



USES AND MISUSES OF KINDERGARTEN READINESS ASSESSMENT RESULTS

ELLIOT REGENSTEIN

MAIA CONNORS

RIO ROMERO-JURADO

JOYCE WEINER

OVERVIEW

Federal and state leaders have increasingly focused on early learning as a key strategy for improving overall school performance.¹ Because a significant percentage of the achievement gap opens before kindergarten entry, early learning is an essential investment to help young children at risk of poor outcomes achieve their long-term potential.² When investing in the early years, policymakers should be focused on ensuring that early learning programs are designed and funded to achieve educational outcomes—including language and cognitive development, social emotional development, self-regulation, and other skills for learning.³

As early learning systems have grown and matured, there has been an increase in the use of kindergarten readiness assessments (KRAs) that measure where children are developmentally as they transition into kindergarten. These KRAs are generally designed to rely on data reported by classroom teachers and collect a broad range of information about a child's abilities. Kindergarten readiness assessments have been a valuable tool for leaders and practitioners to identify gaps in children's knowledge and skills and to enhance teaching and learning. Data from KRAs has also been used to measure kindergarten readiness of children statewide, to inform policy decisions about early learning resources and systems.

Two states have chosen to use kindergarten readiness assessment results as part of their accountability systems for early learning providers, to rate early learning programs in part on whether they produce stronger KRA scores. Policymakers in other states have also considered using KRA results in this manner. But this is a practice with potentially serious negative consequences. The current kindergarten readiness assessments have not been designed for this purpose, and states are attaching to assessment results high stakes that are inappropriate for these tools. The federal Every Student Succeeds Act (ESSA) requires the use of assessment scores for school accountability in 3rd grade and up, but the assessments used to produce those scores were designed to be used in school accountability⁴—unlike kindergarten readiness assessments. As states continue to refine their accountability and assessment policies, it is important not to misuse KRA results in ways that draw incorrect conclusions from results and potentially discredit a tool that can help support instructional improvement.

This paper is meant to help state and local education policymakers and leaders understand how to use KRA results appropriately—and how to avoid stretching them too far. First, the paper describes what kindergarten readiness is and how it is assessed. The paper then discusses the beneficial uses of kindergarten readiness assessment, which include providing system-level policy information and supporting improved instruction across that system. Finally, it explains why states should not use kindergarten readiness assessment results as part of an accountability system for individuals—children, early learning providers, or teachers.*

As states continue to refine their accountability and assessment policies, it is important not to misuse KRA results in ways that draw incorrect conclusions from results and potentially discredit a tool that can help support instructional improvement.

* The analysis in this paper is based on the best available research as of the time of its publication. The field of kindergarten readiness is rapidly evolving, with new tools being developed and new research being conducted. We assume that some of the research and best practices discussed here may change significantly in the coming years, and we hope to update our analysis in response to new information as it becomes available.

I. KINDERGARTEN READINESS AND HOW TO ASSESS IT

A. WHAT IS KINDERGARTEN READINESS?*

The vast majority of states have definitions of children’s kindergarten readiness, but these definitions vary in content.⁵ Early learning experts agree that “kindergarten readiness” is a complex concept to define and measure for several reasons:

- > Readiness encompasses multiple domains of growth and development. State early learning standards generally take this into account, as states typically have early learning standards that address a wide range of domains.⁶
 - o The core areas of kindergarten readiness, as defined by the congressionally appointed National Education Goals Panel (1998), include the domains of language and literacy development, cognition, and general knowledge, including early science and mathematics concepts, approaches toward learning, physical well-being and motor development, and social and emotional development.⁷
- > There is high variability in what is considered the “normal range” of child development.⁸ Young children are constantly developing and acquiring new skills, but the rate at which early learners acquire new concepts and skills varies significantly between children.
- > Young learners develop skills and abilities across all of these developmental domains in a highly connected, interrelated manner, building confidence and expertise as new competencies are mastered. But children often progress unevenly within and across domains, meaning that multiple data points may be needed to get a sense of a child’s developmental trajectory.⁹
- > Child development and readiness are impacted by home, community, and prior early experiences, and are influenced by the blending of all of their previous cognitive, social-emotional, and behavioral experiences.

Given these complexities, definitions of kindergarten readiness are numerous and varied. However, formal state definitions and those from leading national organizations are similar in important ways. They often include not only key knowledge and skills that are part of a child’s readiness for school but also the readiness of schools, educators, caregivers, and communities to provide optimal learning environments that support children’s diverse and evolving learning and development needs. Still, it is important to note that families, educators, schools, and state policy may differ in the ways they determine whether a child is ready for school, and thus there is not one standard definition for kindergarten readiness.

* We use the term “kindergarten readiness,” but some states use the term “school readiness” instead.

EXAMPLES OF FORMAL KINDERGARTEN READINESS DEFINITIONS ¹⁰

KENTUCKY: “School readiness means that a child enters school ready to engage in and benefit from early learning experiences that best promote the child’s success. Families, early care and education providers, school staff and community partners must work together to provide environments and developmental experiences that promote growth and learning to ensure that all children in Kentucky enter school eager and excited to learn.”

GEORGIA: “School readiness must be defined within the context of families and how they live. It must be defined within the context of communities and the services they provide. And, it must be defined within the context of schools and their readiness for children. A child’s readiness for school is when: possible health barriers that block learning have been detected, suspected physical or mental disabilities have been addressed, enthusiasm, curiosity, and persistence toward learning is demonstrated, feelings of both self and others are recognized, social and interpersonal skills are emerging, communication with others is effective, early literacy skills are evident, and a general knowledge about the world, things, places, events, and people has been acquired.”

