hawaiʻi State Assessment (HSA) reading assessment. This reading assessment is an adaptive criterion-referenced assessment, and the scores reported at the end of each school year are used to determine schools' adequate yearly progress (AYP). The assessment is administered online. Although the HSA reading assessment is administered to students in grades 3 through 8 and grade 10, we use scores only from grade 3 in our analyses.[1]

The analyses include all of the elementary schools in the state with available test-score data, numbering 192. Each year of data includes test scores of between 13,000 and 15,000 students, with the number growing during the period represented in these data. The analyses include seven years of third-grade reading-score data; they include every student in the state aged seven to ten who took the HSA in each of the included years, totaling 98,909 students. We deleted data for children whose reported ages were six or younger and 11 or older, as the expected age for third-graders is eight or nine years old.

We compare reading scores across time starting with the spring of 2008 and ending in the spring of 2014. The analysis model is a student-level difference-in-difference model for students at all Hawaiʻi elementary schools. This analysis compares changes over time in student third-grade reading scores within demonstration site elementary schools to changes in those test scores for students in other elementary schools (see Cameron and Trivedi, 2005, for a discussion of difference-in-difference estimates). Our data represented 47 schools in complex areas that participated in the P–3 initiative. These models include school-level statistical controls for other time-varying factors that might influence outcomes such as the fraction of students qualifying for free or reduced-price lunch. The additional covariates that we include in the model are the student's gender, race and ethnicity (Asian, Pacific Islander, white, other), birthdate (year and month), free or reduced-price lunch (FRPL) eligibility, whether the student is designated as an English-language learner, whether the student repeated the third grade, whether the student was receiving special education services, a year indicator, a school indicator variable, the school rate of FRPL-eligible students,

---

[1] The last year of data that we use (school year 2013–14) was the final year that the HSA was used as the state language and mathematics assessment. In the spring of 2015, HIDOE began using the Smarter Balanced assessment for math and English language arts. See HIDOE, undated-a.

and the school percentage of Pacific Islanders. The models that we estimate are further detailed in Chapter 4, which discusses the third-grade test results.

Study Limitations

Study results must be considered against a number of limitations. First, due to limited project resources, we were able to conduct in-person site visits only once a year in the first four years, and none in the final year. To help overcome this limitation, we communicated with the demonstration sites by phone at least two additional times a year. We also met regularly with P–20 staff over the course of the study: mostly biweekly. In addition, consistent with our original design, we worked with P–20 and the sites to identify measures of progress that could be collected and monitored remotely.

Findings might not be generalizable statewide because of the self-selection of the funded demonstration sites, which, with one exception, applied for P–3 funds as part of a competitive process. Hence, the sites included in the study represent highly motivated sites which are not necessarily representative of a "typical" location for a P–3 initiative in the state. Additionally, P–20 also considered the needs of local communities in choosing among the demonstration site applicants, prioritizing locations that had the worst educational outcomes. Reflecting their high need, some of the demonstration sites had a history of externally funded initiatives (which in some cases dated back to a decade) that included features of P–3 work. Furthermore, one site was invited to join P–3 after being selected to receive RTTT funding. Hence, the experience of these P–3 "early adopters" cannot be assumed to represent the outcomes that might be seen were other complex areas to engage in P–3 initiatives.

Another consideration in interpreting the results in this report is that pinpointing a precise date when the P–3 work started is difficult. Identifying a start date for the first two demonstration sites is complicated by the fact that they had been receiving P–3 funds for two years as part of a group of 17 P–3 pilot projects, before the P–3 demonstration site work officially "started" in 2009. The first two sites were also advised to use the first year of the demonstration site grant primarily for planning, which further complicates the establishment of a start date.

Besides having different start dates and different histories of work related to P–3, the sites also differed in terms of the level of resources that they were receiving from

HIDOE during the evaluation period. In the two RTTT sites, huge amounts of resources began to flow into K–12 schools, and some of those funds supported early childhood work. For example, preschool subsidies allowed children who had not previously attended PreK to do so at no cost to their families. At the same time, RTTT complicated some P–3 efforts, particularly those in K–3. Elementary schools had to comply with new RTTT requirements around assessment and at least initially teachers were expected to teach a longer day. Time spent on RTTT activities necessarily competed with time available for some of the P–3 activities.

Despite the mandate to conduct work in all focus areas, demonstration sites also differed in terms of both the specific activities they pursued within each site and the activities to which they devoted the largest share of their time and resources. As discussed in more detail below, most sites focused their work within the PreK-to-first-grade period. A typical site supported PreK-to-kindergarten teacher visits and joint coaching. One site stood out in directing considerable efforts toward creating professional development opportunities in early childhood for HIDOE K–3 teachers. Another site focused its efforts on recruiting families into early childhood programs. Because of this diversity in activities and focus, it is impossible to regard the P–3 initiative as a single entity. This in turn makes it difficult to explain how and why the initiative produced the outcomes it did.

Other characteristics on which sites differ further complicate the assessment of P–3 outcomes. The sites clearly vary in length of exposure to P–3, in experience with P–3 concepts and activities, and in available resources and focus. They also differ in other ways; for example, some sites were very rural with very limited PreK infrastructure, some have high percentages of Native Hawai'ians or immigrants, and some face the challenges associated with being located on an island other than Oahu.

A different type of study limitation might be found in the lack of a child-level assessment of kindergarten readiness or school performance prior to grade 3. Such data would have been highly desirable because they would have been measured without a significant time lag between the implementation of activities focusing on early childhood outcomes and those that focused on child outcomes. This reduced lag time would have conferred the advantage of both reducing attrition among children exposed to the P–3 activities and increasing confidence that effects were due to P–3 activities rather than those of other initiatives. Instead, we had to wait until children exposed to

28

P–3 reached grade 3 to assess reading scores that are routinely collected by HIDOE. Since fall 2013 was the first year in which a cohort exposed to the complete set of P–3 activities reached third grade, our analysis was undertaken at a time when only a fraction of students would have had full exposure to P–3 activities and therefore could benefit from their full effects. Further, if there are threshold effects operating where a certain level of intervention is necessary to show an effect, then the more recent sites are even less likely to show effects. Threshold effects have been found in studies of the effects of prekindergarten classroom quality on a range of child outcomes (e.g., Burchinal et al., 2010). Moreover, because the measure was collected years after many of the early childhood components of the initiative, there was inevitably some attrition of students who were exposed to P–3, and new students were tested who had no P–3 exposure. These latter considerations are likely to downwardly bias our estimates of the effect of the P–3 initiative on students' reading scores.

The substantially different amounts of time that sites spent participating in P–3 work suggest that outcomes by site might differ. The first two sites began P–3 work well before the other three sites. If one "counts" the early work before the 2009 RFP, the two original sites received P–3 funding for close to seven years. In contrast, the newest site first began funded P–3 work after Hawaiʻi was awarded its RTTT grant in 2010. Innovations typically take years to be fully implemented, raising the question of what is reasonable to expect in terms of producing measurable improvements in the more recent sites. These differences might have particular implications for the third-grade reading-score analysis: Third-graders in the more recent sites had very limited exposure to P–3 activities, as most of those activities focused on the PreK-to-first-grade level.

4. Site-Level Implementation and Outcomes

This chapter examines demonstration site performance from three very different perspectives. The first perspective relies on reports from site team members and other participants in the P–3 work, on observations conducted during site visits, and on documents produced by the sites, particularly progress reports imbedded in site logic models. We also include P–20 and stakeholder views. The second perspective draws on the results of a School Information Form, which was completed online by principals in each site (see Chapter 3 for details of SIF administration and response rates). Principals' views are important, as one goal of the P–3 work was to build relationships with K–12 and encourage principals and other K–3 staff to better understand and support efforts to prepare children for kindergarten and build the capacity of K–3 teachers in early childhood. A third perspective analyzes third-grade reading-score data collected by the state in all schools. These analyses allow us to examine whether students in demonstration sites achieved the ultimate P–3 goal of improved reading in third grade, compared with children who were not in these sites.

In the first section, we focus on the work of the five demonstration sites. In the next section, we present the SIF results. In the final section, we present our analysis of third-grade reading scores.

The Work of the Five Demonstration Sites

We review and analyze the work of the five demonstration sites by addressing the four research questions listed below. For ease of readability and being mindful that this report is likely to be read outside of Hawai'i, we have chosen not to present demonstration site analyses separately. Instead, we discuss the important overall issues and the ways in which demonstration sites managed them, highlighting particularly interesting and informative activities and responses. Appendix B describes some of the work of the sites in more detail for those who might wish to learn more about what activities the sites chose to pursue. We relied on the following research questions to guide the analyses:

1. Which strategies did the P–3 demonstration sites employ, and did these align with best practices in P–3 at the time?
2. Did the demonstration sites successfully implement those plans?
3. Did the demonstration sites develop plans likely to promote long-term effects or sustainability of their P–3 efforts?
4. Did the sites engage in activities and processes that promote system change? Did system change occur?

We describe the overall structure that P–20 devised for the demonstration site work first, then examine each of these research questions in turn.

Structure for the Demonstration Site Work

P–20 imposed on the selected sites a set of seven focus areas (later reduced to six when P–20 reclaimed the "Data" focus area for itself). These focus areas, shown in Table 4.1 below, incorporate the fundamental features that define most P–3 initiatives and include:

- Greater access to PreK for three- and four-year-olds
- Support for transitioning from PreK to kindergarten
- Alignment of curriculum, standards, and assessment from PreK through third grade
- Training and instructional practices that include training in early education and child-centered learning approaches as well as teacher-directed approaches
- Parent engagement in learning from PreK through third grade
- Use of data for quality improvement and accountability
- Emphasis on the "whole child" concept of learning (Kauerz, 2008).

The P–3 demonstration site Request for Applications asked potential sites to describe their planned work in terms of a framework defined by seven focus areas selected on the basis of available research evidence. The first two sites could choose the focus areas in which they would work, but soon P–20 asked these sites and subsequent sites to engage in activities in all focus areas. In planning their activities, sites tended to choose activities that reflected local needs; consequently, they did not emphasize the focus areas equally. Each of these focus areas included a number of suggested activities. In Table 4.1 below, we display the seven focus areas along with some illustrative examples of activities in each area. (See Zellman and Kilburn, 2011, for more detail

about the focus areas and a description of the activities of the first two sites by focus area.) There is a clear and close alignment between the focus areas in this table and the features listed above that were considered central to P–3 initiatives at that time. Hence, by adopting this structure for the demonstration site activities, P–20 ensured that sites' work would encompass what are considered the essential elements of P–3.

Table 4.1.

Seven Focus Areas for Demonstration Sites and Illustrative Activities

Leadership for Literacy	Convening of P–3 leadership, including early childhood administrators and elementary school principals; administrators exposed to research-based strategies to promote student learning
Standards, Curriculum, and Assessment	PreK and K–3 teachers meet to review PreK through third-grade standards and align educational expectations; curriculum mapping.
Instruction	Selection of common instructional assessment tool; training on selected tool
Teacher Professional Development	Scholarships for K–3 teachers seeking Early Childhood Education (ECE) certificate; peer coaching model
Comprehensive Early Learning Services, Access to 0 to 5 Services	Map of local services; school-based strategy for service referrals; increased opportunities to participate in PreK
Family School Transitions and Partnerships	Development of transition plans for incoming kindergartners; parent early literary education
Data	Collection of available student-level ECE data; discussion of future data needs

In the first year of funding, the two selected demonstration sites were asked to devise and submit a detailed P–3 plan using the structure of the seven focus areas. These demonstration sites were given considerable autonomy in terms of which focus areas they would work in and the activities they chose to pursue. This decision to

provide the sites autonomy over the P–3 activities they chose to pursue was considered the best way to ensure that the selection of activities met local needs and drew on local strengths. P–20 believed that the clear overarching goal of the initiative and the objectives implicit in the focus areas would be the most effective way to "move the needle" toward that goal. The autonomy given to the sites fostered deep discussion among the site teams over which activities to pursue and which outputs and outcomes of those activities should be expected and monitored to increase the odds of reaching local objectives and P–3 goals.

Over time and with changes in P–20 staff, P–20 became more prescriptive, asking all sites to engage in activities in the first six of the focus areas shown in Table 4.1. This requirement had little direct bearing on activities in the operating sites; site teams reported that when they sat down to address this somewhat daunting new requirement, nearly all discovered that they were already engaging in activities in each focus area. Consequently, this new requirement did not change the unique emphases of each site. Sites continued to emphasize different focus areas and pursue very different activities within focus areas. For example, efforts to increase PreK enrollments in one site involved an extensive campaign to knock on every door to inform parents of local PreK options available and encourage them to enroll in one. Other sites attempted to meet this objective by developing a listing of all local PreK programs and their basic characteristics. Two sites with limited PreK options opened small parent–child Play and Learn programs to provide more children with some PreK experience. (See discussion below for further detail.)

P–20 also asked the sites to develop a logic model that described the activities that the site planned to undertake in each focus area and the resources that would be used to carry out the activities. Sites were asked to define activity outputs and outcomes. P–20 required that annual progress reports be imbedded in each site's logic model, which meant that each site had to describe progress in each of the six focus areas. P–20 expected site teams to carry out the activities included in their logic models and meet the milestones the sites themselves had set. These reports were carefully reviewed to see how well each site was meeting its own goals and milestones. From P–20's perspective, meeting these milestones in each of the focus areas represented important progress on the fundamental features of P–3. P–20 did not impose any additional site

performance expectations, nor did it ask sites to consider or plan for ways to continue the P–3 work after funding ended.[1]

Even before the demonstration sites were selected, P–20 imposed a strategic demand on applicant sites. To be eligible for funding, demonstration site proposals had to include the signatures of both PreK and HIDOE leaders. This mandate set in place a clear expectation that P–3 work in the sites would engage both stakeholder groups, a necessary condition to enable successful P–3 work, particularly alignment of PreK with K–12. Indeed, the building of PreK and K–12 partnerships was expected to be an ongoing activity in each of the sites.

Summary Description of Sites' Work

Which strategies did the P–3 demonstration sites employ and did these align with best practices in P–3 at the time?

The joint signing requirement and the expectations around the focus areas essentially compelled the demonstration sites to conduct work that would further the initiative's P–3 goals, which are defined through the focus areas. The focus area requirements and the expectation for cross-institution partnership-building resulted in new work for many of the sites.

Leadership for Literacy. In every demonstration site, the P–3 team brought together both HIDOE and PreK representatives. Representatives of these groups became members of the P–3 site team, which met frequently to plot its course, launch and oversee activities, work through budget and administrative issues, and produce progress reports. In some sites, the site teams also included private school representatives and representatives of other community organizations.

HIDOE representation varied across sites, and reflected in part both the level of commitment of HIDOE staff and the nature of the work that was being implemented in the site. In three sites, the complex area superintendent, the highest-ranking HIDOE administrator in the area and the person responsible for management of all schools in the complex area, became a member of the site team and attended site team meetings. In one site, the two principals in whose schools many of the P–3 activities were

---

[1] Considerably later, P–20 asked sites to produce an issue brief that highlighted a notable activity with measurable outcomes. Sites did not conform to this request, as discussed later in this chapter.

implemented were active team members; principals sat on site teams in several sites. In the fifth site, the site coordinator was well-connected with HIDOE and frequently drew on the support of a CAS who was strongly committed to the work.

The inclusion of HIDOE staff on site teams, the requirement of HIDOE involvement as a condition of P–3 funding, and a growing sense in the state of the importance of both PreK experience and PreK to K–12 alignment led to stronger PreK to K–12 relationships in all the sites. It was not always easy to forge these relationships. In some sites, cooperation of this sort was new, and stakeholders had to spend time getting to know and trust one another. In the two RTTT sites, HIDOE had to cope with a raft of new requirements, including the use of new observational assessment tools, which made it more difficult to accede to site team requests for meetings and particularly for CLASS observations of K–3 teachers. Staff turnover also affected these relationships. In one site, a CAS strongly committed to P–3 left as the work was just beginning; the replacement was less engaged. But in other sites, new blood brought more enthusiastic commitment. In still other sites, the site team made it its business to educate the CAS about the importance of the P–3 work. Support from P–20 for travel to national P–3 conferences, strategically offered, was effective in turning some CASs and other HIDOE administrators into converts.

Standards, Curriculum, and Assessment. Most of the work under this strategy focused on training for and implementation of the CLASS. This reflected the fact that CLASS implementation came to be required by P–20 after a period when sites were required to adopt an observational assessment but were given autonomy concerning which one to use. The earlier search for (and in some sites, implementation of) another tool and subsequent implementation of the CLASS was time-consuming. Nevertheless, sites engaged in other activities in this focus area as well, as did P–20. These efforts are described briefly below.

All but one of the demonstration sites implemented the CLASS in at least some classrooms.[2] In the two RTTT sites, the CLASS work met with some resistance because of the mandate to implement the Danielson FFT. However, in one of these sites, the site team worked with volunteer HIDOE teachers, who generally found the CLASS to add

---

[2] That site was unable to recruit and train CLASS observers who met CLASS reliability standards. It hoped to import CLASS observers from another site, but these reliable observers are rare and in demand so using them proved to be impossible.

value. P–20 initially wanted all demonstration sites to use the CLASS because it can be used in both PreK and elementary school settings, but accepted the Danielson when it became apparent that HIDOE teachers might resist engagement with an additional observational assessment.

P–20's goal was for at least half of PreK and K–3 classrooms in the demonstration sites to be observed at least yearly using either the CLASS or Danielson tool. The sites have fallen short of this goal for a variety of reasons, including challenges in getting observers trained and reliable (after completion of training, some do not meet reliability criteria). In addition, demonstration site staff shared concerns about the longer-term value of CLASS. Many doubted that CLASS could be sustained after P–3 funds ended because of high costs associated with observations and reliability maintenance. Nevertheless, the albeit limited use of CLASS in nearly all sites introduced participating teachers to the idea and process of evaluation, feedback, and improvement through coaching. Site team members reported that teachers found the feedback and coaching to be enormously helpful.

In addition to CLASS, demonstration site work in this area included PreK to kindergarten articulation meetings in one site; collection and distribution of curriculum used in PreK, kindergarten, and grade 1 classrooms in another; and work on the Peabody Picture Vocabulary Test (discussed below) in a third.

Instruction. P–20 considered the implementation of common tools for assessing quality instruction to be an important component of the P–3 work from the beginning, arguing that such tools support quality teaching and pinpoint areas for improvement and professional development support. Work under this strategy was conducted in all sites, but one site was unable to implement CLASS-based coaching of teachers who had been CLASS-observed.

Instructional support was provided in other ways as well. In the site that was unable to implement CLASS, the curriculum specialist worked with kindergarten teachers on developmentally appropriate practices (DAPs). PreK and kindergarten teachers engaged in cross-visitations. Other sites worked on setting up policies to share student information, either through schools or through parent completion of shared forms. Work was also conducted on ways to teach to Common Core standards using DAP.

Teacher Professional Development. One site placed its emphasis on better-trained teachers. The belief of the site team was that teachers who understood the value of instruction assessment and feedback and early childhood development would be able to provide more-effective instruction to their students. This site created a series of five early childhood courses to inform teachers about important aspects of early childhood development and instruction. With support from P–20, these courses were accepted by HIDOE as eligible for salary points, an important incentive for HIDOE teachers. The availability of salary points will likely increase the number of HIDOE teachers who take these courses.

Several sites promoted professional development through the provision of joint PreK–kindergarten trainings as well as some grade-specific trainings. Another professional development activity that occurred in two sites was the collection and distribution of scholarship and loan information to those interested in pursuing early childhood coursework or programs.

Comprehensive Early Learning Services/Access to Services. Several sites emphasized provision of early learning services, and worked especially to encourage access. Sites approached this work in different ways. In one site, workers knocked on doors and provided parents of young children with information about available programs. Sites researched and collected information on available programs; some also provided information about openings and enrollment processes. A number of sites opened programs that provided some PreK exposure to incoming kindergartners. These programs typically involved children and parents, so that parents who might be reluctant to allow their young children to go to programs were more willing to participate.

Family–School Transitions and Partnerships. This focus area was a natural outgrowth of the joint work of the site teams. In one site, the CAS mandated the creation of kindergarten transition teams in each elementary school. All sites supported visits between PreK and kindergarten teachers, with some sites providing joint training by a curriculum consultant. All sites worked to educate parents about supporting their child's transition to kindergarten. This support was offered in a variety of ways, including transition programs at elementary schools, the distribution of materials that parents could use to help their child learn to hold a pencil and a book, and the distribution of bags that included age-appropriate books.

Did the demonstration sites successfully implement those plans?

Plan implementation and outcomes were assessed by comparing plans articulated in site logic models against yearly progress reports. In face-to-face and telephone meetings with site teams, we received updates, explored site progress in more detail, and sought to understand any gaps between plans and activities each year, as well as the reasons for changes in activities or strategies. The evaluation team also provided yearly internal reports to P–20 and to each of the sites about their progress. The discussion below reflects the sites' P–3 work during their participation in the P–3 initiative.

