Developing Decision-Making Tools through Partnerships

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

Summary

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

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

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

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

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

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

Readiness Assessments in the United States

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

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

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

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

Assessments as Decision-Making Tools

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

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

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

The Virginia Kindergarten Readiness Program

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

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

Partnership History

Virginia defines school readiness as:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Challenges of the Work

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

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

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

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

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

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

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

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

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

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

VKRP's Successes

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

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

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

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

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

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

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

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

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

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

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

Recommendations for Researchers and Policy Makers

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

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

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

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

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

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

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

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

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

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

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

Conclusions

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

Endnotes

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

2. Ibid.

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

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

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

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

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

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

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

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

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

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

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

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

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