VIRGINIA: “A child’s school readiness involves all aspects of development, including the ability to follow directions, hold a crayon, speak understandably, identify shapes/letters/numbers, share with others, and separate from parents without being upset.”

Whether or not states have a formal definition of kindergarten readiness, they all have a de facto definition of kindergarten readiness in their learning standards. That is, all 50 states have learning standards that cover the pre-kindergarten years and are articulated to some degree with standards for kindergarten.¹¹ These learning standards from both early learning and K-12 provide a guide to educators about what kindergarten readiness should entail.

Importantly, once kindergarten readiness is defined, it can be measured. Indeed, states have made significant headway in measuring children’s kindergarten readiness.

B. ASSESSING KINDERGARTEN READINESS

The growing interest in tracking school readiness and the rising use of assessment with young children highlight the need to ensure that new assessments are being developed and implemented appropriately. Kindergarten readiness assessments—many of which are scored based on teachers' observations and interactions with children—must be administered in a manner that reflects best practice in assessing young children. States have made significant headway in implementing KRAs, frequently using pilot tests before expanding assessments statewide.¹²

According to leading experts in the field and studies such as the National Research Council's 2008 report on child assessment,¹³ best practices in assessing kindergarten readiness include the following:

- > ***KRAs should be designed appropriately for the population being evaluated—including being culturally and linguistically responsive—and developed based on the intended purposes of the assessment.*** Administrators should select an assessment process that is most appropriate for the purposes a state, district, or school may have; use more than one tool as necessary if there are multiple objectives; carefully assess the reliability and validity of the assessment in relation to these purposes; and ensure that the process is sensitive to cultural and linguistic contexts and other characteristics of the children being assessed.¹⁴ The assessment should also align with the state's learning standards.

NAMING KINDERGARTEN READINESS ASSESSMENTS

We use the term “kindergarten readiness assessment” because we think it is the best descriptor of these assessment tools: they measure whether children are ready for kindergarten. The term “kindergarten entry assessment” (KEA) is also used frequently, including by the US Department of Education in the regulations for the Race to the Top–Early Learning Challenge program.¹⁵ While we use the term “kindergarten readiness assessment” rather than “kindergarten entry assessment,” we use the department’s definition to define what these tools are. Specific names of assessments vary across states, as many states give their KRA (or KEA) state-specific branding.¹⁶

- > ***Assessments to support improved instruction should be conducted by teachers in the child's classroom environment throughout the kindergarten year.*** Frequent assessments conducted throughout the school year in a child's natural setting and with adults familiar to the child are a critical component of responsive instruction. This process can include collecting examples of children's work and observing children's interactions, behavior, and performance in the classroom. Kindergarten-age children demonstrate developing social and intellectual skills through play and social interactions with others.
 - o Given the rapid pace at which children learn concepts and skills, administering a KRA one time at the beginning of the school year will not provide teachers with the current information they need to support classroom instruction throughout the year. There are progress-monitoring tools and formative assessments created specifically for this purpose that can provide rich information about an individual child's growth over time. It is important that schools and teachers understand the differences between assessments administered once at the beginning of the year and assessments administered throughout the year, and use the results of each kind of assessment appropriately.
 - o Assessing children early in the school year also will not allow children to acclimate to their new setting. At the beginning of the school year, the kindergarten classroom may not be the natural environment that it is for most children by the end of the year, and the pace at which children adjust to their new surroundings can vary substantially.
- > ***It is important that information be collected on multiple areas of a child's development, including social and emotional growth.*** Assessments that focus only on one developmental domain do not capture other important areas of competencies that are closely related to each other. The assessment of a child's kindergarten readiness should involve either a tool that provides a comprehensive view of a child's development and learning or multiple focused tools that collectively provide the same breadth of perspective.¹⁷
- > ***Thoughtful consideration should be given to how information is gathered and shared in ways that reflect the diverse special learning needs and abilities, cultural heritage, and linguistic background of the children being assessed—and the resulting data should be used to design language-, culture-, and ability-responsive learning and assessment experiences.*** Data from the assessment should then be communicated effectively with parents on a child's growth and how families can reinforce classroom learning at home. Additionally, staff should be provided the resources that support their ability to effectively implement and share results of the assessment in culturally and linguistically responsive ways (see sidebar on ensuring equity in kindergarten readiness assessments).¹⁸

ENSURING EQUITY IN KINDERGARTEN READINESS ASSESSMENTS

Developing and implementing a kindergarten readiness assessment that accurately assesses young children coming from culturally, linguistically, and ethnically diverse backgrounds is critical considering the value of KEAs in guiding practice, identifying an individual child's abilities and needs, and informing policies and resource decision-making. Given the limited number of assessment instruments that have been designed or adapted with these populations in mind, experts have noted important considerations and challenges to ensuring validity and reliability in assessments.¹⁹ For example, young dual-language learners may be assessed in multiple languages, which raises issues around assessment availability in other languages and the cultural familiarity and language skills of the assessor. Assessors need to be fluent in both English and the preferred or dominant language of the child and have appropriate versions of the same assessment to be able to administer the assessment effectively. Moreover, assessors need training on the indicators of typical language development for children speaking another language and guidance on how best to report and interpret such results from multiple languages, and they must be knowledgeable about the children's cultural and community contexts.²⁰ These considerations, along with others experts have identified, must continue to be explored and addressed to ensure valid findings and well-guided policy decisions that are equitable and truly benefit young children.