Sites were diligent in meeting their own goals. Across sites, teams successfully completed their goals in nearly all of the work in which they engaged. Three of the five sites met all goals outlined in their logic models. The few exceptions (e.g., inability to complete CLASS observations in one site; failure of one coordinator to attend community events to promote early learning services in the area in another; failure to complete a Directory of Services; and submission of fewer than five articles on early literacy to the community newspaper) were explained in detail in progress reports. For example, in the case of not completing CLASS observations, several individuals who completed training did not meet CLASS reliability criteria, and no one else was locally available who could undergo CLASS training. P–20 accepted sites' progress reports and the achievements they reported with little question once shortcomings were explained. P–20 did not attempt to compare sites or aggregate site achievements.

Our reviews of site logic models indicated a great deal of activity, most of which led to the expected outputs and outcomes. It was clear over time that PreK and K–3 actors were spending a good deal of time together; site teams, which always included PreK and HIDOE representatives (and in some sites, community organizations as well), met frequently: Meetings in some sites were often convened more frequently than originally planned.

Our review also revealed that many of the measurable outcomes found in all of the sites' logic models documented activities completed, usually through reports on participation levels. For example, one site was successful in getting all elementary principals to attend at least two professional development events. In that same site, 100 percent of PreK classrooms were CLASS-observed, and eight PreK CLASS-reliable observers were trained. About half of HIDOE schools and Early Childhood Education

(ECE) classrooms in that site implemented transition activities for incoming kindergarten students. Twenty-one HIDOE teachers visited 15 PreK classrooms. In another site that focused on increasing the number of students entering kindergarten with some PreK experiences, team members reported that they had knocked on 976 doors and successfully recruited 90 children or families into ECE providers. Annual conferences held by two sites attracted hundreds of attendees. Another site reported that 500 literacy activity material packets had been distributed at community events.

Other reports of measurable outcomes provided data on participants' reports of new knowledge gained or attitude change. One site reported that 100 percent of principals who attended professional development activities said they came away with a better understanding of P–3 principles. Another site report indicated that 90 percent of those who attended the Hawaiʻi Island Early Childhood Conference (HIECC) completed an evaluation form indicating that they came away from the conference with new information.

Some measures of success relied on observations or products. For example, in two sites a majority of CLASS-observed teachers improved their scores on the instructional support domain after observation and coaching. All participating teachers in these two sites saw improved scores on the emotional support and classroom organization domains. In one site, all (23) elementary schools completed transition plans in cooperation with PreK partners. One of the elementary schools that was in restructuring in one site made AYP in school year 2012–13 for the second year in a row, which removed it from restructuring under No Child Left Behind.[3] Site team members attributed some of this improvement to the work that their curriculum specialist had done with kindergarten and grade 1 teachers to teach them about and encourage the use of DAP.

Sites reported achieving a number of outcomes that were more difficult to quantify. Nearly all sites reported that relationships with HIDOE were better. Several noted that there was more of a shared understanding of the importance of PreK: several members of different site teams believe that now, HIDOE and other stakeholders more clearly see that good PreK classrooms produce school-ready kids. Site team members in another

---

[3] Under the No Child Left Behind Act of 2001, a restructuring plan must be developed when a school or district has not made adequate yearly progress in the percentage of children meeting grade-level standards for five consecutive years. See Learning Point Associates, 2006.

site noted that stakeholders now better understand that early learning is critical to K–12 success. There is much more buy-in from principals and support for efforts to align PreK to kindergarten practice, and for K–3 teachers to understand and integrate developmental theory into teaching. Several site team members from different sites reported that now, in contrast to what happened before P–3, they can pick up the phone and call a principal and the call is well-received.

Team members in one site noted that the flexibility of P–3 funds was critical[4]: The site coordinator could quickly write a check, if necessary, and this really helped to build relationships. For example, in this site P–3 fronted funds for Head Start teachers to pursue bachelor degree programs that Head Start paid for only when the program was successfully completed.

Sites also noted the importance of P–3 funding covering multiple years: Site team members in some of the earlier sites claimed that it took years before relationships were developed and groups began to volunteer to participate willingly. One site team member noted that despite early support from the CAS and resource teachers, the principals' first response to P–3 was, "…this too shall pass." After some years, the same site team reported witnessing "aha!" reactions from principals: They began to understand the point of P–3's efforts. As a result, principal buy-in reportedly increased substantially.

Site team members believed these partnerships will endure even without P–3 funds. Site teams in every site asserted that the P–3 work left a legacy of schools with more engagement with parents and community partners. HIDOE is now taking on some tasks that P–3 began (see the Long-Term Effects subsection below). The P–3 project was also able to provide PreK teachers with some professional development, which they would not have received otherwise.

Did the demonstration sites promote long-term effects or sustainability of their P–3 efforts?

The long-term effect of time-limited funding such as P–3 deserves examination. Given the substantial funds that WKKF provided, it is important to explore the extent to

---

[4] Since all of the fiscal agents for the demonstration sites were not-for-profit organizations, their procurement process was less restrictive than that of state entities. Site teams were allowed to allocate funds as needed within activities. Total expenditures for activities were monitored, but if they were within the annual budget, P–20 did not question them.

which the funding will have an enduring effect. Here, we discuss the long-term effect of the site-level work. In the next chapter, we discuss the effect of the state-level efforts.

P–20 did not actively encourage sites to consider long-term effects of their work or sustainability of their activities in their planning, despite strong evidence in the research literature that long-term effects and particularly successful sustainability of externally supported activities are facilitated when sustainability is considered from the very beginning of a funded project (McLaughlin and Mitra, 2001; Buchanan et al., 2005; Tibbits et al., 2010).

Our analysis indicates that sites were determined to make the most of their P–3 funds and to also leave behind a legacy of support and activities. Indeed, in developing their work plans and logic models, nearly all site teams developed ideas about what activities were likely to have the most long-term effect or be sustained; these ideas helped them to choose where to focus their efforts. Most considered how best to move the site toward the ultimate goal of the P–3 initiative—all students reading at grade level by grade 3—in a way that left something behind when the P–3 funding ended. For example, one site considered the literature on the effects of PreK exposure on kindergarten readiness and focused its efforts on increasing PreK exposure. In the process of doing this, the site team argued, the community would come to understand and value early learning, and the norm about early learning would change. In another site, the research on the importance of DAP instruction in improving child outcomes led to a focus on developing and implementing courses on early childhood for elementary school teachers. HIDOE acceptance of these courses as being eligible for salary points was a signal achievement in terms of the long-term effect of the P–3 work, as discussed below. In two other sites, a focus on CLASS was believed to be an effective way to help teachers understand the value of assessments and provide them guidance in improving their instruction. In another site, an emphasis on PreK to kindergarten cross-visitation and data use to improve practice would encourage both PreK and kindergarten teachers to reorient their classrooms and their expectations in a way that would endure.

Attention to sustainability of P–3 work after the project ended was also evident in the sites. For example, one site decided early on that the P–3 project would build on local staff and resources and would not bring in people who offered to come in and set up programs because they would no longer be available when the funding ended. Site team members deliberated about delivering PreK services in the remote areas of the

42

complex. They made an early decision not to provide gas money for parents to bring children to these programs, because this was not sustainable once the grant ended. Instead, the site team worked with schools and community organizations to find locations in local communities for their Play and Learn groups.

Several sites noted that the requirement that P–3 funding proposals be supported by both early childhood and HIDOE created an opening to work with HIDOE staff, often those attached to Parent Community Networking Centers (PCNCs) in their area, and transfer activities such as parent programs to them after P–3 funding ended. This kind of outcome was not necessarily anticipated by P–20.

The relatively long P–3 funding time frame—from three to five years—allowed relationships to be built and norms to change over time. For example, this enabled one site to set in place a professional learning community (PLC) approach, which gained traction with principals over time. In the principals' view, this approach was more successful than one that would have required their participation, which might have met with resistance.

At least a few sites modified their activities over time with an eye to long-term sustainability. For example, without outside support, the CLASS could not be sustained in elementary schools or private PreKs because of the ongoing training and testing requirements for observers. CLASS is also very time-intensive. The fact that teachers, once trained, can collect data and use the data to modify instruction without additional support led the P–3 site team in one site to view the Peabody Picture Vocabulary Test (PPVT) as being more sustainable in the long run, although it represents a child-centered rather than teacher-centered approach to improving instruction[5]. P–3 offered trainings that taught participants how to administer the PPVT to students and how to use the resulting data. A website provides DAP lessons based on students' PPVT scores. According to site staff, most teachers have integrated the use of data and assessment into their practice because of their experiences with CLASS and PPVT. P–3 site staff hoped that HIDOE resource teachers will support these efforts and reinforce this approach.

Several sites attempted to win new outside support to continue the P–3 work. The site that has a long history of outside support applied for and received funding from the

---

[5] The PPVT is an untimed test of receptive vocabulary for Standard American English. It is intended to provide a quick estimate of verbal ability and scholastic aptitude (Dunn and Dunn, 2007).

Robert Wood Johnson Foundation for a community schools project, which includes some P–3 work. In that site, INPEACE continues other aspects of the work. One site was awarded state funding through Grant-in-Aid, which will enable it to continue some P–3 activities, including an annual transition conference, on a short-term basis. Two sites applied for funds to keep the work going, but were not successful. Site team members in two sites that did not seek outside support explained that they lacked the capacity to successfully compete for outside funding.

Overall, we noted only a few specific examples of attention to sustainability of the P–3 initiative: it was not a focus for P–20 and only a few sites specifically considered it in making programmatic decisions. However, most sites used P–3 funds for work that might have long-term and broader effects. The examples show efforts that potentially contributed to changed norms around the value of early learning and PreK participation. Principals now understand that what goes on in PreK affects kindergarten readiness; that PreK exposure is important; and that K–3 teachers with some early childhood background might be in a better position to teach young children.

Did the sites engage in activities designed to promote system change? Did system change occur?

The system framework that drives this analysis is described in Chapter 2. The framework posits that for systems to work well, five components must be considered and well aligned (Zellman et al., 2009; Ryan and Martinez, 2008):

1. Setting explicit goals and expectations
2. Identifying the responsibilities of key system actors and entities
3. Establishing incentives and appropriate consequences for meeting (or failing to meet) goals and expectations
4. Monitoring and evaluating the performance of key system actors and entities and reporting on progress in a transparent way
5. Ensuring that key actors and entities have the capacity to carry out their respective responsibilities.

We discuss each of these components below, addressing how P–20 responded to the sites with respect to the component and how the component played out in the demonstration sites.

Setting Explicit Goals and Expectations. A key strength of Hawaiʻi's P–3 initiative was the unanimity and clarity around its ultimate goal: all children reading at grade level by third grade. Since the inception of the evaluation, every individual with whom we spoke about the initiative indicated that this was the goal of Hawaiʻi P–3. This unanimity is notable, as education reform initiatives are not always able to devise and communicate specific measurable objectives with this level of clarity and support.

The unanimity observed surrounding this goal did not always apply to the means and methods for achieving the goal. As noted above, P–20 relied on objectives-based contracts with the sites that imposed a limited number of strategic mandates on them while allowing sites considerable flexibility in devising and implementing activities within the focus areas of the P–3 work. Perhaps the most notable mandate was the requirement that each demonstration site proposal include signatures of both PreK and HIDOE administrators, signaling an expectation that these two important entities would work together to further a fundamental P–3 goal of alignment.

Within the first two demonstration sites, P–20 initially believed that the best approach to achieving the third-grade literacy goal was to allow the demonstration sites to develop their own plans to reach the ultimate goal by proposing and engaging in activities in one or more P–3 focus areas that drew on local assets and were tailored to local needs. By the time the subsequent three sites were funded, P–20 asked sites to propose activities in each of the seven focus areas shown in Table 4.1 (the seventh focus area, Data, was assumed by P–20 when it realized it had more capacity than the sites in that area).

Site teams presented their goals, the activities they planned to meet them, and expected outputs and outcomes of these activities in logic models that they, P–20, and the evaluators used to assess progress. The logic models were created within a template that P–20 developed which was organized by focus area. Under each area, site teams described the activities that the site planned to undertake and the resources that would be used to carry out the activities. Sites described measurable activity outputs and outcomes, and how these were designed to move toward the ultimate goal of improved grade 3 reading. P–20 required that annual progress reports be imbedded in each site's logic model, which meant that each site had to describe progress in each of the six focus areas. (See discussion above on logic model development and monitoring by P–20 and

Zellman and Kilburn, 2011, for further discussion of the process of logic model development in the first two demonstration sites.)

P–20's decision to give site teams autonomy to select activities they believed would promote a consensual outcome—improved grade 3 reading scores—is consistent with a trend in public and nonprofit management toward orienting work toward outcomes rather than specifying activities (Chinman, Imm, and Wandersman, 2004; Friedman, 2005). The contracts with the demonstration sites specified that the sites aim to improve third-grade reading scores by identifying activities in the six focus areas that represented the components of the initiative's P–3 approach. Despite the general shift away from the idea of activities-based contracting in education and social service work, contracts have tended to continue to dictate activities to be implemented. P–20's approach to contracting with the P–3 sites is uncommon.

This objectives-based contracting strategy meant that P–3 avoided the common problem of trying to roll out a "one size fits all" approach in sites that differ in fundamental ways. However, any strong local autonomy policy runs the risk that sites will continue doing "business as usual."

The optimal level of local autonomy to afford demonstration sites was discussed from the beginning of the RAND evaluation. Over time, driven in part by changes in P–20 staffing, P–20 moved toward the imposition of more requirements on the original sites and similar requirements for the newer sites as a condition of inclusion in the initiative. Much of the standardization discussion focused on the use of classroom assessments. P–20 came to require all sites to implement assessments as a tool to improve instruction. Further, P–20 saw efficiencies in encouraging sites to use the *same* tool: Trainings and trainers could be shared, and P–20 would have an easy way to compare sites' progress in terms of number of teachers oriented, number of classrooms assessed, and number of individual coaching sessions held. Eventually, P–20 required all sites to use CLASS.[6] However, since RTTT required that elementary schools in the two complex areas that were receiving RTTT funds use a different instrument, P–20 revised its policy and required only that "*a valid and reliable… assessment tool*" be used.

---

[6] The original sites were aware of this shift, but were unclear if these agreements applied to them as well at first. All the sites wondered whether P–20's movement away from local autonomy around assessment would have implications for other aspects of their work.

P–20 also expected and encouraged the sites to make good on the "P–3" aspect of their work by including teachers of grades 1–3 in their efforts. P–20 staff was clear that the initiative was not a "PreK to kindergarten" or "PreK to grade 1" initiative and expressed concern that much of the early work in the two original sites was focused on PreK to kindergarten. P–20 experienced some pushback: Several sites indicated that they felt it was more appropriate to work first with kindergarten (and sometimes grade 1) teachers and then move upward once some level of success had been achieved. Ultimately, few teachers in grades 2 and 3 participated in P–3 activities, although many professional development and other activities were open to them.

P–20 monitoring responsibilities vis-à-vis the demonstration sites included requesting and receiving information concerning activities and outputs from the sites on a regular basis through the consolidated reporting tool, which combined site logic models and progress reports. Staff evaluated this information and provided sites with feedback and technical assistance when appropriate. P–20 continued to take an active role vis-à-vis the sites through 2013, but began to pull back in the final funding year, reportedly to encourage sites to work on sustainability activities and to consolidate gains in their final year of funding.

In all sites, the P–3 team became an important convener of stakeholders. In some sites, these meetings brought together people who had not worked together before, despite their compatible work and shared goals. In all sites, P–3 team members took responsibility for moving forward the agenda and activities that had been agreed to. Often this meant working with local stakeholders and meeting with HIDOE, which was a major actor in every site and was particularly important in the rural sites with limited early childhood infrastructure.

Establishing Incentives and Appropriate Consequences for Meeting (or Failing to Meet) Goals and Expectations. Although it had potential power to do so, P–20 articulated few consequences or rewards related to demonstration site performance. P–20 staff assumed that site teams were intrinsically motivated to perform and use their WKKF funds in the best possible way.

P–20 did implement some requirements in the area of process performance, such as requiring sites to complete reports and other contractual requirements before they were paid. Additionally, P–20 made the use of a classroom observation tool a condition of funding. However, there were no consequences attached to failing to do this; sites that

did not comply offered a range of reasons that P–20 accepted. P–20 staff worked with the sites to support implementation of the requirement or to modify it. One requirement that was imposed was the completion by each site of an Issue Brief, which P–20 considered an important tool for increasing the visibility of the work and promoting sustainability by informing policymakers and potential funders of site activities. No site complied, citing lack of time and resources. P–20 offered support for the writing of these briefs, and P–20 staff even offered to write them for the sites using site input if, as claimed, the sites lacked time and resources to do so. No site took advantage of this offer. Nor was P–20 able to get sites to participate in a public presentation of the work when the site work was ending. No consequences were imposed on the sites for their lack of cooperation even though the lack of dissemination efforts might have had consequences for the work; the demonstration site work was not well known, and stakeholders had difficulties articulating P–20's role in many of the changes that had occurred in the early childhood sector in the state.

As denoted by its title, the P–3 initiative was to focus on PreK through grade 3. Fairly early on, P–20 noted that few sites were including even grade 1 in their activities, let alone grades 2 and 3. P–20 raised this issue; sites explained that their work was path-breaking in many cases and that their efforts needed to focus on the PreK to kindergarten connection before they could engage the higher grades. P–20 accepted this feedback and imposed no sanctions when sites continued until the end of the grant period to focus on PreK to kindergarten, with some grade 1 inclusion.

Sites also relied heavily on intrinsic motivation rather than incentives to push their work forward. This made sense given the far lower levels of resources that the sites had, compared to those of P–20. However, funds were used selectively in a few sites to encourage teachers to participate in professional development activities. The provision of substitute teachers to cover for teachers who agreed to participate in P–3 activities was a widely used and important participation incentive. Food was a frequent incentive as well; a number of site team members described the Hawaiʻian culture as hospitable and food-oriented. Providing lunch to workshop participants sent an important, welcoming message. In one site, gift cards were provided to teachers who participated in PPVT training and who administered their own assessments to their students. Site team members were confident that once teachers used the PPVT, they would feel empowered and communicate its value to other colleagues.

Site teams also tried to emphasize the value of trainings. As site team members noted, training can easily become "one more thing" that busy teachers have to do that does not directly help them in their work. The team made sure that that training aligned with teachers' and children's identified needs; this increased the perceived value of many of the available training opportunities.

More often, site team members tried to build intrinsic motivation in particular actors they wanted to work with. In one site, P–3 team members spent considerable time with the CAS, helping her to understand the value of a PreK–3 perspective. She became a very strong advocate for the work. In the RTTT sites, elementary school teachers were disinclined to get involved with the CLASS, as they faced a Danielson mandate and generally felt overwhelmed. In one of these sites, P–3 team members made an effort to find just a handful of teachers who were willing to try the CLASS. Those teachers found it both helpful and different enough from the Danielson that they became advocates for the CLASS with their peers.

Transparency of Monitoring and Evaluating the Performance of Key System Actors. To evaluate success, it is necessary to have in place mechanisms for monitoring performance. The development of sophisticated and detailed logic models in each demonstration site that included required activity categories provided P–20 with a tool to facilitate performance monitoring.

Through their logic models, the sites identified measures that they believed reflected the core objectives of their work, were fair, available, and could be collected at a reasonable cost. (See Zellman and Kilburn, 2011, for further discussion of the process of logic model development in the first two demonstration sites.) Data collection for these measures began during year 2 of the evaluation for the first two sites, with continuing collection in each succeeding year. Data collection for the three new sites began during these sites' first P–3 year. As noted above, P–20 monitoring efforts increased over time. In 2012, a staff member was hired with responsibility for monitoring the sites. She actively examined the indicators that each site had developed to assess its progress, relying on both reports from the sites and the HIDOE database (e.g., for Hawai'i State School Readiness Assessment [HSSRA] and HSA data). At the same time, once questions were raised or changes requested, P–20 follow-through was limited, and sites generally continued to do what they had been doing.

Demonstration sites increased their monitoring of their activities in response to P–20 requests and the need to document progress. In one site, for example, in response to P–20's expectations of measurable outcomes, the site began to monitor a number of activities, including the uptake of learning materials by parents, number of website visitors, and attendance at parent groups.

Ensuring That Key Actors and Entities Have the Capacity to Carry Out Their Respective Responsibilities. P–20 spent considerable time and resources to increase capacity among those working on the P–3 initiative. As discussed below, P–20 supported travel to national P–3 conferences and workshops for a number of stakeholders and HIDOE staff. P–20 staff provided professional development and technical assistance to the site. P–20 staff also convened regularly scheduled P–3 site coordinator meetings during which site coordinators could exchange ideas and discuss progress and issues with P–20 staff and with each other.

Some demonstration sites focused their P–3 activities on capacity building; all performed work in this area. The focus on professional development led to reported increases in staff capacity as well as documented improvements in CLASS scores in some sites. Teachers who participated in trainings are reportedly more comfortable with being assessed and have come to recognize the value of observational data in improving instruction. Many now are using data to identify their students' learning gaps and differentiate their instruction accordingly. Use of assessment tools reportedly helped PreK teachers to better support the development of their students' skills. Visits across PreK and kindergarten provided teachers at both levels an opportunity to learn about the expectations that their students face and enabled teachers to provide feedback to their colleagues at the other level. According to P–3 team members in one site, these visits helped elementary school staff to better understand and incorporate a developmental perspective in their work.