The US Department of Education definition of kindergarten entry assessments in its Race to the Top- Early Learning Challenge Fund guidance offers similar principles. The guidance it has outlined explains that assessment should be administered to children during the first few months of their admission into kindergarten; cover all core Essential Domains of School Readiness as recommended by the National Education Goals Panel; be used in conformance with the recommendations of the National Research Council reports on early childhood; and be valid and reliable for its intended purposes and for the target populations and aligned to the Early Learning and Development Standards.²¹ The Early Learning Challenge guidance documents provide a framework for rating systems for early learning providers and kindergarten readiness assessment but do not contemplate integrating KRA results into those rating systems.

C. CURRENT STATE AND DISTRICT PRACTICES

Spurred by new legislative and funding opportunities such as Race to the Top–Early Learning Challenge,* a significant number of states have moved forward with exploring, developing, or implementing a kindergarten readiness assessment. While reports have estimated that 43 states have such assessments, states are at varying points in the implementation process — and also have meaningful differences in their actual assessment practices with regard to what tools are used, what areas of children’s learning is assessed, and the purposes of the data collected.²² For example:

- > Colorado has recently moved from a gradual phase-in to requiring all kindergarteners in the most recent 2015–2016 school year to be assessed using a list of approved teacher observational assessments that districts can choose from that cover developmental and academic domains. The Colorado kindergarten readiness assessment is used to inform instruction and appropriate support to students.
- > Vermont has been collecting information since 2000 on children’s readiness across five domains by surveying kindergarten teachers as part of a collaborative effort of the state Agency of Education and Departments of Children and Families and Health.²³
- > In Utah, districts determine the kindergarten readiness assessment tool of their choosing. The purpose of the tool is to support Utah local education agencies in assessing the needs of students enrolled in their kindergarten programs and report the progress of these programs to the Utah State Board of Education.²⁴ Most recently, there have been efforts in Utah that proposed replacing existing kindergarten readiness assessment—which, as noted, can vary from school to school—with a uniform state assessment as part of a bill to expand full-day kindergarten.²⁵ The proposal, HB 42, was not funded in the 2016 Utah legislative session.²⁶

In 2016, the US Department of Education released a report highlighting some of the successes and challenges experienced by four states—Maryland, Oregon, Pennsylvania, and Washington—in implementing kindergarten readiness assessment.²⁷ The report highlighted some key crosscutting findings:

- > In adopting KRAs, states consider numerous criteria, including “reliability and validity, appropriateness for all students, usefulness for informing classroom instruction, usefulness for informing early learning policies and program improvement, feasibility of administration by teachers, and cost.”

* As of this writing, the Department of Education has awarded Early Learning Challenge grants to three rounds of grantees, but there are no plans for an additional round. <http://www2.ed.gov/programs/racetothetop-earlylearningchallenge/awards.html>.

- > States provided professional development to support teachers learning about KRA and put in place systems to support consistent scoring.
- > District officials in turn sought to provide support to teachers implementing KRA, including new technology.
- > While schools and teachers intended to use KRA results to improve instructional practice and communicate with parents and preschool programs, at the time of the report those practices were largely not yet implemented.²⁸

As states implement their kindergarten readiness assessments, some are focused on making ongoing improvements to ensure the assessment system's strength. The Enhanced Assessment Grants program includes a consortium of states—Connecticut, Indiana, Massachusetts, Michigan, Nevada, and Ohio—that have come together with a shared goal of enhancing a multistate, state-of-the-art assessment system.²⁹ Maryland and Ohio did surveys and outreach, collecting feedback and additional recommendations from teachers to make improvements to their kindergarten readiness assessment that have led to reducing the number of items on the assessment to ensure efficiency and reliability of results.³⁰ Continuing to engage in ongoing improvement processes that include making necessary refinements informed by key stakeholders and best practice is important to maximizing the value of kindergarten readiness assessment.³¹

D. HOW KRAS ARE VALIDATED, AND THE VALIDATION OF TOOLS CURRENTLY IN USE

The best measure of the quality of a KRA is the extent to which it meets or exceeds various psychometric standards, including reliability and validity.

- > Reliability refers to the extent to which the KRA results in the same score when administered by different people to the same child, or to the same child at two time points in close succession.³²
- > Validity refers to the extent to which KRA scores accurately capture what they intend to measure and produce data that are appropriate for the purpose for which they were collected.³³

The validity of KRAs can be examined in several ways, including the ability of KRA scores to predict future outcomes (e.g., 3rd-grade test scores or grades) or the degree to which KRA scores are correlated with other valid and reliable assessments of the same domain of interest. Importantly, validation and psychometric studies should be conducted within a population similar to the one in which the KRA will be used; tools that are psychometrically sound within one population may or may not have similarly strong measurement properties when administered to a different population.

KRAs have the strongest reliability and validity when they are used in a manner consistent with the specific purposes for which the tools were designed and tested. An assessment's psychometric properties are compromised—that is, it will not do as good a job of reliably measuring what it says it does—when the assessment is used for a purpose for which it was not designed. Therefore, when testing the psychometric properties of a KRA, it is important to use the tool in the same way(s) and for the same purpose(s) as ultimately intended, including consideration of who assesses the child, when, how often, how much training and support is necessary, and how the results will be used.³⁴

As of this publication, some KRAs—including California's Desired Results Developmental Profile-School Readiness,³⁵ Illinois' Kindergarten Individual Development Survey,³⁶ and Maryland's Kindergarten Readiness Assessment³⁷—have already undergone rigorous psychometric testing and demonstrate high degrees of reliability and validity. Many others—including Arizona's Kindergarten Developmental Inventory, Delaware's Early Learner Survey, and Texas' Kindergarten Entry Assessment—are still in earlier stages of development and testing.

II. APPROPRIATE USES OF KINDERGARTEN READINESS ASSESSMENT

There are two broad categories for which KRAs can be used: improving practice at the school level and analyzing system-wide needs at the policy level. Existing KRAs have generally been designed for one of these two purposes.

A. USING KRA TO IMPROVE PRACTICE

- > A study of states that have put into effect some form of kindergarten readiness assessment found the majority indicated that improving individual instruction was a key purpose for implementing their assessments.³⁸ School-level studies have found similar results.³⁹ Within the broad category of improving practice, there are numerous specific efforts that KRAs can support:
- > **Developing Ready Schools.** KRA results can be used to develop “ready schools”—that is, schools that are well-designed to meet the needs of their incoming kindergartners.⁴⁰ Using KRAs in a school can inform a community conversation about the needs of the incoming kindergarten cohort, which can inform early learning program design and transition planning for children and families.⁴¹ The implementation of KRAs should also include teacher support to ensure quality and consistency.⁴²
- > **Supporting Aligned Teaching Practice and Program Planning and Improvement.** Kindergarten readiness assessments can support early learning and kindergarten classrooms in developing closer connections and aligning their practices by informing

teachers about the overall strengths and weaknesses of each cohort's skills at kindergarten entry in relation to what will be taught in kindergarten, in order to better target instruction and strengthen transitions.⁴³ In Oregon, the kindergarten readiness assessment has stimulated discussions about how early learning connects to later learning, and also ways to better serve children and families before kindergarten entry. For instance, Malheur County used the assessment results to make a call to action to enhance partnerships between local early learning providers and elementary schools—and in two communities of McMinnville and Gladstone, educators are utilizing assessment results as a part of their work to monitor progress of children's learning from pre-k through 3rd grade. The assessment data collected has also played an important role in informing Oregon's state and local level resource allocation, including a significant investment of \$100 million in early learning and family supports in the 2015 legislative session.⁴⁴

- > In addition to helping to direct investments, KRA results can be a useful tool for informing the continuous quality improvement of programs.⁴⁵ Early learning programs can benefit from receiving results of the KRA data by incorporating the information into their improvement practice efforts so that more children will arrive at school ready to learn. This may be especially useful data for Head Start programs, which are required under the Head Start Program Performance Standards to set school readiness goals.⁴⁶
- > **Delivering Individualized Instruction.** KRAs can also support teachers' understanding of children's learning and advance their knowledge of child development as they document each child's knowledge and competencies and identify a child's strengths and areas in need of improvement.⁴⁷ One Maryland teacher noted the kindergarten results helped to assess which students were demonstrating knowledge in academic areas and met developmental skills; teachers were also able to see differences between students who came from pre-kindergarten and those who did not.⁴⁸
- > **Supporting Teacher-Parent Partnerships.** Kindergarten readiness assessments can also foster teacher-parent partnerships. Families can be included as part of the process of sharing what they know about their child's learning and also given information about their child's development that can support learning needs at home. For example, California has included a parent-report component in its statewide assessment system; the report gives parents a measure of their child's development and learning on a continuum as it relates to a particular domain. Moreover, Illinois' kindergarten readiness assessment offers families the chance to share feedback on their child's progress across multiple domains that are later discussed at parent-teacher conferences. Washington state has also integrated a family component to its assessment, which connects teachers with parents to better understand the context and experiences of a child; teachers meet with parents to discuss important information such as the child's family structure,

living situation, and home language. These strategies not only engage families in the process but also, importantly, build early on a foundation for strong relationships between the home and school.⁴⁹

- > **Screening for Special Needs Assessment.** KRAs can provide useful information—along with results from screenings and other observations—to identify children who should be referred for evaluation of potential developmental delays or other special needs.⁵⁰ In addition, Lane County in Oregon is using aggregate kindergarten readiness assessment results to identify communities that would benefit from the expansion of an evidence-based intervention program that has been shown to positively impact the school readiness skills of children with developmental disabilities and delays, those in foster care, and children living in poverty.⁵¹

These practices can be very beneficial to educators and parents, but not all states have been able to implement them at scale. Aligning teaching practice with preschools⁵² and supporting parental engagement⁵³ are among the strategies that have proven to be challenging for states. Supporting these practices at the local level will be a key element of successful statewide KRAs. Moreover, many schools have tried to use KRA results for multiple purposes, and states may want to provide supports to schools to ensure that assessments are being used only for appropriate purposes.⁵⁴

Successfully implementing KRAs for any instructional purpose requires ongoing training and supports for teachers conducting the assessment; teachers administering the assessment should not only be trained on the tool(s) being used but also knowledgeable in child development, children's capabilities and cultural and linguistic backgrounds, and guiding principles for assessment.⁵⁵ Teachers also need training in how to use assessment data in order to make decisions and plan activities to support child improvement.⁵⁶ Conducting ongoing cycles of assessment to inform instruction is time consuming for teachers and resource intensive for schools, and it demands that school leaders protect the time necessary to do the work correctly if the results are to be meaningful. Thus, an intentional investment and recognition of the benefits by administrators and teachers is imperative for meaningful formative use of KRAs to improve instruction.

ASSESSMENT IN EDUCARE SCHOOLS



Utilizing data in support of continuous learning and improvement of practice is a core feature of Educare schools, which serve children from birth through kindergarten entry.⁵⁷ Educare schools collect, share, and engage in regular dialogue about data to identify strategies and inventions for strengthening program and teaching practices. The information gathered is then used to inform all quality improvements: in the classroom, in individual work with families, for program policies and systems, and for professional development. As part of this process, parents are also engaged in ongoing communication with program staff concerning their child's screenings and assessments and are provided with activities designed to enhance their child's development based on their child's assessment results. To support programs in modeling research-based program practices, Educare leaders play an integral role in fostering a culture of data utilization that facilitates ongoing use of real-time data within the program and promotes regular staff feedback loops to inform practice and professional development.