Summary and Conclusions

The requirement that PreK and HIDOE leaders jointly commit to the P–3 work in the sites set clear expectations that there would be cooperation between these key groups in the work. Site plans were required to include work in all six focus areas developed by P–20, and all site plans complied with this requirement. Since these focus areas

represent the cornerstones of P–3 theory, site plans directed site teams to do the P–3 work.

Sites developed plans that reflected local needs and skills. Site teams and their partner organizations worked diligently and strategically to implement the activities in their plans and to build relationships among key local stakeholders through this work.

Much of this work appeared to change attitudes and norms, and it certainly brought people together. Some of the activities produced system change at the local level, but the continuity of many of these changes was dependent on a decision by a single individual (usually a CAS) to institutionalize it; CAS tenure—and policies--can end at any time. CASs in more than one site publicly embraced a P–3 focus. One CAS required all of the elementary schools in the complex to develop kindergarten transition plans. Another CAS created a place on the monthly principal meeting agenda for P–3. Another began to talk to her principals about the value of early childhood background in K–3 teachers. Staff of PCNCs agreed to pick up some of the parent work of the P–3 site team in one site. But much of this change depends on incumbency: When CASs and principals change jobs, system changes dependent on CAS or principal support might not endure. Or they might flourish and expand: We saw both in the demonstration sites over time.

Site teams believe that the new norms and expectations created through the P–3 work will promote continuity of support for P–3 goals and P–3 work. Several site team members noted, for example, that the CLASS work had educated teachers about the value of observation, data, and coaching. They believe that this new understanding will motivate continued efforts to find and use assessments, and this in fact seemed to be happening in one of the sites. But some work, particularly CLASS, cannot easily be continued without outside funding support. The CLASS work is an area where better sustainability planning might have led P–20 and sites to make different choices about how to use limited resources.

There was little institutionalization of the activities we observed in the sites. This reflects the many challenges involved in promoting lasting change in complex systems, and particularly when funding is relatively short-term. The lack of institutionalization also reflects the reality that a number of sites did not actively seek this sort of change, nor did P–20 push them to consider or seek to sustain their work, as discussed in Chapter 5. Site staff were grateful for the new contacts made in the course of the P–3

work and the changes in attitudes they witnessed. Many believe that continued cooperation around shared norms and goals will continue to further the work. One notable exception was a site that was able to institutionalize its work with the help of P–20: HIDOE accepted the 15 units of early childhood coursework it developed as eligible for PDE3 funding. This means that HIDOE teachers who successfully complete the five courses can get salary points for doing so. This will undoubtedly increase the number of K–3 teachers with some early childhood background. Putting the courses online, which P–20 helped to do, has made these courses available to teachers across the state.

More enduring system change is more likely to occur at a higher level of the system, which has more power to change policy. In Chapter 5, we examine system change at the P–20 level.

School Information Form Results

As described in Chapter 3 on methods, principals in participating demonstration sites provided information about their schools' engagement in P–3 via an online form. In 2011, principals from the original two sites completed the form; in 2013, principals from all five demonstration sites completed the form; and in 2014, principals from four of the five demonstration sites participated. We were not able to distribute the form in 2012 because a change in HIDOE research review procedures slowed approval beyond a reasonable data collection window.

Given the small number of demonstration sites and the small number of schools in some sites, formal statistical tests of change over time would be unlikely to identify significant changes if they were present. This small-sample challenge is exacerbated by the fact that one of the demonstration sites could not participate in the final administration. Consequently, we focus on evidence related to three issues:

- Are principals aware of P–3 demonstration site work in their complex area?
- Are schools implementing P–3 strategies?
- Do principals have a positive attitude toward P–3 work in their complex area?

We first present results related to these questions for the final year of the initiative, and then compare these results to earlier years to assess the degree to which there appear to have been improvements over time.

We assessed the degree to which principals were aware of P–3 work in their complex area by this question: "How much do you know about the Hawai'i P–3 initiative?" Principals could reply on a three-point scale: —"Never heard of it," "Have heard something about it," or "Know a lot about it." Across all waves of SIF administration, every principal had heard of the P–3 initiative. By the final year of administration, 60 percent of principals reported that they "Know a lot about it," with the remainder replying that they "Have heard something about it." We examined whether principal tenure was related to principals' familiarity with P–3 and found that principals who had been at the school for more than three years were more likely to know a lot about the initiative. We also asked principals whether they had attended meetings where the P–3 work in their demonstration site was discussed. These results varied substantially by demonstration site: In one site, 0 percent of principals reported attending such meetings, whereas nearly all principals in another site had attended (95 percent). Attendance at meetings where P–3 was discussed was intermediate in the other two sites: 20 percent and 78 percent of principals in those sites attended such meetings. School leadership support for P–3 activities is an important contributor to the success of the demonstration site work, and the results from the 2014 SIF show a high degree of awareness among principals but varying levels of engagement with P–3 as measured by meeting attendance.

In addition to asking principals to report their awareness of P–3 activities in their demonstration sites, we asked them to assess, at the end of the P–3 funding period, "How much of each of the following has P–3 provided to you, your school, or your teachers?" Their responses are shown in Figure 4.1. More than half the principals indicate that the P–3 initiative provided "A lot" or "Some" support for activities that are features of most P–3 initiatives. These include support for early childhood–elementary school partnerships, information about best practices, support for the adoption of quality instructional tools, and data sharing.

Figure 4.1—Principal Reports on the Extent to Which P–3 Provided Various Types of Support to Their School (2014 SIF)
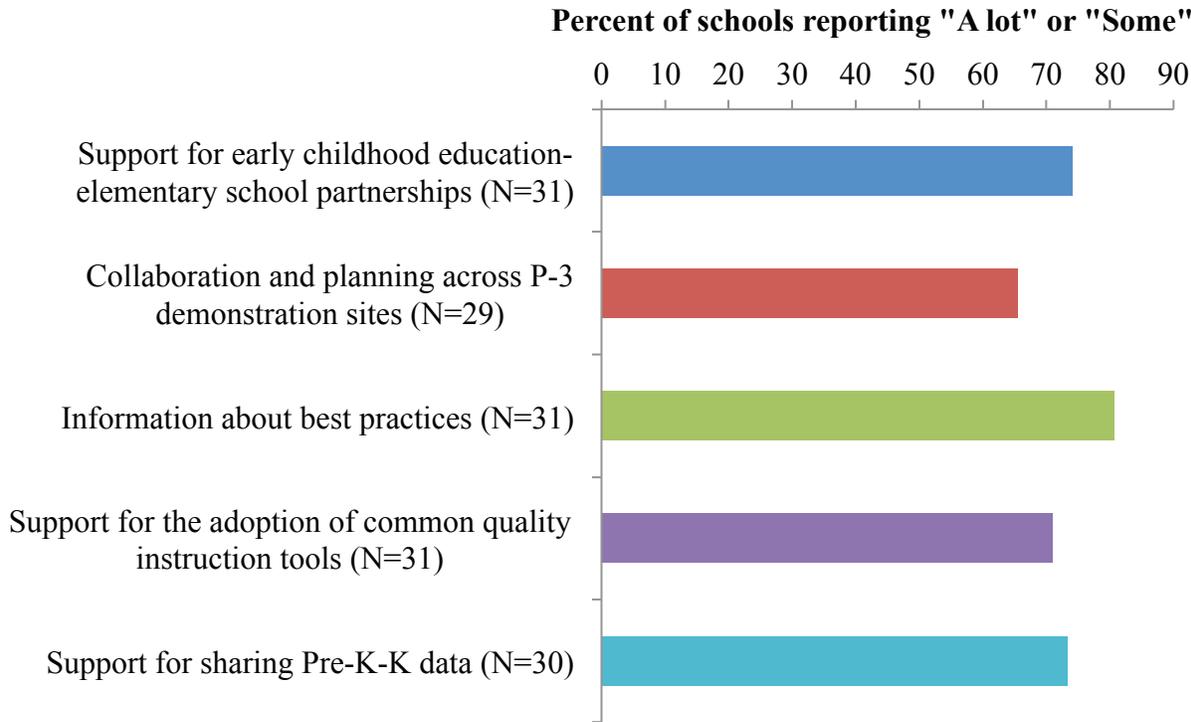
We also asked principals about their attitudes related to the value of P–3 support by asking them to "Please indicate your agreement with each statement in the table below by selecting the option in the column that comes closest to your view." As shown in Figure 4.2, responding principals generally had a favorable view of P–3 in such areas as improving literacy among K–3 students, being a good use of time, helping staff become sensitive to developmental issues, and not interfering with other school activities.

Figure 4.2—Principal Attitudes Related to the Value of P–3 Support (2014 SIF)
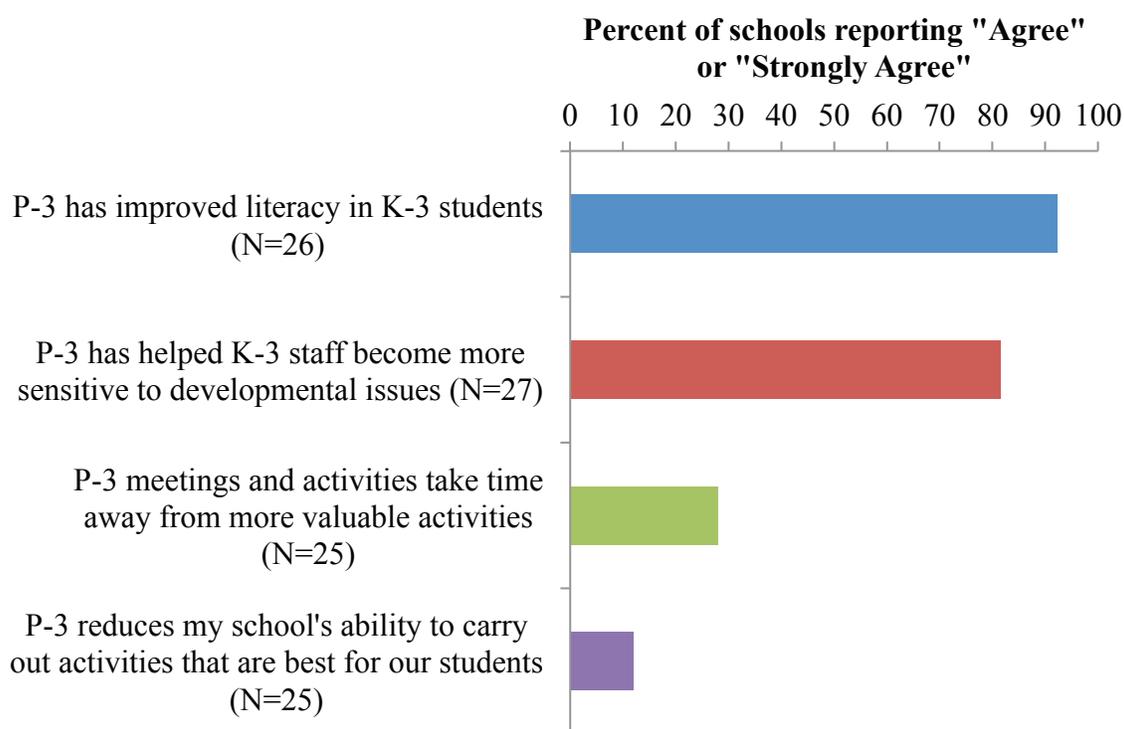
One of the central P–3 strategies is to encourage engagement in professional development and certification related to early childhood education. To address this aspect of the demonstration site work, we collected information from principals about their K–3 teachers, including their attendance at the annual Hawaiʻi Association for Education of Young Children (HAEYC) conferences, whether they participated in University of Hawaiʻi coursework or other professional development programs, and whether they had an early childhood certification. Thirty-four principals responded to these questions in 2014; 21 principals reported that teachers from their schools had attended HAEYC in the previous year, and 18 principals reported that teachers in their schools had participated in coursework or professional development programs. Nineteen principals reported that some K–3 teachers in their schools had early childhood certification. As a whole, the levels reported in 2013 were very similar. These results indicate a relatively high level of teacher engagement in this P–3 strategy, with similar participation levels in the last two years of the SIF administration.

We also asked principals, "Did any of the practices listed below occur at your school during the previous school year (2013–14)?" Examples of these 13 practices include the use of teacher assessments to improve instructional practices, obtaining child development training, vertical alignment meetings across grade levels, and teachers in PreK and elementary school observing each other's classrooms. Figure 4.3 shows results from the 34 principals who responded about the 13 practices in 2014. The majority of schools report engaging in at least half of these practices. It is noteworthy that the practice that is most often undertaken—Danielson assessments—is the one that is required by HIDOE, whereas the others are generally optional. The results are similar to those in the 2013 administration.

Figure 4.3—Principal Reports of P–3 Practices Taking Place at Their School (2014 SIF)
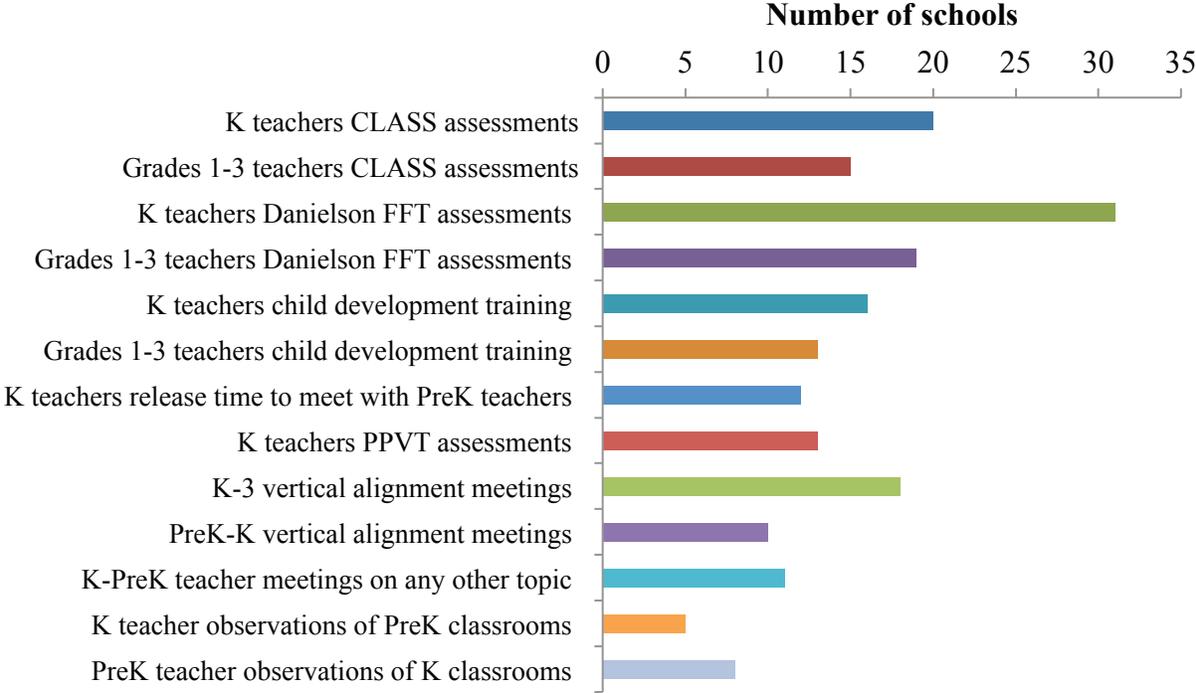
Another practice that is an often-cited feature of P–3 initiatives is transition programs for young children as they begin kindergarten. Principals provided information on transition programs at their schools. Out of 33 principals reporting this information in 2014, 30 indicate that they offer an orientation program for incoming

students. The next most common transition program is an open house for incoming students, which 27 schools offered. Less than half of the schools also have a summer transition program for incoming students. The provision of summer transition programs in 2014 also is not much different from that reported in 2013.

Another key objective of P–3 initiatives is increasing access to PreK. We asked principals, "About what fraction of kindergarten students who entered your school in fall 2014 had had some PreK experience?" These results are reported in Figure 4.4. This bell-shaped graph shows that the most likely percentage was between 41 and 60. The figure for 2013 is very similar.

Figure 4.4—Principal Reports of the Fraction of Kindergartners with Preschool Experience (2014 SIF)
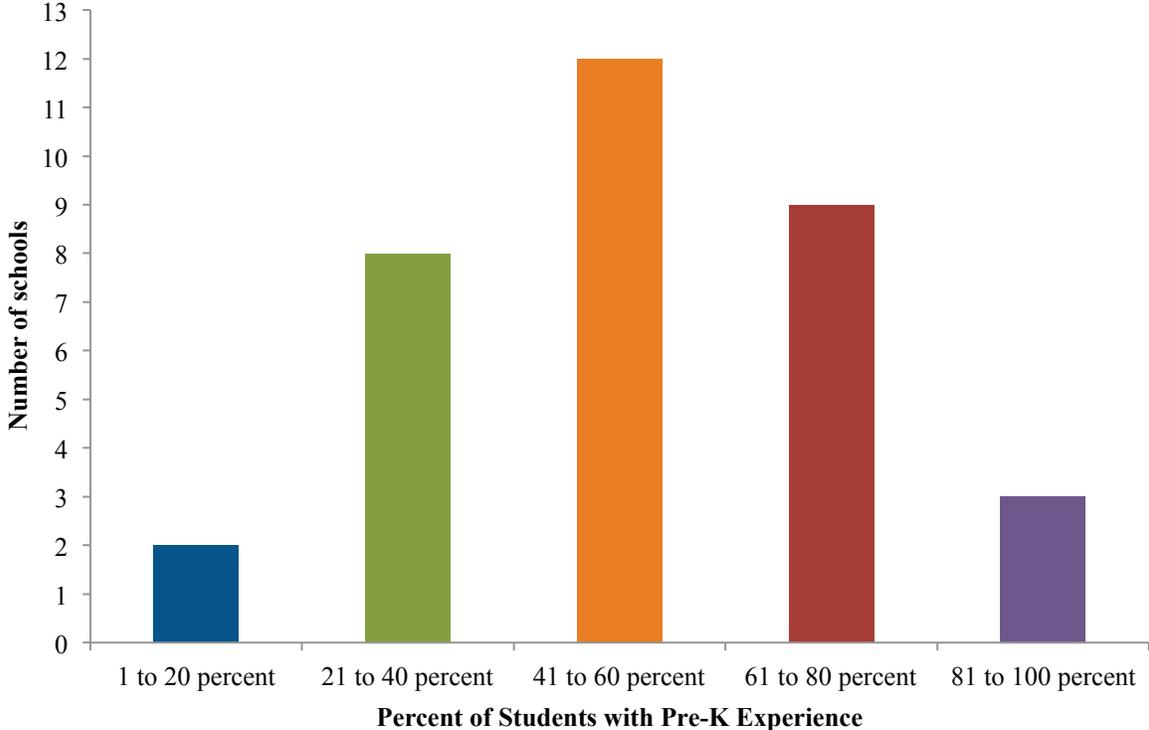
In general, the results from the 2014 SIF indicate high levels of awareness among principals and favorable attitudes toward the P–3 initiative in their complex areas. The SIF results also demonstrate relatively high levels of P–3 activities being undertaken at the schools. However, we did not observe measurable increases in these activities

between 2013 and 2014 for the four demonstration sites with data from both of these years. Similarly, for the two demonstration sites that also have data from 2011, we do not observe meaningful increases over that longer time frame. Furthermore, based on discussions with P–20 staff and other stakeholders about elementary schools they are familiar with that are not located in a demonstration site, it appears that many of the P–3 activities are occurring in schools outside demonstration site complexes. Hence, although we can conclude that there is a large amount of adherence to P–3 principles in the demonstration site schools, we are not necessarily able to attribute this to the P–3 initiative work in the demonstration sites.

Third-Grade Reading-Score Findings

The ultimate goal of Hawai'i's P–3 initiative was to ensure that all children read at grade level in third grade. The outcome analysis for this evaluation examines whether children exposed to P–3 in demonstration sites have a greater likelihood of reaching this goal than students whose complex area did not participate in P–3.

The outcome examined in this portion of the analysis is third-grade reading achievement scores collected by HIDOE. We compare seven years of reading scores across time, starting with spring 2008 and ending in spring 2014. While students in demonstration sites will have been exposed to P–3 activities for as long as three years by the time they reach third grade in 2013, students who reach third grade prior to 2013 will have been exposed to P–3 activities for only their PreK and part of their early elementary years. Since fall 2013 is the first year in which a cohort exposed to the complete set of P–3 activities will reach third grade, our analysis is undertaken at a time when only a fraction of students will have had full exposure to P–3 activities and therefore could benefit from their full effects. Moreover, the literature on implementation of education reform makes clear that it takes time to fully implement reforms, and even when fully implemented, it takes time for them to realize their full impact (Gill et al, 2005). For these reasons, the full effects of P–3 activities on student outcomes will probably not be realized until after our study has been completed, and our analysis is likely to capture only partial effects of the P–3 initiative.

Sometimes administrative data contain errors, such as missing values, incomplete samples, or other quality problems. Before analyzing these data, we undertook checks to ensure that the data are of high quality—i.e., that variable values were within range,

the number of schools and students was correct, and that there were not a large number of missing values for any variables. This exercise convinced us that these data were of very high quality.