PRACTICAL STEPS FOR SUCCESSFUL KRA IMPLEMENTATION**FOR STATES**

States can set up districts to succeed by:

- Choosing an assessment tool in collaboration with local leaders
- Ensuring that the assessment tool is designed to meet the state's goals
- Aligning the assessment tool to the state's definition of kindergarten readiness
- Pilot testing assessments before rolling them out at scale
- Providing adequate time for a statewide rollout
- Communicating clearly about the intended uses of KRA results
- Maintaining systems to ensure reliability across raters
- Designing supports such as effective training and guidance for districts and teachers that ensure they have the capacity to succeed
- Integrating KRA results into state data systems and ensuring access to data
- Producing reports that make it easy for teachers to utilize and parents to understand KRA data

FOR DISTRICTS

Districts can set up teachers to succeed by:

- Providing training on how to conduct assessments, including considerations for young children with diverse backgrounds and needs
- Ensuring early childhood teachers understand child development
- Offering supports for utilizing assessment results to improve instruction
- Supporting elementary school principals in providing instructional leadership in developmentally appropriate instruction
- Avoiding unrealistic expectations of teachers that reduce their capacity to effectively administer assessments
- Using KRAs to facilitate communication between kindergarten and preschool teachers

FOR TEACHERS

Teachers can succeed in KRA by:

- Understanding the purposes for which the KRA is being administered
- Accessing supports and training on how to administer KRA
- Using KRA results to inform instruction and practice improvement
- Communicating with parents about KRA results

B. SYSTEM-WIDE MEASUREMENT OF KINDERGARTEN READINESS

In many states and cities, policymakers and advocates have developed a dashboard of metrics that measure the quality of an entire early childhood system.⁵⁸ These system-wide efforts are generally designed to measure whether the investments made in early learning are having the intended results and whether the state or city has made a sufficient investment in early learning to meet its overall educational goals.

There is wide acknowledgment that it is very difficult for schools to help students catch up when they fall behind their grade-level expectations.⁵⁹ Unfortunately, many children enter kindergarten already behind; socioeconomic gaps in development begin opening at birth and can be measured even in the first year of life.⁶⁰ Understandably, system leaders at the state and local levels believe that early learning can be an important strategy to close the kindergarten entry gap. But in order to address that gap efficiently and effectively, it is helpful for system leaders to first accurately quantify the gap. Aggregated data from kindergarten readiness assessments can play a critical role in that process.

Maryland, for example, has developed the Maryland Model for School Readiness, which provides parents, teachers, and early learning providers with a common goal and language for what children need to know and are able to do upon starting kindergarten. The Maryland model incorporates research-based instruction, age-appropriate assessment of children's learning, skills, knowledge, behaviors, and academic accomplishments across a wide variety of curriculum areas, and ongoing professional development for teachers and providers to implement their teaching practices. By systematically measuring how children are doing in kindergarten this effort has helped increase understanding by policymakers, school leaders, and the public of the need for children to receive support before they enter kindergarten; the evaluation data teachers collect is included in a report to the General Assembly on the level of kindergarten readiness statewide.⁶¹ This reporting has helped support increased public investment in early learning in Maryland—and scores have consistently risen over time.⁶² States that do not yet have kindergarten readiness assessment data available for this purpose have been trying to build systems like Maryland's so that policymakers, school leaders, and the public can have a sense of how many children are entering kindergarten at readiness levels.⁶³

Maryland has made new changes to administering the kindergarten readiness assessment to allow for more instructional time⁶⁴ after teachers pushed back on the assessment, arguing that it took too long and was developmentally inappropriate.⁶⁵ Still, system officials continue to find ways to maximize the data to make improvements. For example, school leaders in Washington County, Maryland, found they ranked last in kindergarten readiness across all counties in the state's latest report. School officials there used the results to shape county-wide plans for partnership and action that included developing a task force specifically focused on kindergarten readiness, making available additional preschool classes, and collaborating more effectively with the community.⁶⁶

In reporting data, states can present statewide aggregate data and data broken down by subgroups (income levels, demographic characteristics, and geographical regions to name three).⁶⁷ Examining these data can help illuminate a more detailed picture of the needs within a state and over time can be used to identify patterns and trends (both positive and negative) and inform the allocation of resources and supports. Importantly, reporting aggregate data in a way that is easy to understand and interpret provides public information without attaching to the KRA results any specific consequence to children, providers, or teachers. States and localities may decide to focus resources on populations with low KRA scores, or jurisdictions with low KRA scores may choose to increase their investment in early childhood—decisions that are informed by data but not automatic consequences of the data.

When administered by trained data collectors (teachers or others) according to the guidelines of the tool, KRAs are extremely well-suited to being used in this way. KRAs are generally much shorter than traditional assessments used by researchers to measure children’s learning and skills, and require less training to administer. Although this means that KRAs provide a less precise and detailed measurement of each individual child’s development, it also makes a KRA a good and economical choice when the goal is to use it to establish a “snapshot” of the kindergarten readiness skills of an entire cohort of children — or to identify population-level patterns and trends, as described above. Indeed, this is something that KRAs have been designed to do.

Importantly, reporting aggregate data in a way that is easy to understand and interpret provides public information without attaching to the KRA results any specific consequence to children, providers, or teachers.

EXAMPLES OF QUESTIONS THAT KRAS CAN HELP ANSWER

KRA results can be used to answer basic descriptive questions about the kindergarten readiness skills of students entering kindergarten. These might include:

- Is this year's kindergarten cohort more or less school ready than last year's cohort?
- Do some racial or ethnic groups demonstrate stronger kindergarten readiness skills than others?
- Do some neighborhoods or communities demonstrate stronger kindergarten readiness skills than others?
- How large is the gap in the kindergarten readiness skills demonstrated by students from low income families compared to their higher income peers?
- State- or district-wide, which skill areas tend to be strongest and weakest for incoming kindergarten students? How do these areas of strength and challenge differ across communities or student subgroups?

In addition, KRA results can be used to answer more targeted questions about initiatives that states or districts have already implemented or that they might implement in the future. These could include:

- Do students from communities in which we made substantial investments in a kindergarten transition program demonstrate stronger kindergarten readiness skills than those from communities in which we did not invest in these programs?
- Which communities might benefit most from additional kindergarten transition resources?



- Which communities might benefit most from additional early childhood services capacity or quality improvement supports?
- State- or district-wide, which skill areas might kindergarten teachers need to focus on most? What kinds of professional development and support might they need most to do so?

For example, the state of Oregon articulated the following uses for its aggregate KRA results:

"Statewide results are used to identify opportunity gaps in order to inform decision-making in allocating resources to the communities with the greatest need and to measure statewide progress in the years to come. District and school level results help districts and schools, in partnership with local providers of early learning services, better understand the strengths and gaps in selected domains of development for the population of Oregon's entering kindergarteners overall and by student subgroups. These results can be used to monitor patterns in district- and school-level data over time and identify opportunity gaps among student subgroups. This information can inform instructional decisions and be used to target professional development, resources, and supports on the areas of greatest need."⁶⁸

DISTRICT VS. STATE ASSESSMENTS

While some states have a single statewide KRA tool, other states require districts to use KRA without specifying a particular assessment.⁶⁹ In addition, many districts voluntarily administer a wide range of KRA tools, regardless of whether or not the state requires or provides a single KRA. If administered well, these district-selected KRAs can be very valuable to improving practice, but they do not provide a system-wide measurement of kindergarten readiness. One struggle for states attempting to implement statewide KRAs has been resistance from districts that see statewide KRA as duplicative.⁷⁰ States seeking statewide KRA data should be prepared to work with districts to chart a path to consistent statewide data that still minimizes the burden on districts, schools, and teachers.

C. CONCLUSION

When used as intended to inform practice and policy, KRAs are important components of high-quality early education. First, by providing teachers and schools with valuable information about individual children's learning and development, KRAs support program planning and individualized instruction, stronger connections between early childhood programs and K-12 schools, and partnerships between teachers and parents. Second, by providing districts and states with valuable information about the learning and development of large groups of children, KRAs support policymakers in tracking system-wide trends and patterns, identifying and addressing kindergarten readiness gaps, effectively allocating resources and supports, and making data-informed policy decisions.

III. MISUSES OF KINDERGARTEN READINESS ASSESSMENT RESULTS

There are three major potential misuses of kindergarten readiness assessment results that either have been implemented or considered by policymakers and practitioners. The first is using KRA results to hold individual children out of kindergarten; the second is using KRA results as an accountability measure for early learning providers; and the third is using KRA results as a factor in teacher evaluations. This section addresses each of those issues in turn.

A. CHILD-LEVEL CONSEQUENCES

For many years, leading experts in the field have cautioned against using kindergarten readiness assessment results to hold children out of kindergarten. The assessments were not designed for this purpose; while kindergarten readiness assessment results may be valid for multiple purposes, they are not valid for the high-stakes purpose of excluding children from kindergarten. Indeed, children who score poorly on readiness assessments may be the children who would benefit most from the kindergarten experience.⁷¹

The Early Learning Challenge prohibited using KRA results for this purpose, and districts implementing KRA under its grants have understood the importance of this limitation.⁷² Unfortunately, some schools do appear to use KRA results in this manner.⁷³ State training protocols and materials can play a valuable role in ensuring that this practice is avoided at the local level.⁷⁴

B. SITE-LEVEL ACCOUNTABILITY

The fact that states are considering using KRA results in accountability systems for early learning sites stems in part from a convergence of trends between K-12 accountability and early learning accountability. Increasingly, states are looking at how to incorporate measures of child success into their early learning accountability, which has long been a practice in K-12. This subsection provides a brief summary of those trends to help explain the context in which states are considering using KRA results for site-level accountability. It then explains the reasons why KRA results should not be used for site accountability: The KRAs themselves have not been designed for that purpose, and the practical reality of today's early learning system makes it impossible to utilize KRA results for accountability.

1. Trends in Site-Level Accountability

In recent years, states have been developing Quality Rating and Improvement Systems (QRIS) to measure the effectiveness of their early childhood providers, which frequently build on state child care licensing regulations.⁷⁵ To date these have been comprised primarily by what are

known as elements of “structural quality”: observable quality standards such as adult-child ratio, daily schedule, curriculum, materials, and health and safety standards.⁷⁶ One important trend in site-level accountability*—particularly in QRIS—is a move toward the inclusion of what are known as elements of “process quality”: professional practices often measured through external reviews of the quality of instruction and adult-child interactions. For example, the most recent reauthorization of the federal Head Start program included scores from the Classroom Assessment Scoring System observational measure of the quality of classroom organization, emotional support, and instructional support in the criteria for grantee recompetition.⁷⁷

Relatedly, the passage of the Every Student Succeeds Act (ESSA) in 2015 ushered in a new era of accountability for K–12 schools. Prior to the ESSA, schools were held accountable primarily based on the percentage of their students achieving proficiency on standardized tests; the ESSA changes the equation, requiring states to include new measures of school or student success.⁷⁸

The trends in K–12 and early learning program accountability are arguably converging on an approach in which schools and sites are held accountable for a mix of factors that includes measures of (1) educational quality and (2) child outcomes. A previous Ounce Policy Conversation paper by two of the authors of this paper, “A Framework for Rethinking State Education Accountability and Support from Birth Through High School and Changing the Metrics of Turnaround to Encourage Early Learning Strategies,” proposed an accountability framework based on those two major categories of metrics. In early learning educational quality measures are already a core part of accountability systems, and the Ounce paper argues in favor of a shift toward educational quality metrics – which includes “process quality” measures such as measures of professional practice (see sidebar on the benefits of professional practice metrics).[†]

* We use the term “sites” here because many publicly funded providers operate multiple sites, and accountability systems generally show results for each individual site.