The approach we use to analyze student outcomes analysis is a student-level difference-in-difference model for students at all Hawai'i elementary schools. This analysis compares changes over time in student third-grade reading scores within demonstration site elementary schools to changes in those test scores for students in other elementary schools (see Cameron and Trivedi, 2005, for discussion of difference-in-difference estimates). These models include school-level statistical controls for other time-varying factors that might influence outcomes, such as the fraction of students qualifying for free or reduced-price lunch. This approach enables us to determine whether schools participating in the P–3 demonstration projects had different student outcome growth patterns than other schools. In sum, using student-level data for third-grade reading for all students in the state, we test the degree to which the students in schools participating in the P–3 demonstration projects realize greater gains in scores over time compared to students at schools not participating in the demonstration sites, holding other characteristics constant.

We analyze student-level third-grade reading scores as measured by the HSA third-grade reading score. The additional covariates that we include in the model are the student's gender, race, and ethnicity (Asian, Pacific Islander, white, other); birthdate (year and month); whether the student is FRPL eligible; whether the student is designated as an English-language learner; whether the student repeated the third grade; whether the student was receiving special education services; a school-year indicator variable; a school indicator variable; the school rate of FRPL eligibility; and the school percentage of Pacific Islanders. Note that we adjust the standard errors of these estimates to account for the fact that students are "clustered" in schools, and hence test scores of students from the same schools should not be treated as independent. If we did not make this adjustment, we would overstate the true amount of variation in the data and calculate confidence intervals that were too small (Shadish, Cook, and Campbell, 2002).

There are approximately 192 elementary schools in the state. In each year of data, between 13,000 and 15,000 students have HSA scores, with the number growing during the period represented in these data. We analyzed seven years of third-grade reading-

score data, which include every student in the state who took the HSA in each of the covered years, totaling 98,909 students. Two schools began P–3 demonstration site activities in 2009, two more were added in 2010, and the last site commenced in 2011. As shown in Table 4.2, sites were first included in the test-score analysis the school year after P–3 work began in the site. The table also shows the number of HSA scores available for students in P–3 demonstration sites by year and the total number of HSA scores in the district by year.

Table 4.2.

Number of HSA Scores from P–3 Sites and Total, by Year

School Year	Number of HSA Scores from P–3 Sites	Total Number of HSA Scores	Number of P–3 Sites
2007–08	0	13,223	0
2008–09	0	13,727	0
2009–10	1,217	13,128	2
2010–11	2,714	14,255	4
2011–12	3,090	14,718	5
2012–13	2,603	14,758	5
2013–14	2,712	15,100	5
Total	12,336	98,909	5

Our analyses tested two hypotheses:

1. Reading scores will grow more for students exposed to P–3 than for students who attended schools in non-demonstration sites

2. Reading-score growth will rise with the number of years of exposure to P–3.

The ability to draw firm conclusions about the effects of site-level P–3 activities on reading-test scores in the time frame of this project is constrained by several factors. Although the analysis includes all schools and all five demonstration sites, we recognize that the analysis will be more likely to capture the effects of P–3 activities in the original sites, since they began earlier and students attending school in those sites were exposed

to P–3 for a longer period. A number of other factors also limit our ability to draw firm conclusions from these analyses. Students in the more recent cohorts might have received even less exposure than timing alone suggests, since many of the activities in the P–3 sites were focused on PreK and kindergarten and the students in the more recent cohorts would have been in first through third grade during the P–3 intervention at their schools. Moreover, it is important to note that in the P–3 demonstration sites, other educational initiatives were being undertaken at the same time as the P–3 initiative, so effects cannot be attributed solely to the P–3 activities.

Findings

P–3 demonstration site schools were found to have lower mean HSA scores in 2008 than comparison schools that never participated in P–3. Since in 2008, no demonstration site had begun its work,[7] this difference indicates that the P–3 initiative was targeted to lower-performing schools. In 2008, the mean HSA score for students that never were part of a P–3 site was 307, whereas the mean HSA score for students in schools located in complexes that later became P–3 demonstration sites was 298, as shown in Figure 4.5. By 2014, this gap had narrowed slightly; the difference in mean HSA scores between schools that had and had not participated in P–3 had been reduced to a five-point difference. Mean scores for schools that had no P–3 involvement rose to 315, whereas the mean HSA score for students who had been part of a P–3 site rose to 310. In sum, over the seven years covered by our data, the gap in scores between schools that never participated in P–3 and those that did narrowed by four points, from nine to five points.

---

[7] We do not consider the small P–3 grants received by the first two sites prior to their being P–3 demonstration sites to be large enough to be significant in these analyses. The early grants were very small, and those grants included a planning year when activity implementation was neither expected nor planned.

Figure 4.5—Mean HSA Scores for Non-P–3 Schools and P–3 Schools, by Year
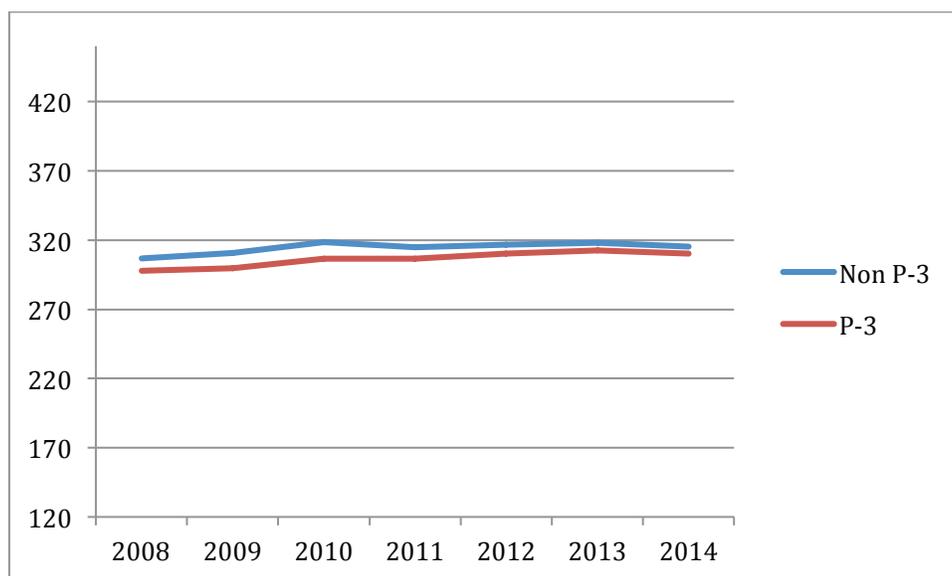
The results from estimates of the multivariate models, which include the control variables outlined above, largely echo this graph. We find that being in a P–3 site is unrelated to the HSA score, after controlling for the other variables. However, in the model that allows the effect of being in a P–3 site to vary by the amount of time the student has been in the P–3 site, we find that each year of being exposed to the P–3 initiative is associated with a 0.7-point increase per year of P–3 exposure. This latter estimate implies that being exposed to P–3 for five years is associated on average with a 3.5-point gain on the HSA, which is equivalent to an effect size of about 0.1. (This estimate is statistically significant at the 0.032 level.) This gain is equivalent to an effect size of about 0.1. Note that a 3.5-point gain is very close to the four-point narrowing of the gap described above for comparisons with no controls. The complete set of multivariate results is provided in Appendix C.

Another measure of reading performance is reading proficiency, which is considered to be present when a student achieves a scale score greater than 300 on the HSA reading assessment. This measure most closely approximates the overarching goal of P–3, which was to have all children reading at grade level by third grade. In 2008, 64 percent of students in schools that were never part of a P–3 demonstration site were proficient in reading, whereas 53 percent of students in schools that eventually were part of P–3 sites were proficient, as shown in Figure 4.6. This 11-point difference

again underscores the fact that the P–3 initiative targeted complex areas with lower-performing schools. By 2014, the percentage of reading proficiency in schools that did not participate in P–3 had grown from 64 to 67 percent. The percentage of students proficient in reading in the P–3 sites still lagged behind other schools in 2014, but the gap had narrowed; the percentage of reading proficiency in those schools grew from 53 percent to 60 percent, leaving a seven-point gap.

Figure 4.6—Percentage of Students Proficient in Reading for Non-P–3 Schools and P–3 Schools, by Year
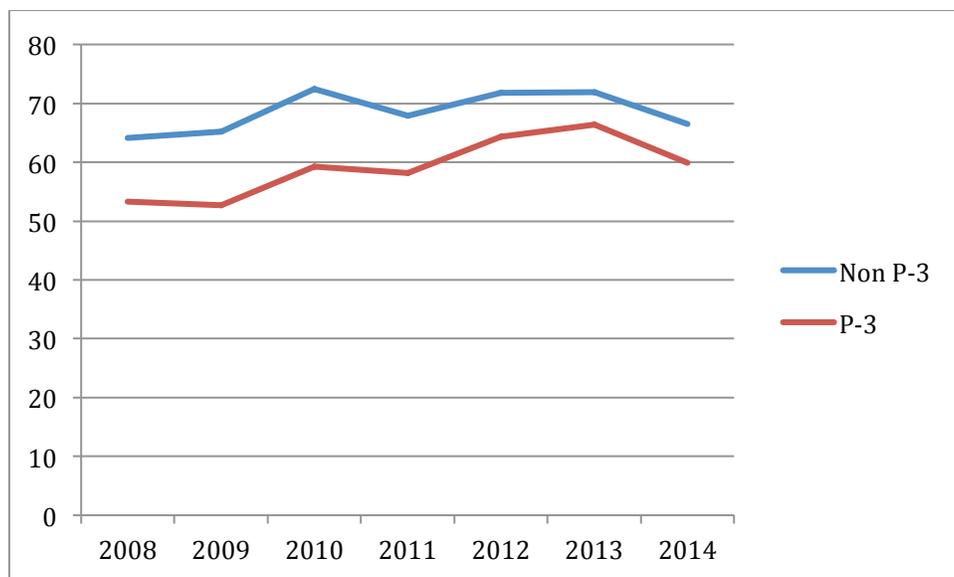
The multivariate estimates, which include the same set of control variables as the scale score model above, show that simply being in a P–3 site is not associated with a greater likelihood of being proficient relative to being in schools that were never part of P–3 sites, controlling for the same factors as in the scale score estimates. However, the model that accounts for the number of years of participation in the P–3 initiative indicates that each year of participation is associated with a 4-percent higher chance of scoring in the proficient range. This latter estimate is significant at the 0.049 level.

While the size of the reading score and proficiency gains made by students in the P–3 sites might seem modest, the size of this effect compares favorably with the sizes of effects of many educational interventions as measured by scores on general tests like the HSA. For example, Hill et al. (2008) estimate an average effect size of 0.07 for a

sample of randomized trial evaluations of educational interventions for mainstreamed students in elementary school. This figure is also comparable to estimates of the effects of nine additional weeks of schooling for elementary school students (Chingos, Whitehurst, and Gallaher, 2013). Furthermore, these findings could be considered encouraging given the small number of years that the majority of P–3 demonstration sites had been under way; the low dosage most students in these data experienced; and because the demonstration sites emphasized PreK to kindergarten in their work, and more students in the data would have been exposed to P–3 in the higher grades (grades 1–3), which emphasized P–3 less than in the lower grades. Education research suggests that many interventions take three or more years to be fully implemented (Fixsen et al., 2005), so only the original demonstration sites in this initiative might be considered to be beyond the initial implementation stage. The students in the original two demonstration sites would have received the most years of exposure to P–3 activities, but this would have been about five years at most. In sum, despite the low dosage of the P–3 intervention experienced by most treated students in these data, we find evidence that more years of participating in the P–3 initiative raised student reading scores and increased the likelihood that they scored proficient on the state reading test.

5. P–20 State-Level Work

In this chapter, we review and analyze the state-level work of P–20 and examine it from a system change perspective.. The state-level work is most easily defined as all of the activities in which P–20 staff engaged that did not involve support for the five demonstration sites (discussed in Chapter 4). We address the following research questions:

1. Which strategies did the P–3 initiative employ in its state-level work and did these align with best practices in P–3 at the time?
2. Did P–20 successfully implement its strategies and plans?
3. Did P–20 develop plans likely to promote long-term effects or sustainability of its P–3 efforts?
4. Did P–20, in its state-level work, engage in activities and processes that promote system change? Did system change occur?

Summary Description of P–20 Work

Which strategies did the P–3 initiative employ in its state-level work and did these align with best practices in P–3 at the time?

Like the work of the demonstration sites, P–20's state-level work was guided by a logic model that was similar in form to that used by the demonstration sites, in that it defined resources, activities, outputs, and outcomes. However, the focus areas included in the P–20 logic model were different than those for the sites, and there were fewer of them. The P–20 logic model included four focus areas:

1. Engaging P–3 community teams
2. Evaluating and chronicling successful community strategies
3. Aligning and informing statewide efforts (curriculum, standards, culture-based education, and longitudinal data collection)
4. Increasing teacher and leadership capacity.

These focus areas are consistent with strategies that have come to be understood as building blocks of successful P–3 planning and approaches (see, e.g., Kaurez and

Coffman, 2013). The first two focus areas emphasize support for the work of the demonstration sites, whose logic models focus on P–3 best practices, as discussed in Chapter 4. The latter two focus areas directly address two cornerstones of P–3 frameworks: alignment and increased capacity in the P–3 workforce. Within the focus areas, P–20 engaged in activities that address the areas identified by Kaurez and Coffman (2013). Below we discuss P–20's work in each area and the alignment of these areas with P–3 best practices.

Engaging P–3 Community Teams. P–20 adopted several strategies to support and promote demonstration site work, which would contribute to initiative goals. Most of these strategies were implemented after new P–20 leadership arrived in 2011. As discussed in Chapter 4, in the early years, the first sites were given a good deal of autonomy and were subjected to minimal oversight, although they were expected to carry out and report on the work proposed in site logic models. Over time, P–20 began to convene site coordinator meetings and more actively monitor the sites' work. Sites were also provided opportunities to engage in relevant professional development.

Evaluating and Chronicling Successful Community Strategies. As part of this focus area, P–20 sought to identify and disseminate " . . . issues, needs, and promising practices . . . through . . . demo site projects" to inform system building.[1] P–20 sought to do this through identification of best practices, support for and dissemination of issue briefs from each of the demonstration sites that presented a successful practice; continued updating of the P–3 website to reflect current activities; and tracking progress on P–3 demonstration site indicators. P–20 points to practices that were carried out in multiple sites as evidence of replication, such as cross-visitation of kindergarten and preschool teachers to each other's classrooms, CLASS coaching, leadership symposia, and book bag rotation projects. However, most of these practices, and particularly CLASS, which was required, were so fundamental to the sites' work that it is hard to think of them as representing successful instances of replication.

P–20 represented Hawai'i P–3 at state-level ECE and other meetings and at national conferences and workshops. P–20 also brought national P–3 experts to Hawai'i to provide guidance to demonstration site teams.

---

[1] See Hawai'i P–3 Initiative, undated.

Aligning and Informing Statewide Efforts (Curriculum, Standards, Culture-based Education, and Longitudinal Data Collection). While the bulk of Hawai'i P–3 initiative funds were used to support the work of the local demonstration sites, the P–20 initiative also undertook a number of activities designed to influence the statewide P–3 context in hopes of effecting lasting changes that would affect other local communities in Hawai'i and the state as a whole. These activities, which were laid out in P–20's logic model, developed early childhood infrastructure and promoted policies that supported P–3 goals, such as more PreK enrollments; increased early childhood capacity; brought best practices into standard early childhood practice, for example, through training on and implementing CLASS assessments in PreK and K classrooms; and increased alignment through the HELDS and the longitudinal database work. P–20 also worked to increase early childhood capacity by encouraging the development of a set of five early childhood courses that would qualify for salary points (that is, a raise) for HIDOE teachers, and through the establishment of and access to a PreK–3 Graduate Certificate program in early childhood. More generally, significant efforts were made to promote a greater understanding and acceptance of the value of early education and to facilitate continuing investments in the early childhood sector. Unlike the work of the demonstration sites, the state-level P–20 work continues because WKKF provided P–20 with a three-year no-cost extension of the grant during which time P–20 will spend out the remaining WKKF grant funds.

When the P–3 initiative began there was scant state infrastructure for early childhood; government involvement in early childhood was housed in the Department of Human Services and focused on facilitating parental employment. From the beginning, P–20 state-level activities tried to build support for early childhood education and a culture in the state in which PreK was understood to be a valuable part of a child's education. These goals paralleled those of the demonstration sites, as discussed in Chapter 4.

P–20 itself engaged in a variety of efforts to promote a culture of support for early education. One example was travel support offered to HIDOE and other stakeholders, including legislators, for trips to national meetings that attempted to convince education policymakers to increase support for ECE, including Harvard's PreK–3 Institute and WKKF-sponsored Learning Lab meetings. Several stakeholders later told RAND that they believed these experiences were transformative for them and other

decisionmakers: more than one HIDOE leader reported that the Harvard meeting led them to redirect the focus of their work to connecting to PreK providers and bringing early childhood knowledge to their K–3 teaching staff.

Once the Executive Office on Early Learning was established in June, 2012, P–20 coordinated closely with the office on a number of initiatives to promote a culture of support for early education. The two entities, along with HIDOE, sponsored a Governor's Symposium on early learning in January 2014, supported by the National Governor's Association Policy Academy. P–20 has worked closely with a range of organizations to help broaden community understanding about the importance of P–3 alignment and its potential short- and long-term benefits as well as other P–3 goals. P–20 also strove to build connections between PreK and HIDOE at the policy level. HIDOE has adopted a number of P–3 goals as its own. For example, several of the P–3 indicators described in the original P–3 grant proposal to the WKKF have been included in the DOE's strategic plan, including increasing the numbers of children entering kindergarten with PreK experience, increasing the numbers of children who are kindergarten-ready, and adopting third-grade reading scores as a separate indicator of proficiency.

As P–20 encouraged the demonstration sites to work on sustainability of local initiatives at the community level in the final years of the WKKF-funded work, P–20 provided more hands-on support for a number of statewide efforts that depended on P–3 initiative funds and P–20 staff. These are discussed briefly below.

HELDS. P–20 was actively involved in the creation and implementation of the Hawai'i Early Learning and Development Standards. These research-based standards identify expectations for what children should know and be able to do from birth to kindergarten entry. According to P–20 staff, this work was considered extremely important by the early learning community and HIDOE because the HELDS would promote alignment between PreK and K–3, a central P–3 goal. The standards were to align with HIDOE's K–12 standards, including the Common Core State Standards, the General Learner Outcomes, the Hawai'i Content and Performance Standards III, the Hawai'i Preschool Content Standards, and the Head Start Child Development and Early Learning Framework. In collaboration with EOEL, P–20 staff developed a draft of the HELDS and collected feedback from statewide stakeholders through a series of focus groups. The standards were revised based on this feedback, and a final draft was

68

produced. EOEL/P–20 sought and received endorsement for the HELDS from the governor, the Department of Health, the Early Learning Advisory Board, HIDOE, the Department of Health Services, and other agencies. P–20 staff also met with key stakeholders to identify adoption strategies. A series of RFPs were issued to attract organizations that could develop guidance and support for the HELDS, which would include differentiated guidance for important subgroups such as families and Native Hawai'ian Language medium early childhood providers. P–20 also held a series of HELDS trainings and is currently working on a HELDS training module to be used statewide.

Family Partnership Guidelines. These guidelines were viewed as another way to promote alignment between PreK and HIDOE. P–20 staff partnered with EOEL and other ECE stakeholders to develop the Family Partnership Guidelines for Early Childhood Settings, which align with the proposed HIDOE family engagement standards. Facilitation and write-up of focus group sessions that were part of the guidelines development process were supported with WKKF funding.

Kindergarten Entry Assessment. Growing understanding of the link between kindergarten readiness and third-grade performance led to a HIDOE decision to end the use of a classroom-level kindergarten-readiness assessment, the HSSRA, in favor of the development of an assessment tool that measures kindergarten readiness at the individual student level. P–20 has been involved in the selection and piloting of a student-level kindergarten entry assessment (KEA) tool that aligns with Common Core State Standards for kindergarten. P–20 staff, working with the HIDOE Curriculum Office, identified several KEAs in use by other states and selected *Teaching Strategies GOLD®* (GOLD®) for piloting.[2] Schools throughout the state have signed up to pilot GOLD®, and P–20 is supporting the pilot by providing ongoing training and technical assistance to support the implementation of the tool. P–20 has worked to expand the use of GOLD® as a formative assessment in first-grade classrooms as well. More recently, P–20 staff has begun to work on the development of kindergarten-entry standards which will align with HELDS and K–12 standards. The hope is that this work

---

[2] *Teaching Strategies GOLD®* is a work sampling system that allows teachers to associate results on specific indicators with activities that will help children improve scores on those indicators. GOLD® also has family engagement built in and includes activities that families can undertake that are associated with specific results.

will lead to both state standards for kindergarten entry and eventually an assessment tool.