† In this paper, the term “metrics” means data that captures an outcome, and “measures” means the method by which that metric is collected—the same definition used in the Ounce paper cited above.

BENEFITS OF PROFESSIONAL PRACTICE METRICS

Professional practice metrics are measures of educational inputs—such as the routines, instruction, relationships, and interactions between and among adults and children in classrooms and schools—that support learning. Some of the benefits of these metrics include:

- These types of measures can be used at any age of education, starting at birth. Current K-12 accountability systems rely heavily on student test scores and therefore place a disproportionate weight on the tested years (3rd grade and up). This means that the short-term incentives for school leaders will always pressure them to invest in those years as opposed to the earlier years. Under the ESSA, states can use professional practice metrics to create meaningful accountability measures for the K-2 years, which fundamentally changes the incentives acting on local leaders when making decisions about investment in kindergarten through 2nd grade.⁷⁹
- They can focus schools on the practices that really matter. The most important way that early learning schools can contribute to child development is through outstanding instruction. The quality of instruction (and the systems that support it) can be measured by external observation, but previous generations of early learning accountability systems typically did not include a sufficient focus on instruction; the trend in recent years has been to strengthen that focus.
- Relatedly, they provide actionable information. External reviews provide feedback on school systems and instruction, and the only way for schools to improve their performance is to improve their systems and the quality of their instruction. This is a significant departure from previous early learning accountability systems, which incentivized improvements that might not be tied to instructional quality; it is also a significant departure from previous K-12 accountability systems, which provided essentially no feedback to schools on how to improve instruction.

While education quality metrics can play a valuable role in accountability systems, improvements in instruction are ultimately intended to improve child outcomes. Accordingly, valid and reliable information about children’s learning and development across multiple domains is key to our understanding of education quality. Indeed, measures of child outcomes currently feature prominently in K–12 school accountability and may also have a role in early childhood accountability.⁸⁰ However, an ongoing challenge in school accountability is the issue of what child outcome metrics should be used prior to 3rd grade. Our “Framework” paper acknowledged that a great deal of work was likely necessary to develop better metrics for child outcomes in the pre-3rd-grade years.

Some states have considered filling this void by using kindergarten readiness assessment scores as a child outcome metric,* an understandable instinct given the paucity of measures and the prevalence of test scores as a child outcome metric in later years. On the surface it seems logical to use kindergarten readiness assessment results in site-level accountability for early learning programs. After all, school-level accountability should include child outcome measures, and kindergarten readiness assessments are the most prominent measurement of child skills and abilities in the early learning field.

But measuring skills and abilities is not always the same as measuring outcomes: In order to consider skills and abilities “outcomes,” it must be possible to reliably attribute them to the educational intervention being held accountable. The nature of how young children develop and how we currently assess their learning, as well as the complexity of our current early childhood system, make it impossible to reliably make this attribution for all incoming kindergarteners. Thus it is unwise to include kindergarten readiness assessment results as an outcome measure in site-level early learning accountability systems: Kindergarten readiness assessments have not been designed for that purpose, and it is not practical to administer kindergarten readiness assessments in a way that would make its results reliable for accountability purposes.†

* We use the term “child outcome metric” here noting that it can be used to describe both direct outcomes that are meaningful (e.g., graduation rates, avoiding incarceration, or higher lifetime earnings) and also indicators that are predictive of those direct outcomes. In either case, the metric should be objective and comparable across multiple settings. In both definitions, the metric is something that is (a) objective, (b) linked to positive life outcomes, and (c) comparable across children from different classrooms/schools.

† As noted previously, two of the authors have argued that early learning accountability should focus on the quality of early learning practice. Regenstein, E. and Romero-Jurado, R. (June 2014). “A Framework for Rethinking State Education Accountability and Support from Birth through High School,” 26. In that vein, we do think it would be appropriate for schools to be held accountable for implementing best practices in assessment—that is, for the quality of the teacher practices in administering assessments and acting on the results. That kind of accountability would be focused solely on the quality of professional practice, not on the scores obtained from children in the assessment.

YOU GET WHAT YOU PAY FOR

The research base on early childhood programs is clear: High-quality early learning increases the odds of positive long-term results, and lower-quality early learning does not.⁸¹ But many large-scale early learning programs are not designed or implemented in a manner likely to lead to improved long-term outcomes, frequently because they are not funded to meet the standards they are being held to.⁸² Thus, accountability systems for early learning programs will generally need to do more than just inform targeted assistance to underperforming programs—they will need to inform systemic changes that provide programs the supports and resources necessary to actually deliver quality.⁸³

2. Kindergarten Readiness Assessments Are Not Designed to Produce Results Suitable for Use in Accountability Systems

Assessment tools should be used for the specific purposes for which they were designed, and no current kindergarten readiness assessment was designed for the purpose of holding individual early learning sites accountable.* As discussed above, current KRAs have been validated for the purpose of informing instruction, identifying children with learning needs, and providing information to families, but they have not been validated for use as a program- or teacher-accountability tool. Because KRAs were designed to evaluate children's skills to help teachers target and individualize instruction, using KRA results for the entirely different purpose of assessing program accountability would be an invalid use of the KRA instrument.