Common Tools. A major P–20 contribution to early learning has been its focus on alignment through the development and implementation of common assessment tools in PreK and elementary classrooms. The dominant strategy for achieving this goal was to require each demonstration site to include a validated tool in its work. P–20 supported this work by training demonstration (and P–20) staff to be CLASS-reliable, and to monitor demonstration site CLASS implementation. At one point, P–20 was asking all sites to implement the CLASS in PreK and K–3 settings. However, as discussed in Chapter 4, elementary school teachers in the two RTTT sites were obligated to implement the Danielson Framework for Teaching; there was pushback at the elementary school level about the CLASS in these sites because teachers would be observed twice. In one of those sites, P–3 team members worked with volunteer teachers, some of whom came around to valuing the CLASS. P–20 staff accepted the Danielson as meeting the common tool requirement; P–20 staff monitored implementation of it in the RTTT sites through site logic models.

Data Sharing. P–20 encouraged the demonstration sites to develop and execute data sharing agreements between PreK and local elementary schools. When P–20 took back the data focus area, it continued this work itself, providing support and legal expertise as needed. The P–20 partnership has worked to include PreK data in the state's longitudinal data set as another strategy to promote P–3 alignment. A local foundation provided support to map available early childhood data, a critical task as there is no regular established data stream in early childhood. A consultant helped determine what sorts of data the different programs collect; the information was published in a data catalog and shared at a data summit.

P–20 continues to work with early childhood providers on the early childhood information sharing system and integrated data system governance. A coordinating council of organizations that produce or own data (which might eventually include HIDOE as it now operates mainstream PreK classrooms) has convened to develop visions, mission, goals, roles, protocols, policies, and procedures.

Increasing Teacher and Leadership Capacity. From the beginning of the WKKF funding, the P–20 partnership has led or contributed to several efforts designed to develop early childhood capacity in the PreK–3 workforce. The goal has been to

increase the supply of trained and credentialed early childhood teachers and to encourage early elementary school teachers to develop a better understanding of child development. This work includes the development of undergraduate and graduate-level opportunities as well as efforts to make these opportunities more accessible by offering them online.

P–20, through its work with the demonstration sites, supported the development of five Professional Development Experiences that Educate and Empower (PDE3) courses in early childhood subjects that were developed by one of the demonstration site coordinators and her colleagues. Originally, these courses were available only locally. Through its work with the National Governors' Association Policy Academy, P–20 made ECE courses available statewide. By making these courses eligible for professional development credits for HIDOE staff, the likelihood that K–3 teachers will enroll in these courses in much greater.

P–20 staff devoted considerable time to developing and overseeing the PreK–3 Graduate Certificate program.[3] The program was developed collaboratively by the Department of Curriculum Studies and the Department of Educational Psychology faculty with P–20 staff. Students in the program received tuition assistance from P–3 funds. Originally, the certificate program was available only at University of Hawai'i Mānoa, which effectively barred those living on neighbor islands from participating. This was a serious issue as there are few opportunities on those islands to pursue early childhood coursework, and even fewer that lead to a certificate. With some effort and support from P–20, the courses became available online, which should improve access and enrollments.

P–20 also worked to increase the knowledge base of those working in PreK–3 through support for conference attendance of key stakeholders and teachers. P–20 funded the participation of several demonstration site CASs in the Harvard Graduate School of Education's four-day institute "PreK–3rd: The Foundation for Educational Success."

P–20 was contracted by HIDOE to provide professional development for five resource teachers who are responsible for overseeing the work of the PreK teachers in the 20 PreK classrooms in 18 elementary schools that HIDOE has established as part of a

---

[3] The program's formal name is the PK-3 Graduate Certificate Program, offered by the University of Hawai'i College of Education Curriculum Studies Master's Degree Program.

state-funded PreK program. Existing P–20 staff will train resource teachers on GOLD®, CLASS, Family Partnership Guidelines, HELDS, and ASQ (an environmental rating tool), and provide coaching support. These classrooms were established to serve four-year-olds who missed the new kindergarten cutoff dates and qualify based on family income. Whether they will remain after the new kindergarten entry age policy has worked its way through the system is unclear at this writing in May 2015. P–20 staff are also providing professional development support to PreK staff in a small number of charter schools as part of a federal PreK development grant.

Did P–20 successfully implement its strategies and plans?

Engaging P–3 Community Teams. As discussed above, P–20 staff initially provided little oversight of the demonstration sites. As P–20 staff turned over, P–20 took a more active approach. Logic models were created and monitored. Demonstration site staff reported that P–20-convened site coordinator meetings and a newsletter were helpful in providing ideas and networking opportunities. Monitoring of progress reports and provision of feedback to the sites was done, but we saw no instances of change in site approaches or new ideas for the sites as a result.

Evaluating and Chronicling Successful Community Strategies. In the final years of the initiative, work in this area diminished. The website was not updated during 2014, and the issue briefs that each site was to have written to highlight a successful community strategy were not produced. Nor did P–20 actively seek to identify successful local strategies. According to stakeholders, there is very little information out there about what the sites were doing or what they might have achieved. While reports from the sites were carefully read by P–20 staff, no aggregation or dissemination of successful strategies emerged, nor did P–20 do much to support sites in sustaining this work, as discussed in Chapter 4.

At the same time, it is not clear how much of the demonstration site work could have been sustained given the short-term nature of site support and the requirements attached to P–3 funding. Sites were not asked to evaluate most of their activities in the measurable way that attracts funders: Evidence of success, as described in Chapter 4, was largely limited to participation figures. Some measurable outcomes were reported from the use of CLASS, but outside support for this sort of work, which requires continued funding, would be less likely to win outside funding.

Nevertheless, the site-level CLASS work was viewed as an important outcome. According to stakeholders, no other entity had the resources or capacity to do this work. Without the WKKF grant, neither CLASS nor GOLD® would have been implemented at all (beyond its mandated use in Head Start classrooms) in the view of one stakeholder. However, neither CLASS's longer-term impact nor the degree to which it will be widely adopted is clear at this writing in May 2015. Resource issues likely will constrain its maintenance and spread. As discussed in Chapter 4, the sites struggled to train CLASS-reliable observers; turnover-induced training and mandated reliability testing require a continuing source of funds. At the same time, introducing CLASS and gaining some acceptance for it introduced teachers to monitoring and helped them see the value of monitoring and feedback to their teaching and to their students' outcomes. This change might endure even if specific CLASS implementation might not. Nevertheless, as discussed in Chapter 4, sites had some success in transferring responsibility for some of their work to organizations and individuals in the community with more reliable funding, most notably HIDOE.

Aligning and Informing Statewide Efforts (Curriculum, Standards, Culture-based Education, and Longitudinal Data Collection). More than one early childhood stakeholder noted that P–3 has been involved in all of the important changes in early childhood in the state. In particular, awareness of early education has never been higher in the state, and legislative support for early childhood has increased substantially. By publicizing and supporting the ultimate goal for the work—all students reading at grade level by grade 3—the P–3 initiative introduced the concept that it is important to look not just at kindergarten readiness, but also third-grade readiness and third-grade reading skills. Another key contribution of the P–3 initiative to the work of EOEL and the early learning community more generally is its emphasis on the importance of including HIDOE personnel in their efforts. This emphasis came out of the P–3 demonstration site work, whose early goal and consistent push was to bring together PreK and K–3 to promote early learning.

The establishment of the EOEL was a landmark accomplishment, although stakeholders to whom we spoke believe that opinion leaders would be unlikely to credit P–20 with the establishment of the EOEL. This general lack of awareness of P–20's contributions outside of early childhood circles reflects the dispersion of responsibilities in the state-level early childhood work, as discussed below, as well as a lack of branding

on the part of P–20. This stepping back by P–20 might reflect a willingness to credit actors such as the governor with the hard work attached to the EOEL and other initiatives.

P–20 has contributed to a climate of support for early childhood, and there have been successes on some measures. P–20 was instrumental in helping Hawai'i win its RTTT grant. Indeed, several stakeholders noted that without P–20 staff support, the state would have not won that competition. The win has been important to early childhood in the state, because that proposal included some components that provided significant resources to support P–3 goals. For example, in the RTTT sites, full subsidies were provided for PreK attendance.

P–20 has been instrumental in helping to develop an environment that values assessment and improvement by requiring all demonstration sites to include a validated tool in their work. Many believe that experiences with these instruments have convinced teachers of their value. P–20 also was a leader in the development of the HELDs.

Progress in the area of data-sharing has been slow because of the need to develop memoranda of understanding and coordinate data and policy that have never before been coordinated. P–20 has not fully achieved its original goals around data-set integration; some stakeholders feel that more might have been accomplished. Data-set work will continue with WKKF and local funds. This is extremely important pioneering work for the early childhood field. If the original goal of a fully integrated P–20 system is realized, this work will represent an important part of P–3's legacy.

Increasing Teacher and Leadership Capacity. P–20 helped to develop the PK-3 Graduate Certificate program and get the certificate courses online; the latter was a huge benefit to neighbor islands. However, P–20 fell far short of its initial optimistic goal of 85 percent of kindergarten teachers in demonstration site elementary schools completing the program by school year 2013–14. Arguably, the goal was far too optimistic: Much had to be done before the program was in place on O'ahu; it took much longer for it to become available in the Hawai'i island sites.[4] Forty seven teachers from the three Oahu-based demonstration sites applied for and entered the program across three cohorts; 36 completed the program. This completion figure represents 20

---

[4] P–20 chose more realistic goals over time.

percent of demonstration site teachers. P–20 was successful in transferring the sponsorship of the certificate program in early childhood to the Department of Curriculum Studies at University of Hawaiʻi Mānoa.

P–20 staff, demonstration site team members, and the CASs reported that their participation in national P–3 conferences and meetings led to a greater understanding of early childhood frameworks, research, and strategies to strengthen P–3 alignment and programs and the importance of early learning. It has also led in some cases to a demonstrably greater focus in their work on efforts in this area.

Did P–20 develop plans likely to promote long-term effects or sustainability of their P–3 efforts?

Substantial portions of the work in which P–20 participated at the state level are likely to endure; some of the work, such as the development of the certificate program, can be clearly attributed to P–20. But other work is more difficult to define, assign credit for, or predict its long-term endurance. In general, it is easier for the P–20 work to endure because it is product-focused: The certificate program is now part of the University of Hawaiʻi curricular offerings. The HELDS have become an important tool that is being used statewide. More than one stakeholder reported that P–20 and its resources were critical for producing the standards: No entity had the staff or money to produce them even though their value was clear. Their use will make a significant contribution to PreK–3 alignment. The early childhood certificate program is now housed at the University of Hawaiʻi at Mānoa, where its status is secure. The longitudinal data-set work, once completed, will become an important asset to the field. Indeed, principals and others are already beginning to ask when those data will become available. These inquirers often note that data on transitions from high school to post–high school are available and being widely used, and the inquirers would like to be able to access similar data for the PreK to kindergarten transition.

More work could be done to sustain the shift in attitudes supporting the importance of PreK and the value of early childhood knowledge. As noted in Chapter 4, key aspects of the demonstration site work depended on the support of key HIDOE personnel, most often the CAS or principals. Bringing principals along and getting them to devote severely constrained time and resources to alignment, cross-sector visiting, and transitions from PreK to kindergarten takes a good deal of work. More than once, turnover led to new people in these positions who were less convinced of P–

3 value. Demonstration site teams talked about the desirability of implementing policies that might increase the odds that people committed to P–3 theory and work would replace existing committed staff when they leave. For example, principals have the right to move teachers in their schools into whatever grades their credential covers. If there were a policy that teachers in kindergarten through grade 1 should have or be pursuing early childhood coursework, it might better support the P–3 agenda. Similarly, considering early childhood commitment in potential CASs would help as well. HIDOE must consider many factors in setting personnel policy, of course, but if these sorts of things could be considered, local sustainability in particular might be facilitated.

Did P–20, in its state-level work, engage in activities and processes that promote system change? Did system change occur?

The system framework that was applied to the demonstration sites in Chapter 4 was applied to the P–20 state-level work as well. The framework components point to strengths in the initiative as well as challenges in moving P–3 goals forward. The analysis must be understood in the broader early childhood context: P–20 lacked authority over the major players, HIDOE, the governor's office, or PreK providers. But it brought both resources and a new understanding about how to consider PreK, school readiness and how to assess effects.

Setting Explicit Goals and Expectations. Unlike the demonstration sites, which looked to P–20 for guidance, monitoring, and support, P–20 received limited input from WKKF, as discussed below. Staff turnover and changing priorities at WKKF led to reduced WKKF engagement with the P–3 initiative over time. Instead, it set its own expectations through the development of a logic model, updated yearly, that described resources, activities, outputs, and outcomes in each of the focus areas.

As noted in Chapter 4, a significant strength of Hawai'i's P–3 initiative was the unanimity and crystal clarity around its ultimate goal: all children reading at grade level by third grade. Another strength of the P–3 initiative was that stakeholders across the state shared an understanding of the general rationale behind P–3 investments and of the key components of P–3 work. We observed that most stakeholders could describe *what* P–3 work would entail. That is, they could list the primary strategies and activities that P–3 sites might pursue: for example, increase access to preschool, improve

kindergarten transitions, align PreK and K–3 curricula, and improve instruction in PreK–3. We also found that most stakeholders could articulate *why* P–3 activities might have an impact on the ultimate goal. That is, they could outline the basic logic behind P–3 reforms.

Specifying the Responsibilities of Key System Actors and Entities. A new governor who assumed office in December 2010 prioritized early childhood. As a result, the profile of early childhood expanded in the state. In 2012, this governor requested, and the Legislature approved, an Executive Office on Early Learning.

This office created a new highly visible leadership role for state government in early childhood; the EOEL quickly became a leader in state early learning policy. Not surprisingly, its creation led to some redistribution of leadership responsibilities among early childhood stakeholders. It immediately became apparent that there was some overlap in the responsibilities of the P–20 partnership and the EOEL. This lack of clarity about responsibilities was consistent with the somewhat informal way that responsibilities for leadership of various early learning efforts were being allocated among the various early childhood players, particularly P–20, HIDOE, and the EOEL. Many responsibilities were "assigned" on a case by case basis, according to the respective entities' capacities and staff strengths. While the level of cooperation inherent in such sharing is laudable (indeed, one site director told RAND that she had never before in her career seen so much cooperation among people involved in early learning), one consequence of this informal allocation of responsibilities is that people might become confused, and some entities might not receive the credit they deserve for the work they have done. While P–20 was actively involved in most of the major early childhood activities that have been implemented during the WKKF funding period and continues to pursue some of this work in the extension period, much of this work was not highly visible. While early childhood insiders would say that P–20 has made major contributions, the lack of "branding" or clear P–20 responsibilities beyond the work with the demonstration sites has made P–20 a critical but largely unheralded player. A good example of the importance of P–20's contributions might be found in the RTTT grant proposal effort. P–20 staff devoted considerable time and effort to the design and drafting of the RTTT grant proposal. That work benefited P–3, because the design provided benefits to early childhood in the RTTT sites such as early childhood subsidies. But unsurprisingly, P–20 receives little public credit for this work.

By the end of 2014, a new governor had been elected for whom early childhood was not a top priority. This left the status of the EOEL uncertain, at least in the short term. Because EOEL staff are political appointees, the director resigned when the new governor was elected. As of this writing in May 2015, a new director has not been appointed. The enabling legislation mandates that EOEL move to HIDOE in July 2015, which assures the office's continuation. It is assumed that a director will be named at that time, if not sooner.

The near absence of the EOEL as a player in early childhood in recent months has had immediate implications for the early childhood community and left responsibilities unclear. In a recent legislative hearing on PreK, someone asked, "Who will advocate for early childhood going forward, noting that since its inception it was 'always EOEL'?" The absence of EOEL is evident most directly in possible lost funding opportunities. Hawaiʻi was unable to respond to RFPs to develop early childhood programs issued by the federal government and the National Governor's Association because there was no one in charge of EOEL at the time proposals were due. Other groups that might respond, such as P–20, GBA, or the Center on the Family, cannot sponsor a proposal because they are not attached to the Governor's Office. HIDOE, an obvious candidate to sponsor these proposals in the absence of EOEL, has limited involvement in PreK and limited resources to lead proposal responses. But the expectation is that a new director will be appointed soon after the EOEL moves to HIDOE in July 2015.

The P–20 partnership's monitoring responsibilities vis-à-vis the demonstration sites were taken seriously, especially as the new sites came into the initiative, as discussed in Chapter 4. P–20 continued to take an active role vis-à-vis the sites through 2013, but began to pull back in the final funding year, reportedly to encourage sites to work on sustainability activities and to consolidate gains in their final year of funding.

Establishing Incentives and Appropriate Consequences for Meeting (or Failing to Meet) Goals and Expectations. P–20 did not actively use incentives to manage the work of the demonstration sites. In particular, sites did not engage in dissemination efforts P–20 put forward without any consequences to them. Their refusal might have had consequences for the work; the demonstration site work was not well known, and stakeholders had difficulties articulating P–20's role in many of the changes that had occurred in the early childhood sector.

The use of incentives was generally lacking in the state-level work as well. This reflects general views throughout the education sector that many if not most will pursue improvement and reform out of intrinsic motivation to better serve the children in their care (e.g., Bastick, 2000).

P–20 did recognize that willingness to engage in education and training activities might be facilitated by the availability of incentives, given the out-of-work time involved in taking classes. For example, P–20 provided tuition assistance to those who enrolled in the PK–3 Graduate certificate program when it was under P–20's supervision. It also worked to institutionalize rewards for teachers who take early childhood courses so that teachers get salary credit.

Nevertheless, P–20 did use its considerable funds to promote important P–3 goals. A number of stakeholders pointed to the travel support that P–20 supplied to encourage key actors to attend P–3 conferences, Harvard's PreK–3 Institute, and WKKF-sponsored Learning Lab meetings.

P–20 was also able to use funds to provide continuing support to activities that other partners lacked the staff to lead, such as HELDS, which is recognized as a major contribution. But funds were not used to incentivize other players or other work. Perhaps more contingent use of resources could have brought about more change.

Transparency of Monitoring and Evaluating the Performance of Key System Actors. To evaluate success, it is necessary to have in place mechanisms for monitoring performance. Like the demonstration sites, P–20 developed a logic model to guide its work and identify performance expectations and goals. Some of these goals were reported to WKKF, but as noted above, there was little feedback.

As noted in Chapter 4, P–20 increased its site monitoring after the initial years of the initiative. It did much of this by standardizing and systematizing progress reports, which were linked to demonstration site logic models.

The P–20 partnership's project to include PreK data in the state's longitudinal data set represents an important opportunity to put in place a quantitative monitoring mechanism. Some locally funded pilot work was carried out through a partnership with a large multisite preschool that is an active participant in one of the P–3 sites. Data provided by this preschool were used by P–20 staff to conduct an initial match as part of P–20's goal of integrating PreK data into the longitudinal database. P–20 staff attempted to match early learning program information with HIDOE information; the

match rate was 94 percent using Head Start data, and 82 for data from KCCA Preschools of Hawaiʻi, a nonprofit provider of preschool. This effort mirrors work that has already been done to match college students with their high school data. Other data work is progressing, but slowly. P–20 is working on data governance, using some remaining WKKF funds to convene a coordinating council comprised of organizations that collect or own data. This is extremely important pioneering work for the early childhood field, and if PreK data can be included in the state's longitudinal data set, they might well serve as part of P–3's lasting legacy.

Ensuring That Key Actors and Entities Have the Capacity to Carry Out Their Respective Responsibilities. From its inception, P–20 has targeted the PreK–3 workforce and has worked to develop its early childhood education capacity, both by increasing the supply of trained and credentialed early childhood staff and by supporting early elementary teachers in gaining a better understanding of child development and early learning. It also came to support demonstration site staff capacity-building at all levels. This work included developing graduate-level certification opportunities, as well as efforts to make them more accessible by offering them online. P–20 also ensured their sustainability by arranging to move the certificate program to University of Hawaiʻi Mānoa. P–20 has paid attention to the need to provide incentives to the workforce to promote additional education and training. For example, it provided tuition assistance to enrollees in the PK-3 Graduate Certificate program while it was under P–20's aegis.

This work, some of which has been institutionalized, represents a major accomplishment for P–20. Like all such initiatives, it must deal with the realities of staff turnover and its effects on knowledge and support. Policies to ensure some level of capacity, such as early childhood courses for HIDOE salary credit, are a major step in this direction. Other ideas, such as requiring some early childhood exposure for K–3 teachers, should also be considered.

Summary and Conclusions

Like the demonstration sites, P–20 organized its state-level work according to the focus areas it had developed, which reflected the P–3 theory of change and best practice at the time. Since these focus areas represent the cornerstones of P–3 theory, P–20's work plan included a set of activities that promoted P–3 theory and P–3 work.

P–20 worked hard to implement the activities in its logic model; much of that work has produced observable outcomes, such as the HELDS and the Early Childhood Graduate Certificate program at the University of Hawai'i.

Given its work at the state level with organizations that make and can change policy, P–20 is in a much better position than the demonstration sites to be able to sustain its work through policy change, new products, and new programs. And unlike the demonstration site work, these outcomes depend to a much lesser degree on incumbency, although the fate of EOEL at the present time (but hopefully not in the future) is clear evidence that policymaker commitment matters at every level.

P–20 work changed attitudes and norms, brought people together, and leveraged P–3 resources by bringing in other support through Hawai'i's successful RTTT proposal. But as was the case in the demonstration sites, some of the attitude change might not endure as leadership changes and organizations redefine their mission and activities. However, institutionalization of many aspects of the state-level work should help to support the attitude change and continue to build norms around the importance of the P–3 agenda.

System analysis reveals that clear goals, the potential for active monitoring through a data set that includes early childhood data, and the institutionalization of capacity-building programs represent current and potential longer-term strengths of the P–20 work. Given that P–20 had no authority to compel change, these aspects of the system that P–20 helped build have been the most effective system elements in carrying the work forward and hold the most promise of leaving lasting effects.

Unlike the work at the demonstration sites, the work of P–20 continues as of this writing in May 2015. As a result, P–20 has not yet faced the end of its WKKF support, although staff are acutely aware that it will end. Therefore, it is more difficult to consider Hawai'i without P–20's state-level P–3 work at this point. An important unknown in imagining a future without P–3's state-level presence is how EOEL will evolve. When it was housed in the governor's office, it quickly became a major player in early childhood, and in that location could engage in many aspects of the work. In its new home in HIDOE, with a new director, it might be constrained in its scope because of its tie to a department of education, or it might function more effectively once it is at greater remove from the political sphere.

6. Conclusions and Implications

In this final chapter, we summarize our findings regarding the main research questions. Then, we outline the implications of the findings for future P–3 initiatives and evaluations of them.

Findings Related to Main Research Questions

Did P–20 and the demonstration sites develop plans that aligned with best practices in P–3 at the time?

When this initiative commenced in 2009, several components were commonly understood to be at the core of a P–3 approach. In general, the framework that P–20 developed to guide its own work and the work of the demonstration sites captured these components. The demonstration sites were required to develop plans that covered seven "focus areas" (later reduced to six), which spanned the components that were considered best practices in P–3. P–20's requirement that the demonstration sites engage in activities in all focus areas appears to have played a large role in the demonstration sites' work, which covered all aspects of best practices in P–3 rather than simply focusing on selected components.

Did P–20 and the demonstration sites execute their plans as intended?

P–20 and the demonstration sites executed the majority of the activities outlined in their plans; the exceptions tended to represent activities that were modified or added to plans to respond to unforeseen opportunities or changes in the P–3 landscape. An example is the unanticipated HIDOE adoption of the Danielson assessment of instructional quality, which reduced the necessity for the P–3 initiative to promote the use of CLASS in K–3 classrooms. Participants in the P–3 initiative were not able to anticipate all the changes that might be encountered during its tenure, and, as would be expected, sites balanced their commitments to stated plans with adaptations in the face of new opportunities and implementation challenges.

Did P–20 and the demonstration sites develop long-term plans likely to promote sustainability of their P–3 efforts?

Much of the statewide and demonstration site work is likely to contribute to long-term change in Hawai'i's P–3 landscape. For example, the initiative helped change attitudes about the early childhood context in the state. Stakeholders uniformly reported that awareness of the important contribution of early learning to later academic performance grew substantially during the project period. Most attributed at least some of the change in these views to the work of the demonstration sites and to P–20's efforts, such as engaging key stakeholders in national P–3 conferences. The initiative also produced changes in knowledge and skills of many individuals engaged in P–3 services in the state, and many of these individuals will continue to work in the P–3 realm long after the initiative has formally ended.

There will be long-term impacts of the initiative due to the activities undertaken by P–20 during the project period, particularly the HELDS, the PK-3 Graduate Certificate program, and the HIDOE PDE3 courses. But only a few of the site-level initiative activities will continue past the end of the WKKF grant. There are instances where HIDOE or individual CASs have decided to continue elements of P–3 activities that the demonstration sites have undertaken, but there are limited numbers of activities that have been institutionalized to the extent that they will be sustainable when the grant funding ends. At the same time, some of the state-level products of the initiative, such as the HELDS, the PK-3 Graduate Certificate program, and the PDE3 early childhood courses that HIDOE staff can now complete for salary credits, are likely to endure.

There has been no systematic effort to plan for a next phase of the P–3 initiative or identify funding to continue P–20's activities in this area. P–20 and the sites engaged in little strategic communication regarding their accomplishments, and so stakeholders often attribute the P–3 initiative with having an important role in some of the changes that occurred during the project period, but are vague on the specific activities undertaken as part of the P–3 grant per se.

Did P–20 and the demonstration sites engage in activities that promoted system change?

Both P–20 and the demonstration sites implemented most of the elements of system change required to effectively promote it. New expectations for cooperation were put into place by P–20 when high-level HIDOE and PreK administrators were required to

jointly commit to the P–3 work in demonstration site proposals. P–20 gave sites considerable responsibility for defining and organizing the work, as site contracts were outcome-based rather than activity-based. Within sites, responsibilities were negotiated among team members and the planned work was completed for the most part. P–20 did not offer incentives or impose consequences for demonstration site performance; in the demonstration sites, incentives played little role in motivating work. Nevertheless, stakeholders noted that the availability of funds and time that P–3 afforded made an enormous difference.

At first, P–20 engaged in very limited monitoring of the demonstration sites. While monitoring increased over time, it focused on implementation of site logic models; few efforts were devoted to identifying and disseminating innovative approaches or on sustainability of the P–3 activities. P–20 itself appeared to be monitored lightly by funders and the three organizations that formed the P–20 partnership. As noted above, P–20 worked effectively in the area of capacity building at the state level, sending HIDOE and staff of other organizations to national P–3 gatherings that were revelatory for many. It developed and institutionalized the PK-3 Graduate Certificate program and the HIDOE PDE3 courses, and supported the establishment of the EOEL. It also supported the demonstration sites in their capacity-building work.

At the state level, P–20 engaged in a range of work that promoted alignment, such as the longitudinal database work. HIDOE's new involvement in delivering PreK to four-year-olds outside of early intervention preschools has brought it more actively into the early learning space, which might increase alignment. Unlike the work at the demonstration sites, the P–20 work continues as of this writing in May 2015. Although P–20 staff are acutely aware that the project is ending, the future of P–3 work being undertaken by P–20 is unknown. One question is how EOEL will evolve without P–3's state-level presence when P–20's WKKF funding ends. When EOEL was housed in the governor's office, it quickly became a major player in early childhood, and in that location could engage in many aspects of the work. In its new home in HIDOE, with a new director, it might be constrained in its scope because of its tie to a department of education, or it might function more effectively once it is at greater distance from the political sphere.

Did the P–3 initiative increase the fraction of third-grade children reading at grade level?

The ultimate goal of Hawai'i's P–3 initiative was to have all children reading at grade level by third grade. Unlike other P–3 evaluations, we were able to analyze reading-score data to assess the extent to which the initiative reached this ambitious goal. We find evidence that more years of participating in the P–3 initiative raised student reading scores modestly but significantly and increased the likelihood of scoring proficient on the state reading test. We find that being in a school that was exposed to the P–3 initiative for five years was associated with a 3.5-point increase in students' HSA reading score, which is an effect size of about 0.1. This is comparable to other estimates of the effects of nine additional weeks of schooling (Chingos, Whitehurst, and Gallaher, 2013) and is higher than an estimate of the average effect size for elementary school interventions for mainstream students (Hill et al., 2008). Other reform activities undertaken in the P–3 demonstration sites during this period might have also contributed to these outcomes, but these findings are encouraging in showing evidence that the reading-score test gap between demonstration site schools and other schools narrowed by the end of the P–3 initiative. It is also noteworthy that we found evidence of effects despite several limitations that reduced the likelihood of finding effects, including the small number of years that the majority of P–3 demonstration sites had been involved in the initiative and the low dosage of the P–3 intervention experienced by most treated students in these data, a limitation exacerbated by the emphasis of demonstration sites on PreK–kindergarten activities, to which many students in the data were never exposed.

Implications for P–3 Initiatives and Evaluations

Many of the implications for future P–3 initiatives derive from the limitations we faced as we tracked the implementation of the initiative, assessed its progress, and documented its accomplishments. Below we discuss factors that should be considered in the design, implementation, and evaluation of future P–3 initiatives.

Determine in Advance an Appropriate Balance Between Standardization and Site-Specific Needs and Resources. It is widely understood that when working with and in local communities, one-size-fits-all approaches are bound to be less effective than approaches that are targeted toward communities' needs and take advantage of

community assets (Kilburn and Maloney, 2010). But it can be difficult to determine the right balance between local autonomy and the standardization required to maximize impact in a multisite initiative directed at sites that vary widely. In Hawaiʻi, P–20 began by giving the first two demonstration sites considerable autonomy over the activities they would conduct. Eventually, it recognized the importance of some degree of standardization and required the first two and subsequent sites to implement activities in all six focus areas. It allowed sites to determine which activities to pursue in each focus area and to specify appropriate measurable outcomes in each. This change allowed more diversity in strategies and activities. An ultimate outcome measure that P–20 established to define success—all children reading at grade level by third grade—supported and rationalized this diversity, since all sites would ultimately be evaluated against reading score gains.

Consider Contracts That Specify Outcomes Rather Than Activities. A noteworthy feature of this initiative was that the contracts between P–20 and the demonstration sites specified the objective to be achieved, that is, improved third-grade reading scores through implementation of activities in each of the six focus areas that represented the components of the initiative's P–3 approach. P–20 did not specify the activities themselves. Although public administration has increasingly recognized the value of an outcomes-based approach for services rather than the typical activities-based approach, education and social service contracts have tended to continue to use activities-based contracts. The initiative described in this report demonstrates that multisite initiatives can harness the advantages of outcomes-based contracting, which include local decisionmaking processes and activities tailored to local needs and resources, while continuing to monitor performance.

Establish Measurable Outcomes for the Work. It is important to identify measurable outcomes for all of the activities to be undertaken. This way, monitoring is possible and changes can be made early on if the work is not producing expected outcomes. Standardized outcomes represent a desirable feature for multisite work because they enable cross-site comparisons as well as define a conceptually simple benchmark. P–3 initiatives that require conformity to available frameworks are implicitly standardized, sharing as they do a number of strategies and ultimate outcomes. One of the strongest features of this P–3 initiative was that all stakeholders

were able to clearly state the overarching objective and agree on its importance. This feature led to extremely strong buy-in for the initiative.

Explicitly Plan for Changes in Policy and Personnel Turnover. An initiative like P–3 requires collaboration with a variety of institutions, many of which might alter policies during the course of the initiative. Providing the flexibility to alter course in response to policy changes, rather than adhere to ex ante plans, allows an initiative to take advantage of opportunity and avoid continuing activities that might no longer be needed. However, this flexibility must be balanced against the risk that permitting midstream changes makes accountability more challenging. Personnel turnover was also an important but uncontrollable issue in several of the demonstration sites. Spending time and resources to engage a CAS or a principal in P–3 training was viewed as a critical investment to many site teams; some were dismayed when trained people left. This kind of disruption can be minimized by explicitly planning for staff turnover.

Consider Sustainability from Inception. Now that demonstration site funding has ended, some sites will continue some aspects of the P–3 work, but most of the P–3 activities will cease. This is a source of disappointment to several of the demonstration sites. P–20 offered to help the sites with planning for sustainability or activities related to it, but did not emphasize or mandate this help; the sites generally ignored it. For the most part, sites did not secure funding or institutionalize specific activities or initiatives; some sites claimed to lack the capacity to do this. Moreover, there does not appear to be an effort to "tell the story" of the accomplishments of the P–3 initiative. While many stakeholders have a positive view of the initiative and believe it was an important contributor to the improvement of early childhood education and early elementary education, they were often not able to articulate specific accomplishments that are attributable to P–3. As P–20 engages in activities during the extension of the WKKF grant, it might want to consider adding dissemination activities to the work plan.

Require Explicit Agreements Between Collaborating Parties. One area of uniform agreement among all stakeholders participating in interviews was the value inherent in requiring HIDOE and other partners to be part of the proposal and award agreement processes and to have high-level administrators within partner organizations make specific commitments. One of the hallmarks of this initiative was that it contributed to the building of strong coalitions between the early education and elementary education

sectors that most stakeholders agreed were valuable and unprecedented to this degree in the demonstration sites.

Appendix A: Qualitative Analysis Details

Qualitative Data Collection

Table A.1. Qualitative Data Collection Details

Month and Year	Interview participants	Number of participants	Entity Addressed
February 2010 initial visit	P–20 Director, Director of Early Learning Programs, P–3 Operations Director	3	P–3
	N-W site team members	7	N-W P–3 demonstration site. (At that time, N-W was one of two P–3 sites.)
	Staff of Good Beginnings Alliance	3	P–3, Good Beginnings Alliance
	Lead staff on other P–20 projects, e.g., the longitudinal data set, curriculum instruction	6	P–20
	Vice President for Academic Planning and Policy, University of Hawaiʻi	1	P–20, P–3
	Farrington site team members	6	Farrington P–3 demonstration site. (At that time, Farrington was one of two P–3 sites.)
	P–3 advisory group and partners including private and public sector leaders	8	P–3
August 2010 visit	N-W site team members	6	N-W P–3 demonstration site
	Farrington site team members	5	Farrington P–3 demonstration site

Month and Year	Interview participants	Number of participants	Entity Addressed
	P–3 Critical Friends Group, including public and private sector leaders	7	P–3
	P–3 staff	2	P–3
	Windward P–3 demonstration site team member	1	Windward P–3 demonstration site
	Honokaʻa P–3 demonstration site team member	1	Honokaʻa P–3 demonstration site
August 2011 abbreviated visit	P–20 director and P–3 staff member	2	P–3
	Farrington site team members	5	Farrington P–3 demonstration site
April 2012 visit	Windward site team members	6	Windward P–3 demonstration site
	Windward site team member	1	Windward P–3 demonstration site
	P–20 director and P–3 staff	3	P–20, P–3
	Farrington site team members	4	Farrington P–3 demonstration site
	Former P–20 director	1	P–20, P–3
	Director, Office of Early Learning	1	Governor's Office of Early Learning, P–3
	N-W site team members	5	N-W P–3 demonstration site
	KKPCA site leadership team	4	KKPCA P–3 demonstration site
	Honokaʻa site team members	6	Honokaʻa P–3 demonstration site
	P–3 staff	2	P–3
November 2012 site update phone meetings	N-W site team member	1	N-W P–3 demonstration site
	Farrington site team members	3	Farrington demonstration site

Month and Year	Interview participants	Number of participants	Entity Addressed
	Windward site team member	1	Windward demonstration site
	Honokaʻa site team member	1	Honokaʻa demonstration site
	KKPCA site team members	3	KKP demonstration site
August 2013 trip	N-W site team members	4	N-W demonstration site
	Farrington site team members and partners	7	Farrington P–3 demonstration site
	Windward site team members	4	Windward P–3 demonstration site
	Honokaʻa site team members	3	Honokaʻa demonstration site
	KKPCA site team members	4	KKPCA P–3 demonstration site
	P–20 director and P–3 staff	4	P–3
	Former P–20 director	1	P–3
	Former EOEL director	1	P–3
	Complex area superintendent and staff member	2	N-W P–3 demonstration site
June 2014 end-of-project phone meetings with four sites	N-W site team members and partners	4	N-W demonstration site
	Farrington site team members and partners	6	Farrington demonstration site
	Windward site team members and partners	6	Windward demonstration site
	Honokaʻa site team members	3	Honokaʻa demonstration site

Month and Year	Interview participants	Number of participants	Entity Addressed
November 2014 end-of-project phone meeting with remaining site	KKPCA site team members	4	KKPCA P–3 demonstration site
Spring 2015 phone meetings with key stakeholders	Stakeholder 1	1	P–3
	Stakeholder 2	1	P–3
	Stakeholder 3	1	P–3
TOTAL		161	

NOTES: After our initial visit in February 2010, site visits were conducted approximately annually beginning in August 2010. Because of an earlier than usual site visit in April 2012, we conducted group interviews with each of the demonstration site teams on the phone in November 2012. We returned for a visit in August 2013, which was our last data collection visit. In June 2014 we conducted final interviews with four site teams as they completed their P–3 work. We conducted an interview in November 2014 with a site that had been given a no-cost extension to complete its P–3 work. Throughout the project, RAND project team members met on the phone with our client, P–20, to share progress and hear updates on P–3 work and on developments in early childhood in the state. Through 2012, these calls occurred on a biweekly schedule. After that time, calls were conducted approximately monthly.

Interview Guide for 2014 Demonstration Site Visits

Welcome and Introductions

We are happy to be here, look forward to talking with you. Remind them that they signed consent forms; offer them to those who have not.

Key trip goals are to review your progress in meeting your own goals, talk about future activities and discuss P–20's role in your efforts.

Today's Agenda

Questions for site team about site context and end of project planning and status.
Review of logic model/progress report
Additional questions for site team

Understanding this site

1. Have there been any changes in the leadership or the key players (individuals and organizations) in this site over the last year? Has anything else changed in your site that has affected your work, e.g., new/reduced funding sources? New requirements?

2. What have the EC providers involved in the P–3 work gotten (that uninvolved EC providers have not gotten)? How about elementary schools? Are those that worked with P–3 implementing specific activities? Did those schools receive specific benefits?

3. What have been your key overall goals for your P–3 work? Have they changed over the time that you received P–3 funds? If yes, how? Why? Has your theory of change as articulated in your logic model changed? If so, why?

4. To what extent have you been able to achieve your goals? Please describe your success by goal.

5. What are your biggest strengths as a site? How have these strengths affected your ability to achieve your goals?

6. What are your biggest challenges as a site? How have these challenges affected your ability to achieve your P–3 goals?

7. Looking back on your efforts over the time your site has been funded by P–3, how realistic were your goals? Did you take on more than you could realistically achieve? Did you feel that you were able to use P–3 resources in the most efficient and effective ways? If not, why not? Probe for P–20 demands, expectations, lack of time, resources.

8. To what degree did your work advance yours and P–3's long-term goal of every child reading at grade-level in third grade?

9. Could you have advanced these goals without P–3 support?

10. What will happen to the work that you did with P–3 support? Probe for different activities. Will it continue? Grow? Change? Go away? What plans, if any, do you have for continuing the work that was funded by P–3 in your day jobs or in other capacities?

11. Will some of you continue in the same or similar roles after funding ends? If yes, how will you be able to do this (probe for sources of support or integration of the P–3 work into institutional functioning)?

Review of Logic Model/Progress Report

12. Looking back, how useful was your logic model in designing, implementing and assessing your P–3 work?

13   How useful were the measures of progress that you articulated in directing your activities and holding you accountable for progress?

14. Did the measures of progress that you identified continue to accurately represent your progress towards your long-term goal of every child reading at grade-level in third grade?

Questions on System Functioning

Goals, expectations and standards

The ultimate goal of your P–3 work has always been clear: more kids/all kids reading at grade level by grade 3. We have a few additional questions about goals and expectations concerning the work you are doing.

15. Were you always clear about what P–20 expected of your site? Were their criteria for success understandable? Reasonable given the attributes of your site and the resources made available? Did you believe that carrying out all the activities in your logic model would constitute success to P–20? How focused was P–20 on outputs? Outcomes? How much did P–20 staff help/hinder what you were trying to do? Did P–20 provide any value beyond the funding it provided? If yes, what was it?

16. Was your site able to pursue its most important goals in its P–3 work? If no, why not? What has happened to these goals that you were unable to pursue during the period that your site has received P–3 funding?

17. As time passed, P–20 became clearer about its expectations for the sites, and increased the number of required activities, which fell under every one of the focus areas. Did your site benefit from the wider focus implied by required activities in each focus area? Or did these required activities in each focus area dilute your efforts and potential influence?

Clarifying responsibilities of key actors

18. What roles did P–20 play in your site? Did P–20 staff closely monitor activities, read reports, monitor budgets?

19. How about here in (SITE)? Was it clear who was responsible for what? For example, who completed the forms that P–20 required?

20. Was it difficult for members of the site team to carry out P–3 responsibilities given the many responsibilities associated with their "day" jobs?

21. How important was it to you to have a paid coordinator?

Incentives and appropriate consequences

22. Did P–20 establish incentives to promote good site performance? Were there rewards for meeting/exceeding expectations?

23. If your site underperformed, e.g., didn't turn in a report, or turned it in late, were there any consequences? What if you overspent your budget? What if you didn't train all the people you said you would train? Or you didn't conduct as many meetings as you said you would? What if your site winds up not showing progress on the indicators that were part of your work plan?

24. To achieve your site goals, lots of people needed to work together to get things done. Did your site establish incentives to promote cooperation and other site activities? Were there rewards for meeting/exceeding expectations? If partners were underperforming, e.g., didn't train as many people as promised, were there any consequences?

25. Much of your work requires DOE staff cooperation. But these staff work for an organization that may have different goals which are more salient to DOE staff, e.g., science education. Were you able to come up with ways to get DOE staff to attach sufficient importance to the P–3 work? Were you able to offer sufficient incentives? Did DOE help by doing this in some way, e.g., a CAS declaring P–3 goals to be important?

Monitoring and evaluating performance and reporting on it

26. How did you know if your site was performing as expected? What sorts of feedback were provided?

27. How did P–20 know if your site was performing as expected? Did you get feedback from P–20 on your reports? On your work?

28. Overall, how well did your site carry out its P–3 work? What are the most effective things you did? The least?

Ensuring that key actors and entities have the capacity to carry out their respective responsibilities

29. Did you feel that you had the right mix of talent and capacity to do the P–3 work in this site?
30. Could you have benefited from training? Other support?

31. Does your site have sufficient staff to oversee the P–3 work?

32. Is P–20 available to provide help with record-keeping, reporting, or other such tasks? Does P–20 provide or make available needed TA for these tasks?

33. Have you found P–20 trainings and convenings to be helpful to you in your work?

34. Do ECE providers and teachers in this site have sufficient education and training to enable children to reach P–3's goal of reading at grade level by grade 3? If not, what is the site/P–20 doing to promote such education and training?

A few last questions

35. We know that the Kellogg grant that is funding much of your work is ending soon. What do expect will happen to your site's P–3 work when the money goes away? Are you doing anything now to ensure the continuity of this work?

36. Have you learned any lessons from your P–3 experiences that might be of value to other sites (or other states) that may want to implement P–3 initiatives?

Appendix B: Demonstration Site Details

Site-Specific Information

More detailed, site-specific information about each of the five P–3 demonstration sites appears below. For each site, we present the site's background and history with P–3. We then present a summary of the site's P–3 activities. A final section presents site-specific information about activity implementation and outcomes.

Farrington Site Background

This site received one of the original P–3 site grants in 2007 and focused much of its early work on professional development and collaboration between ECE and HIDOE providers. In its first year of being a demonstration site, P–20 asked the site not to use project funds to create a P–3 coordinator position in order to preserve P–3 funds for activities. This request proved to be counterproductive: After struggling for some time and relying on P–20 staff to do much of the coordination work, it was agreed that some coordination time needed to be budgeted. Once a part-time coordinator was hired, project work moved forward more rapidly and efficiently. All site team members had full-time jobs and could not take on a P–3 leadership role, so a coordinator was needed to organize the work.

Farrington Activity Summary

From the beginning of the P–3 initiative, much of the work of the site focused on professional development and collaboration. According to site staff, the P–3 team came to believe that the most effective way to make progress toward P–3's big goal of all students reading at grade level by third grade was to focus the P–3 work on teaching practice. Consequently, P–3 efforts in Farrington focused mostly on coaching and professional development. P–3 funds were used to provide substitute teachers, which facilitated teacher participation in professional development. Altogether, more than 80 percent of PreK and K–3 teachers attended at least one professional-development activity sponsored by Farrington P–3, and P–3 staff reported that teacher feedback was generally positive. Citing research that found the support of principals to be important in encouraging teachers to improve their practice and imbed developmental knowledge

into their work, team members focused efforts on principals as well. A continuous effort was made to educate principals by exposing them to early care and education experts at conferences and other professional development activities. A 2013 principal symposium during which research on early brain development was presented was viewed as very helpful in getting principals to support P–3 efforts. The site also provided $700 mini-grants to elementary schools that were willing to engage in PreK–kindergarten visits.

This site actively supported the use of CLASS, believing that CLASS assessments inform teachers about how to work with students. However, the site dropped the CLASS requirement for all K–3 teachers because it was unable to schedule the pool of trained and reliable CLASS observers to observe all of them. Instead, team members supported the use of the Peabody Picture Vocabulary Test in HIDOE classrooms. All demonstration site teachers were encouraged and supported by HIDOE to administer their own PPVTs in spring 2014.

With support from a local foundation, the site was able to assess PreK classrooms using the PPVT as well. Site team members viewed the PPVT as appealing in a number of respects: It is norm-referenced and standardized and can be used for PreK as well as early elementary grades. They also found it unique among assessments in that trained teachers are able to administer it, assess data, and use the results to refine their instruction on their own; no observers are needed. The fact that teachers, once trained, can collect data and use them to modify instruction without additional support led the P–3 site staff to view the PPVT as more sustainable in the long run than assessments that require greater infrastructure and support such as the CLASS. The site provides training to teachers on how to conduct PPVT assessments; site staff reported that many teachers in the complex can now administer PPVT on their own. To encourage teachers to undergo PPVT training, which includes introduction to the tool and how to use PPVT data in differentiating their instruction, the site provided a gift card to those who assessed their own students with the PPVT. Site staff said that teachers are now better able to work with data because of these trainings.

N-W Site Background

This site received one of the original P–3 site grants in 2007 and focused much of its early work on increasing the number of children entering kindergarten with PreK

experience and providing parents with information about PreK options to help them access early childhood education for their children. Both were major challenges in an area where a high percentage of families want to keep their young children close to them or in a traditional home-based setting. Unlike other P–3 sites, N-W had an active agency, INPEACE, in place with a history of related work when the P–3 grant was received. INPEACE was successful in convincing P–20 to allow the site team to use P–3 funds to support a project coordinator. This allowed the site to move forward, in contrast to Farrington, where P–20's initial resistance to using site funds to hire a coordinator hampered progress. In addition, INPEACE had pursued and acquired additional resources to support a constellation of related activities to support P–3 objectives. Many of the activities proposed by the site for its P–3 work grew out of the SPARK work, which was funded by WKKF.

In 2010, Hawai'i was awarded a four-year, $75 million federal RTTT grant, and N-W became one of the two sites in the state (along with KKPCA) in which RTTT was to be implemented. P–20 staff targeted N-W for participation in RTTT because of the high percentage of low-performing schools in the complex and because of its ongoing P–3 work; the idea was that there could be good synergy between RTTT and P–3. Site staff reported that the ECE-HIDOE partnership developed through P–3 helped RTTT get off the ground successfully. P–3 strategies such as wraparound services positions were written into the RTTT grant. The new CAS, who came to the site in school year 2011–12, was committed to P–3 efforts and has demonstrated her commitment with a series of policies that motivate or mandate principals to support P–3 efforts in their schools. N-W was also successful in winning a Robert Wood Johnson Foundation Community Schools project that is enabling the site to continue some of the P–3 work.

N-W Activity Summary

Much of the P–3 work in this site focused on PreK enrollments and parent involvement. Recruitment of families into P–3 activities through community canvassing has been a long-term site activity. Between August 2013 and June 2014, site staff knocked on 976 doors and successfully recruited 90 children and families into ECE providers. Despite PreK attendance declines in the state as a whole, data indicate that ECE enrollments have increased in the site.

Kamehameha Schools opened its Ka Pua Maʻili Community Learning Center in August 2014 with 185 slots; the site reported that as of July 2015, 95 percent of these new slots were filled. P–3 site team members believed that the P–3 work was important in building demand for preschool slots through recruitment and community engagement activities. The site leadership team worked to promote family visits to kindergarten classrooms as well as early kindergarten enrollment, which aids elementary schools in planning and teacher hiring and assignments. The leadership team advocated that kindergarten orientation activities at each school include a strong family involvement component. INPEACE continues to collect both early childhood capacity and early childhood enrollment data.

The use of CLASS in this site was eventually limited to PreK programs; all regular elementary schools are implementing the Danielson tool as part of the RTTT mandate. Use of both tools is expected to continue, as will coaching on them, although funds for the CLASS outside of Head Start are not assured.

P–3 was given a standing spot on the monthly principal meeting agenda; P–3 determined the content of this presentation. Its inclusion on this agenda was very important in increasing the credibility and visibility of P–3 and its efforts to promote an early childhood focus in the site to principals. When P–3 funding ended, this slot was filled by a community schools project representative; this work is the next project for several P–3 project staff.

Windward Site Background

This site was selected as the third P–3 demonstration site in 2010 when an RFA was issued for a second phase of P–3. The site included four complexes that line the east coast of the island of Oʻahu. The Windward site is the largest of the five P–3 demonstration sites and included all 23 elementary schools in its P–3 work. The site coordinator was well-known and very active in the early learning community, and both CASs were widely viewed as being very committed to early childhood. Each elementary school in the demonstration site had an early childhood team, and some quarterly principal meetings focused on early childhood. Partners in the P–3 work included a number of preschools, including a Hawaiʻi Community Action Program

Head Start, as well as Kamehameha Schools.[1] This site is a largely suburban area, and the characteristics of the students and the strengths and needs of the schools vary quite a bit across the area.

Windward Activity Summary

This site focused much of its energy on professional development. The site was successful in getting HIDOE to accept the 15 units of early childhood courses developed by the site for HIDOE professional development credit (through PDE3). Teachers can receive a pay increase after completing 15 HIDOE credits approved by their principal. The Windward P–3 site offered courses through the HIDOE system free of charge; completion of PDE3 courses is one of the few routes to a pay increase.

With strong CAS cooperation, some P–3 activities were mandated for teachers. For example, each elementary school was required to send a team comprised of an administrator, a parent, a special education PreK teacher, and a kindergarten teacher to a conference on PreK–kindergarten transitions. P–3 paid for substitute teachers. This conference was part of a long-term effort to promote transition planning. Schools were offered $500 P–3 minigrants to develop a transition plan. To qualify for the full amount, the plan had to include cross-visitation between early childhood programs and the school.

The site team, with the cooperation of the two CASs, was able to alter the kindergarten schedule at the beginning of the school year so that children began learning on the first day of school. Before the change, new kindergarten students received virtually no instruction during the first four to five days of the school year because this time was used for assessments. Ten schools hosted "Keiki Steps to Kindergarten" programs before the start of school year 2012–13; five of these programs were supported with P–3 funds. Six other schools have accessed P–3 support for other summer transition activities.

Throughout the funding period, the site worked to promote the importance of early childhood and early learning and the message that children can learn anywhere, not

---

[1] Kamehameha Schools is a private school system in Hawaiʻi developed to educate children of Hawaiʻian descent. The system serves nearly 7,000 K–12 students at three campuses and at 31 preschools statewide. In addition to its campuses, Kamehameha serves over 40,000 additional learners annually through a range of programs and community collaborations including community charter school support and literacy enhancement programs for public school children (See Kamehameha Schools, undated).

just in school. A key product was a set of literacy cards that parents can use with their children in different environments. Children engaged in a range of literacy activities, e.g., find red food labels in the supermarket to win a sticker at checkout. Data indicated that more than 200 of these literacy cards were taken by parents.

The number of children entering kindergarten with some PreK experience increased according to site staff. The percentage of kindergarten students who attended preschool prior to kindergarten entry improved slightly in the Kailua-Kalaheo complex from school year 2011–12 to 2012–13 (64.7 percent in school year 2011–12 vs. 65.9 percent in 2012–13), but the same figure declined slightly across the two years in the Castle-Kahuku complex (72.0 percent in school year 2011–12 vs. 71.4 percent in 2012–13).[2]

Honokaʻa Site Background

The group in this site, called the "Baby Steps Group," received one of the original P–20 site grants in 2007 to work on PreK–kindergarten transition and literacy in schools. The grant was initially for 18 months. In the course of this work, the group established close relationships with Waimea and Honokaʻa elementary schools, two of the three elementary schools in the complex. Honokaʻa has two feeder PreKs; Waimea has five. There is also a Migrant PreK program. The original grant ended in 2009. In 2010, Honokaʻa was selected as the fourth P–3 demonstration site.

Honokaʻa Activity Summary

The bulk of the work in Honokaʻa focused around four areas: standards, curriculum, and assessment; instruction; comprehensive early learning services/access to 0 to 5 opportunities; and family school transitions and partnerships.

Low kindergarten readiness levels and kindergarten retention rates around 30 percent led the site team to focus its P–3 work on kindergarten readiness. Staff relied on the Kindergarten Readiness Test (KRT), a Scholastic Testing Service (STS) assessment with 36 questions. The test, which includes six subtests[3], classifies children into four categories: children who score in the first two categories—ready-plus and ready—can

---

[2] See Hawaiʻi Department of Education, Hawaiʻi State School Readiness Assessment, Complex Area Results, School Year 2012-2013, website.

[3] The KRT offers six subtests: Vocabulary, Identifying Letters, Visual Discrimination, Phonemic Awareness, Comprehension and Interpretation, and Mathematical Knowledge.

benefit from standard kindergarten instruction without any need for additional support; marginally ready children need some support to benefit from standard instruction. A fourth category of children, not-ready, cannot benefit from standard instruction without considerable support.[4] Site staff noted that there were no structures in place to support children in the not-ready category in their site; they worked on an ad hoc basis to provide such supports. Site staff noted that the number of not-ready children would likely decrease as the new, more restrictive kindergarten enrollment age policy is implemented.

The site's early childhood curriculum specialist used aggregate classroom-level KRT data and subscale scores as a tool in working with PreK and kindergarten teachers to improve children's kindergarten readiness. Both participating elementary schools administered the KRT to all of their kindergarten students in participating kindergarten classrooms: 106 students in Waimea elementary school and 66 students in Honoka'a elementary school. Two years ago, the test revealed that 24 percent of children were not ready for kindergarten; data collected from incoming kindergarten students last year showed a 25 percent improvement in KRT scores for students entering kindergarten in 2013. Site staff felt that the improvement could be attributed to the work that the early childhood consultant does with PreK directors and elementary school teachers in joint articulation meetings. Typically, in these meetings, the consultant reviewed aggregate KRT subscale scores and their implications for instruction in an effort to improve instruction to promote kindergarten readiness. The goal of these meetings has been to highlight areas where children's skills need improvement and to help teachers improve their instruction in these areas. The early childhood curriculum specialist also attended biweekly kindergarten articulation meetings and observed the teaching of kindergarten teachers. These observations formed the basis for instructional coaching that she provided.

Use of the KRT tool was not supported by the P–20 team, which argued that the test lacks reliability and validity data. The Honoka'a site team was able to attract other funding to continue use of the KRT tool after P–20 refused to support the work for the reasons stated above.

---

[4] According to STS, KRT levels of readiness are "…related to national percentiles and stanines" (Scholastic Testing Service, undated).

Cross-site visits between kindergarten and PreK teachers constituted another important site activity. These site visits, followed by focused debriefs, were described by PreK teachers as very helpful in supporting strong relationships between the two systems. Site team members believed that these visits were very important; PreK teachers are shocked each year by the structured nature of kindergarten classrooms, and kindergarten teachers are shocked by the high-functioning classroom behavior skills of four-year-olds in PreK. An alignment review that was conducted of PreK and kindergarten curricula reportedly has been helpful to kindergarten teachers because it gives them guidance about what efforts had already been made to build the skills of kindergarten students who were not kindergarten-ready. According to site team members, all of this work helped to build in kindergarten teachers a stronger developmental perspective on learning in elementary schools. Team members noted that such a perspective is not normally embedded in HIDOE schools.

The Honoka'a site experienced challenges in implementing the CLASS and was unable to meet its CLASS milestones. The site's failure to meet CLASS goals derived from its inability to find and effectively train sufficient numbers of reliable CLASS assessors; a number of those who completed the training failed to meet CLASS reliability criteria.

A major professional development activity is the annual Hawai'i Island Early Childhood Conference (HIECC) that the site organizes and conducts. Site team members along with many others on Hawai'i Island fund, plan, and conduct the annual HIECC conference.

Site staff devoted considerable efforts to increasing enrollments in PreK programs in several ways. The site updated its inventory of ECE programs, including their enrollment capacity. P–3 funds also helped to fund Play and Learn groups.[5] These groups, in addition to providing early learning experiences for children not enrolled in PreK, represent an additional way to reach families to inform them of positive parenting skills, connect them to other families, and provide information on more formal PreK opportunities.

---

[5] P–3 funds supported the playgroup leader in one site; other costs, mainly space, were covered by the host elementary school.

Ka'u-Kea'au-Pahoa Complex Area (KKPCA) Site Background

KKPCA was already involved in RTTT when P–20 asked the site to begin P–3 work in 2011, believing that bringing P–3 resources into the site would amplify the impact of RTTT. The site convened a planning team led by KKPCA's CAS to determine the best use of the P–3 funds at the beginning of the effort. The site coordinator met with P–20 staff to learn what P–3 funds could be used for; she wanted to ensure that RTTT and P–3 did not duplicate their efforts. The RTTT goal related to PreK was to open early childhood learning opportunities to children so that they would be more prepared for kindergarten. RTTT funds supported this goal by providing full subsidies for PreK attendance to families who were subsidy-eligible. P–3 funds enabled site team members to bring an early childhood focus into HIDOE schools: Whereas the RTTT plan for early childhood focused on fully subsidizing parents who enrolled their children in PreK programs, there were no plans to infuse early childhood approaches or developmentally appropriate practice into RTTT schools. Indeed, according to site team members, early childhood was not a part of any RTTT discussion for almost a year. This lack of focus was understandable given that RTTT brought with it a host of requirements that had to be met, including a requirement to implement the Danielson suite of assessments.

P–3 funds could also be used to fund private PreK providers. This was important because there were too few community preschool programs to serve eligible parents when RTTT began. Those running PreK programs in other locations offered to come in and open PreK classrooms in the complex, but P–3 and HIDOE staff agreed that this approach was not sustainable: These providers would certainly leave the complex when P–3 funding ended. New HIDOE preschool programs also were opened on four elementary school campuses to respond to this unmet local need for preschool slots.

Ka'u-Kea'au-Pahoa Complex Area Activity Summary

The site focused its efforts on HIDOE more than other sites did because with limited local infrastructure (particularly ECE programs), schools serve as community centers in the complex. As a consequence, engaging with HIDOE was critical. A highly committed CAS also motivated HIDOE support. She communicated that the "…P–3 project was addressing the early childhood section of the RTTT plan"; there was an expectation for school administrators to interface with the project. School-based efforts

focused on gaining support for the CLASS as a tool to improve instruction and on developing programs that provide children who are not involved in formal PreK activities some exposure to early learning activities. Other efforts to engage parents in the education of their children were refined and expanded.

Implementing CLASS was a particular challenge because RTTT teachers were mandated to undergo Danielson training and observations. Nevertheless, the core team persevered and was able to engage both PreK and a small number of volunteer kindergarten teachers in CLASS. The team was also able to connect CLASS data and professional development content. They used information from two rounds of CLASS observations to design their professional development training for preschool, kindergarten, and grade 1 teachers.

The project core team sought to provide professional development to principals designed to encourage them to coordinate and align P–3 and K–3 efforts. This was a hard sell given RTTT, which imposed many mandates and administrative requirements. Nevertheless, all elementary principals in the complex attended at least one P–3-sponsored professional development activity on leadership (some principals attended several).

Kindergarten teachers in the complex were targeted for professional development. All kindergarten teachers attended sessions that focused on research-based teaching strategies for emergent readers, along with other strategies to build literacy and support development in the early years of school. Kindergarten teachers also made a visit to Kamehameha Schools PreK, which included a briefing and discussion about curriculum, transition, and classroom issues. Training for first-grade teachers was also offered. A set of activities focused on RTTT PreK teachers was implemented as well. These teachers attended a series of professional development sessions that focused on implementing developmentally appropriate instruction.

Given the relatively few PreK programs and other barriers to PreK attendance in the complex area such as geographic isolation, the site focused on building robust Play and Learn groups that met once a week in two communities identified to be most in need of early childhood development opportunities. One of these groups met in a local elementary school; the other met in a community center. But transportation was a problem, and when one coordinator was offered full-time work, one group ended. Noting the value of having flexible P–3 funds, the Play and Learn group was replaced

with a family literacy outreach effort and a two-week kindergarten Play and Learn program in one elementary school. Site staff indicated that the site has seen a tripling in enrollments in early childhood programs; a major share of this increase is a result of RTTT subsidies and the establishment of HIDOE-supported classrooms on school campuses.

Appendix C: Multivariate Estimation Results

As described above, the third-grade reading-score analysis involves estimating student-level difference-in-difference models for students at all Hawaiʻi elementary schools, an analysis that compares rates of change over time in student outcomes on key indicators within demonstration site elementary schools to change rates on those same indicators for students in other elementary schools (see Cameron and Trivedi, 2005, for discussion of difference-in-difference estimates). Another way of stating this is that we examine the degree to which the students in schools participating in the P–3 demonstration projects realize measurably greater gains over time compared to students at schools not participating in the demonstration sites, holding student and school characteristics constant.

We examine two outcome variables: individual students' scale scores on the Hawaiʻi State Assessment (HSA) third-grade reading assessment and a proficiency indicator variable that equals one if the student's score on the HSA is 300 or greater. This indicator of proficiency is established by HIDOE. We obtained data from HIDOE for 2008–2014 for all students in the state as well as the other variables we include in the model. These additional variables include student-level control variables that are associated with test scores in other studies (e.g., Chingos, Whitehurst, and Gallaher, 2013). Including these variables is important to account for the possibility that demonstration site test scores change over time owing to changes in student composition rather than changes in P–3 exposure. The additional student-level variables that we include in the model are gender, race and ethnicity, whether the student is eligible for FRPL, whether the student is designated as an English Language Learner, whether the student receives special education services, age at the time of the test, and whether the student has repeated a grade. We also include some time-varying school-level variables—the percentage of students who are eligible for FRPL and the percentage of Pacific Islanders—to control for the possibility of changing school composition. The "difference-in-difference" estimates are generated by including year and school indicator variables in the model. This holds constant factors that vary across year but not schools, and school-level factors that are constant over time. Note that we

111

would adjust the standard errors of these estimates to account for the fact that students are "clustered" in schools, and so we cannot treat students at the same schools as independent observations. If we did not make this adjustment, we would overstate the true amount of variation in the data and calculate confidence intervals that were too small (Shadish et al., 2005).

Equations

The model that we estimate using STATA version 13.1 is:

$$
\begin{aligned}
Y_{ist} = \alpha_0 &+ \alpha_1 GENDER_{ist} + \alpha_2 RACETH_{ist} + \alpha_3 BIRTHMONTH_{ist} + \alpha_4 FRPLELIGIBLE_{ist} \\
&+ \alpha_5 ENGLISHLANGLEARNER_{ist} + \alpha_6 SPECIALED_{ist} \\
&+ \alpha_7 AGE + \alpha_8 REPEATEDGRADE + \alpha_9 SCHOOL\ FRPLRATE_{st} \\
&+ \alpha_{10} SCHOOL\ PACIFICISL_{st} + \alpha_{11} D_{ist} + \delta_t + \delta_s + \varepsilon_{ist}
\end{aligned}
$$

where i = student, s = school, and t = year.

D is an indicator of the treatment effect, which is turned on for that student in years where the P–3 initiative was active in that school's demonstration site. We also estimated a specification that replaced D with the number of years that the P–3 initiative had been in place in the school to account for students' amount of P–3 exposure. We would expect that student scores would increase with the number of years that they had been exposed to P–3.

The following tables report results for four specifications: for each of the two outcome variables with the P–3 intervention specified as either an indicator variable when the student was in a P–3 school in a year that P–3 was implemented (Tables B.1 and B.3) or a variable indicating the number of years the student had been exposed to P–3 (Tables B.2 and B.4).

Note that this analysis does not account for student movement across schools and assumes that students had been at the same school during their entire tenure. Student mobility would affect the results only if students were more likely to move into or out of P–3 demonstration site schools or the movement over time resulted in changing student composition over time. We do not have information on the former, but our data do not show that the composition of P–3 schools relative to non-P–3 schools changed over time.

Table B.1.

Linear Regression Estimates for HSA Reading-Scale Score with Indicator for P–3 Participation

Variable	Coefficient	Standard Error	P > |t|
In P–3 School in P–3 Year	1.64	0.90	0.07
Year is 2009	3.48	0.52	0.00
Year is 2010	10.78	0.53	0.00
Year is 2011	8.86	0.74	0.00
Year is 2012	9.95	0.60	0.00
Year is 2013	12.57	0.68	0.00
Year is 2014	8.77	0.68	0.00
Female	7.50	0.23	0.00
Asian	3.71	0.38	0.00
Pacific Islander	-6.06	0.34	0.00
White	7.53	0.42	0.00
FRPL	-9.94	0.38	0.00
Designated English Language Learner	-25.09	0.55	0.00
Special Education Student	-42.30	0.66	0.00
Student Age	2.62	0.35	0.00
Student Repeated a Grade	-21.22	1.85	0.00
Percentage of FRPL Students at School	-6.35	8.32	0.45
Percentage of Pacific Islanders at School	14.07	11.56	0.23
Constant	300.43	3.72	0.00

NOTE: Outcome variable is HSA reading-scale score. N = 98,805. Also included in the model are indicators for 207 schools, and these are not shown. Robust standard errors are adjusted for clustering at the school level. The year 2008 is the omitted year.

Table B.2.

Linear Regression Estimates for HSA Reading-Scale Score with Years of P–3 Participation

Variable	Coefficient	Standard Error	P > |t|
Years of P–3 Participation	0.69	0.32	0.03
Year is 2009	3.48	0.52	0.00
Year is 2010	10.88	0.53	0.00
Year is 2011	9.01	0.72	0.00
Year is 2012	10.01	0.58	0.00
Year is 2013	12.50	0.68	0.00
Year is 2014	8.58	0.70	0.00
Female	7.50	0.23	0.00
Asian	3.71	0.39	0.00
Pacific Islander	-6.06	0.34	0.00
White	7.53	0.42	0.00
FRPL	-9.94	0.38	0.00
Designated English Language Learner	-25.09	0.55	0.00
Special Education Student	-42.29	0.66	0.00
Student Age	2.62	0.36	0.00
Student Repeated a Grade	-21.25	1.85	0.00
Percentage of FRPL Students at School	-5.71	8.41	0.50
Percentage of Pacific Islanders at School	14.05	11.46	0.22
Constant	300.35	3.72	0.00

NOTE: Outcome variable is HSA reading-scale score. N = 98,805. Also included in the model are indicators for 207 schools, and these are not shown. Robust standard errors are adjusted for clustering at the school level. The year 2008 is the omitted year.

Table B.3.

Logistic Regression Estimates for HSA Reading Proficiency with Indicator for P–3 Participation

Variable	Coefficient	Standard Error	P > |t|
In P–3 School in P–3 Year	0.092	0.069	0.18
Year is 2009	0.026	0.040	0.52
Year is 2010	0.452	0.038	0.00
Year is 2011	0.286	0.052	0.00
Year is 2012	0.461	0.047	0.00
Year is 2013	0.573	0.052	0.00
Year is 2014	0.130	0.050	0.01
Female	0.460	0.020	0.00
Asian	0.250	0.027	0.00
Pacific Islander	-0.353	0.025	0.00
White	0.405	0.033	0.00
FRPL	-0.577	0.028	0.00
Designated English Language Learner	-1.577	0.038	0.00
Special Education Student	-2.750	0.051	0.00
Student Age	0.090	0.028	0.00
Student Repeated a Grade	-1.396	0.149	0.00
Percentage of FRPL Students at School	-0.061	0.543	0.91
Percentage of Pacific Islanders at School	0.730	0.921	0.43
Constant	0.846	0.298	0.01

NOTE: Outcome variable equals 1 if student achieves HSA reading score of 300 or greater. N = 98,762. Scores from five schools are dropped from the model because no students scored in the proficient range. Also included in the model are indicators for 207 schools, and these are not shown. Robust standard errors are adjusted for clustering at the school level. The year 2008 is the omitted year.

Table B.4.

Logistic Regression Estimates for HSA Reading-Scale Score with Years of P–3 Participation

Variable	Coefficient	Standard Error	P > |t|
Years of P–3 Participation	1.04	0.02	0.05
Year is 2009	1.03	0.04	0.52
Year is 2010	1.58	0.06	0.00
Year is 2011	1.34	0.07	0.00
Year is 2012	1.59	0.07	0.00
Year is 2013	1.76	0.09	0.00
Year is 2014	1.12	0.06	0.03
Female	1.58	0.03	0.00
Asian	1.28	0.04	0.00
Pacific Islander	0.70	0.02	0.00
White	1.50	0.05	0.00
FRPL	0.56	0.02	0.00
Designated English Language Learner	0.21	0.01	0.00
Special Education Student	0.06	0.00	0.00
Student Age	1.09	0.03	0.00
Student Repeated a Grade	0.25	0.04	0.00
Percentage of FRPL Students at School	1.00	0.55	0.99
Percentage of Pacific Islanders at School	2.12	1.96	0.42
Constant	2.30	0.69	0.01

NOTE: Outcome variable equals 1 if student achieves HSA reading score of 300 or greater. N = 98,762. Scores from five schools are dropped from the model because no students scored in the proficient range. Also included in the model are indicators for 207 schools, and these are not shown. Robust standard errors are adjusted for clustering at the school level. The year 2008 is the omitted year.

References

Adams, J. E., and M. Kirst, "New Demands for Educational Accountability: Striving for Results in an Era of Excellence," in J. Murphy and K. S. Louis (eds.), *Handbook of Research in Educational Administration*, 2nd ed., San Francisco, Calif.: Jossey-Bass, 1999, pp. 463–489.

Armstrong, J., "What Is an Accountability Model?" Issue Paper, National Forum on Accountability, Education Commission of the States, Denver, Colo., 2002.

Bogard, Kimber, *Mapping the P–3 Continuum (MAP): P–3 as the Foundation of Education Reform*, Foundation for Child Development, New York, September 11–12, 2003.

Bogard, Kimber, and Ruby Takanishi, *PK-3:* "An Aligned and Coordinated Approach to Education for Children 3 to 8 Years Old," *Social Policy Report: Giving Child and Youth Development Knowledge Away*, Volume XIX, Number III, Society for Research in Child Development, 2005.

Buchanan, D., L. Fitzgerald, D. Ketley, R. Gollop, J. L. Jones, S. Saint Lamont, A. Neath, and E. Whitby, *International Journal of Management Reviews*, Vol. 7, No. 3, 2005, pp. 189–205.

Burchinal, M., N. Vandergrift, P. Pianta, and A. Mashburn, "Threshold Analysis of Association Between Child Care Quality and Child Outcomes for Low-Income Children in Pre-Kindergarten Programs," *Early Childhood Research Quarterly*, Vol. 25, No. 2, 2010, pp. 166–176. As of September 1, 2015: http://dx.doi.org/10.1016/j.ecresq.2009.10.004

Cameron, A. Colin, and Pravin K. Trivedi, *Regression Analysis of Count Data,* New York: Cambridge University Press, 1998.

Campbell, F. A., and C. T. Ramey, "Cognitive and School Outcomes for High-Risk African-American Students at Middle Adolescence: Positive Effects of Early Intervention," *American Educational Research Journal*, Vol. 32, 1995, pp. 743–772.

Cannon, J., and L. A. Karoly, *Who Is Ahead and Who Is Behind? Gaps in School Readiness and Student Achievement in the Early Grades for California's Children*, Santa Monica, Calif.: RAND Corporation, TR-537, 2007.

Chen, H., *Practical Program Evaluation: Assessing and Improving Planning, Implementation and Effectiveness*, Thousand Oaks, Calif.: Sage Publications, 2005.

Chingos, M. M., G. J. Whitehurst, and M. R. Gallaher, *School Districts and Student Achievement*, Washington, D.C.: Brookings Institution, Brown Center on Education Policy, 2013. As of May 12, 2015:
http://www.brookings.edu/~/media/research/files/papers/2013/3/27-school-district-impacts-whitehurst/districts_technical_paper_final.pdf

Chinman M., P. Imm, and A. Wandersman, *Getting to Outcomes 2004: Promoting Accountability Through Methods and Tools for Planning, Implementation, and Evaluation*, Santa Monica, Calif.: RAND Corporation, 2004.

Danielson, C., *The Handbook for Enhancing Professional Practice: Using the Framework for Teaching in Your School*, 2nd. ed., Alexandria, VA: Association for Supervision and Curriculum Development, 2007.

Dunn, L., and D. Dunn, *Peabody Picture Vocabulary Test*, 4th ed., Johannesburg, South Africa: JvR Psychometrics, 2007.

Fiester, L., *EARLY WARNING! Why Reading by the End of Third Grade Matters*, Baltimore: The Annie E. Casey Foundation, 2010.

Fixsen, D. L., S. F. Naoom, K. A. Blase, R. M. Friedman, and F. Wallace, *Implementation Research: A Synthesis of the Literature*. Tampa, Fla.: University of South Florida, Louis de la Parte Florida Mental Health Institute, 2005.

Foundation for Child Development, "Components" web page, undated. As of September 1, 2015:
http://fcd-us.org/our-work/prek-3rd-education/components

Friedman, Mark, *Trying Hard Is Not Good Enough: How to Produce Measurable Improvements for Customers and Communities*, Victoria, Canada: Trafford Publishing, 2005.

Fuerst, J. S., and D. Fuerst, "Chicago Experience with an Early Childhood Program: The Special Case of the Child–Parent Center Program," *Urban Education*, Vol. 28, 1993, pp. 69–96.

Golan, Shari, Marjorie Wechsler, Ashley Campbell, C. J. Park, Dana Petersen, and Mario Snow, *Early Implementation of Ready Schools Miami: Findings of the First-Year Evaluation*, SRI Project 18246, Menlo Park, Calif.: SRI International, December 2008. As of September 1, 2015: http://www.sri.com/sites/default/files/publications/readyschoolsyear1report.pdf

Gormley, W., and D. Weimer, *Organizational Report Cards*, Cambridge, Mass.: Harvard University Press, 1999.

Gormley, William T., and Deborah Phillips, "The Effects of Universal Pre-K in Oklahoma: Research Highlights and Policy Implications," *Policy Studies Journal*, Vol. 33, No. 1, 2005, pp. 65–82.

Graves, W., "PK-3: What Is It and How Do We Know It Works?" New York: Foundation for Child Development, FCD Policy Brief, Advancing PK-3, No. 4, May 2006.

Hawai'i Department of Education (HIDOE), "Smarter Balanced Assessments" web page, undated-a. As of May 7, 2015: http://www.hawaiipublicschools.org/TeachingAndLearning/Testing/StateAssessment/Pages/home.aspx

———, "Student Assessment Section (SAS)" web page, undated-b. As of May 7, 2015: http://sas.sao.k12.hi.us/STATE/SAO/SASWebsite.nsf/By+Category/74AD3443EFFA4FDF0A2576340006576F?OpenDocument

———, "Hawai'i State School Readiness Assessment, Complex Area Results, School Year 2012–2013," web page, 2012. As of May 15, 2015: http://arch.k12.hi.us/PDFs/hssra/2013/Windward/Complex%20Area-8-Castle-Kahuku.pdf

Hawai'i P–3 Initiative, website, undated. As of May 11, 2015: http://p3hawaii.org/

Hawai'i P–20 Partnerships for Education, website, undated. As of May 7, 2015: http://www.p20hawaii.org/

Hernandez, D., *Double Jeopardy: How Third-Grade Reading Skills and Poverty Influence High School Graduation,* Baltimore: The Annie E. Casey Foundation, 2011.

Hill, C. J., H. S. Bloom, A. R. Black, and M. W. Lipsey, "Empirical Benchmarks for Interpreting Effect Szes in Research," *Child Development Perspectives,* Vol. 2, No. 3, 2008, pp. 172–177.

Hill, P. T., and J. Bonan, *Decentralization and Accountability in Public Education*, R-4037-MC, Santa Monica, Calif.: RAND R-4037-MC, 1991.

Kamehameha Schools, website, "About Kamehameha Schools," web page, undated. As of September 1, 2015:
http://www.ksbe.edu/about_us/

Karoly, Lynn A., Peter W. Greenwood, Susan S. Everingham, Jill Houbé, M. Rebecca Kilburn, C. Peter Rydell, Matthew Sanders, and James Chiesa, *Investing in Our Children: What We Know and Don't Know About the Costs and Benefits of Early Childhood Interventions*, Santa Monica, Calif.: RAND Corporation, MR-898-TCWF, 1998.

Karoly, Lynn A., M. Rebecca Kilburn, Jill S. Cannon, *Early Childhood Interventions: Proven Results, Future Promise*, Santa Monica, Calif.: RAND Corporation, MG-341-PNC, 2005.

Kauerz, K., *NALEO Legislative Hearing, P–3: What It Is, Why It's Important, and State Approaches*, September 23, 2008.

Kauerz, K., and J. Coffman, *Framework for Planning, Implementing, and Evaluating PreK-3rd Grade Approaches*, Seattle, Wash.: College of Education, University of Washington, 2013.

Kilburn, M. R., and S. I. Maloney, *Priorities for Investments in Children and Families in Caddo and Bossier Parishes: Application of a Unique Framework for Identifying Priorities*, RAND Corporation, TR-821-CFSB, 2010.

Learning Point Associates, The Center for Comprehensive School Reform and Improvement, *School Restructuring Under No Child Left Behind: What Works When? A Guide for Education Leaders,* 2006. As of August 29, 2015:
http://files.eric.ed.gov/fulltext/ED496104.pdf

Lieberman, A. (ed.), *The Roots of Educational Change: International Handbook of Educational Change*, Dordrecht, Netherlands: Springer, 2005.

McLaughlin, M. W., and D. Mitra, "Theory-Based Change and Change-Based Theory: Going Deeper, Going Broader," *Journal of Educational Change*, Vol. 2, 2001, pp. 301–323.

McLaughlin, M. W., and L. A. Shepard, *Improving Education Through Standards-Based Reform: A Report by the National Academy of Education Panel on Standards-Based Reform*, Stanford, Calif.: National Academy of Education, 1995.

Mead, Sara, *Education Reform Starts Early: Lessons from New Jersey's PreK–3rd Reform Efforts*, Education Policy Program, Early Education Initiative, New America Foundation, December 2009. As of September 1, 2015:
http://www.Newamerica.net

National Center for Education Statistics (NCES), *The Nation's Report Card: Findings in Brief—Reading and Mathematics 2011, National Assessment of Education Progress at Grades 4 and 8*, U.S. Department of Education, Institute of Education Sciences, NCES 2012-458, 2012. As of May 12, 2015:
http://nces.ed.gov/nationsreportcard/pdf/main2011/2012459.pdf

O'Day, J. A., "Complexity, Accountability, and School Improvement," *Harvard Educational Review*, Vol. 72, No. 3, 2002, pp. 293–329.

O'Donnell, C. L., "Defining, Conceptualizing, and Measuring Fidelity of Implementation and Its Relationship to Outcomes in K–12 Curriculum Intervention Research," *Review of Educational Research*, Vol. 78, No. 1, 2008, pp. 33–84.

Organisation for Economic Co-operation and Development (OECD), *Education at a Glance 2014: Highlights*, OECD Publishing, 2014. As of May 12, 2015:
http://www.oecd-ilibrary.org/education/education-at-a-glance-2014_eag_highlights-2014-en

Pande, P., and R. Neuman, *The Six Sigma Way: How GE, Motorola, and Other Top Companies Are Honing Their Performance*, New York: McGraw-Hill, 2000.

Reynolds, A. J., "Effects of a Preschool Plus Follow-on Intervention for Children at Risk," *Developmental Psychology*, Vol. 30, 1995, pp. 787–804.

Reynolds, Arthur, Katherine Magnuson, and Suh-Ruu Ou, "*PK-3 Education: Programs and Practices that Work in Children's First Decade,*" The Foundation for Child Development, FCD Working Paper: Advancing PK-3, No. 6, January 2006.

———, "Preschool-to-Third Grade Programs and Practices: A Review of Research," *Children and Youth Services Review*, Vol. 32, 2010, pp. 1121–1131.

Reynolds, A. J., J. A. Temple, D. L. Roberson, and E. A. Mann, "Age 21 Cost–Benefit Analysis of the Title I Chicago Child–Parent Centers," *Educational Evaluation and Policy Analysis*, Vol. 24, No. 4, 2002, pp. 267–303.

Rice, C., and J. McLaughlin, *Providing Tools Toward Quality: The Status of P–3 Teacher Preparation Programs in New Jersey*, Association for the Children of New Jersey, New Jersey, 2007.

Rothman, R., *Measuring Up: Standards, Assessment, and School Reform*, San Francisco, Calif.: Jossey-Bass, 1995.

Ryan, G. W., and H. Martínez, *Description of Dr. Miriam and Sheldon G. Adelson Medical Research Foundation Collaborative Research Model*, Santa Monica, Calif.: RAND Corporation, WR-604-AMRF, 2008.

Scholastic Testing Service, "Kindergarten Readiness Test (KRT)," web page, undated. As of September 1, 2015:
http://www.ststesting.com/krt.html

Shadish, W. R., T. D. Cook, and D. Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Boston: Houghton Mifflin Company, 2002.

Stecher, B., F. Camm, C. Damberg, L. Hamilton, K. Mullen, C. Nelson, P. Sorenson, M. Wachs, A. Yoh, and G. Zellman, *Towards a Culture of Consequences: Performance Based Accountability Systems (PBAS) for Public Services*, Santa Monica, Calif.: RAND Corporation, 2010.

"Starting Strong in Washington State: Early Learning Lessons and Success Stories," web page, Superintendent of Public Instruction, State of Washington, undated. As of September 1, 2015:
http://fcd-us.org/sites/default/files/StartingStrongBrief_Final_0.pdf

Takanishi, R., and K. Kauerz, "PK Inclusion: Getting Serious About a P-16 Education System," *Phi Delta Kappan*, Vol. 89, No. 7, 2008, pp. 480-487.

Tibbits, M. K., B. K. Bumbarger, S. J. Kyler, and D. F. Perkins, "Sustaining Evidence-Based Interventions under Real-World Conditions: Results from a Large-Scale Diffusion Project," *Prevention Science*, Vol. 11, No. 3, 2010, pp. 252–262.

Welch, J., with J. Byrne, *Jack, Straight from the Gut*, New York: Warner Books, 2001.

What Works Clearinghouse, U.S. Department of Education, Institute of Education Sciences (IES), Washington, DC, 2015. As of September 1, 2015: http://ies.ed.gov/ncee/wwc/default.aspx

Zellman, G. L., and M. R. Kilburn, *Hawaiʻi P–3 Initiative: Findings from the First Year of the Evaluation*, Santa Monica, Calif.: RAND Corporation, TR-921-HDED, April 2011. As of September 1, 2015: http://www.rand.org/pubs/technical_reports/TR921.html

Zellman, G. L., G. W. Ryan, R. Karam, L. Constant, H. Salem, G. Gonzalez, N. Orr, C. A. Goldman, H. Al-Thani, and K. Al-Obaidli, *Implementation of the K–12 Education Reform in Qatar's Schools*, Santa Monica, Calif.: RAND Corporation, MG-880-QATAR, 2009.