According to the National Education Goals Panel, "using student and school-level data for accountability purposes requires a higher level of technical accuracy."⁸⁴ Assessment validity is achieved when instruments accurately measure the variables they were designed and intended to evaluate. Thus, the selection or design of an instrument should be based on its intended purpose.

* We express no opinion here as to whether this is possible or advisable.

The panel found that using assessment data for nonvalidated purposes—such as with high-stakes decision-making—provides misleading information. High-stakes assessments are those in which results are used to make impactful decisions, such as those used to determine children’s entry into kindergarten, funding or licensing for early learning programs, or teacher or school evaluations. The panel strongly recommends, “Before age 8, standardized achievement measures are not sufficiently accurate to be used for high-stakes decisions about individual children and schools. Therefore, high-stakes assessments intended for accountability purposes should be delayed until the end of third grade (or preferably fourth grade).”⁸⁵ Since the panel issued its report, other researchers have noted the extreme level of care necessary in designing high-stakes accountability and the challenges of including child assessment results from young children in high-stakes accountability⁸⁶—including at the school, principal, and teacher levels.

Given these challenges, to our knowledge at this time no state or other entity has ever even attempted to develop an assessment valid for high-stakes program accountability use with children under age 8—including the states that are using assessments for high-stakes purposes.

3. It Is Not Practical to Administer a Kindergarten Readiness Assessment in a Way That Would Make Its Results Reliable for Provider Accountability Purposes

If a state did develop a kindergarten readiness assessment suitable for use in accountability systems, it would still face multiple practical challenges in implementing that assessment in a way that produced reliable data. These include ensuring consistency among all kindergarten teachers in conducting ratings, safeguarding against incentives for teachers to exaggerate child assessment results, and developing a scoring scale that appropriately accounts for proficiency or growth in the early learning assessment environment.

a. Ensuring That Teacher-Reported KRAs Are Administered Consistently Would Require an Enormous Investment

If states take the time and energy to develop KRAs meant for early learning provider accountability, they would need to implement statewide systems to ensure reliability in how those assessments were administered and children’s learning evaluated and scored. This scoring reliability would have to look very different than the systems used to score existing accountability tests used in 3rd grade and up.

As discussed above, KRAs are typically teacher reported—that is, they are administered by individual kindergarten teachers observing and collecting children’s work samples, often based on state learning standards. Unlike with direct assessments (such as standardized tests) often used for older children, obtaining reliable information from teacher-reported assessments requires ensuring not only that the assessment tool itself is reliable but that the teachers who use it also meet reliability requirements to do so. That is, if KRA results were being used for program accountability purposes, then in fairness to the rated early learning providers—and to collect reliable, accurate data—states would have to have extensive training and reliability systems in place to ensure that all kindergarten teachers in the state score assessment results consistently. Otherwise, the inevitable inconsistency among teacher reporters could have dramatic effects on individual child scores,⁸⁷ and if those child scores are being used to rate the early learning sites children attended, that inconsistency among teacher reporters would lead to inaccurate and unfair ratings.

Because to date the primary purposes of existing kindergarten readiness assessments have not been high stakes, successfully implementing those assessments has required less-rigorous training and reliability standards, and thus there have not been statewide efforts to create such a high level of data reliability.⁸⁸ For example, in order for information about a child’s kindergarten readiness to be useful to a teacher or school, it matters less whether that information was collected in exactly the same way that another teacher in another part of the state collected equally useful information about a different child. For aggregated information about a group of children to be useful to a school district’s planning processes, it is important that the information was collected in the same way from each child within the district—but it is less important that the information be collected in exactly the same way as it is in another school district.

Of course, when the goal is to aggregate scores across districts to examine kindergarten readiness statewide, it becomes more important that teachers in all districts are closely following assessment administration guidelines. Still, in each of these examples, there will be some measurement error introduced because of slight variations in the ways teachers interpret the assessment tools and collect information about children. When kindergarten readiness information is used as described above to inform program improvement and understand system-wide needs, small amounts of measurement error are relatively inconsequential; the information is useful even when it provides only a ballpark estimate of children’s actual skills—surrounded by what is called a confidence interval, or the range within which a true score is most likely to fall.

However, when kindergarten readiness information is used for high-stakes accountability, this kind of measurement error becomes deeply problematic. When small differences in scores lead to highly consequential policy decisions, there is no room for error. To be valid for these high-stakes uses, the assessments themselves must be sensitive enough to capture small differences and improvements in skills and must be administered at the same time and in exactly the same way from child to child and teacher to teacher. To ensure such consistency across a large group of assessors is no small task. Developing a system of reliability training involving every kindergarten teacher in a state would require an enormous investment of funds and sophisticated training far more involved than any professional development currently being attempted in states.

In sum, if a KRA existed that could in theory be used for early learning provider accountability, implementing that assessment and collecting data for accountability purposes would require an unprecedented investment in ensuring consistency across assessors. That expensive effort to ensure consistency in rating would constantly be battling the fact that teachers know exactly which colleagues have a stake in the outcome of their assessments, meaning that community pressures could weigh heavily on them during the process. In a cost-benefit analysis of the investments necessary to implement a reliable scoring system for KRAs to use in program accountability, it appears likely that the costs would be extremely high and the benefits of consistent scoring would be unlikely to materialize—meaning that this investment likely would be seen by early learning professionals as having a high opportunity cost given other needs and priorities in early learning professional development. And even that assumes that having genuinely consistent KRA results could actually form the basis of a well-designed provider accountability system—which, as discussed further in section c below, is not the case.

b. Safeguarding Against Incentives for Teachers to Exaggerate KRA Results Would Require New Kinds of Assessment Tools

No matter how much training is provided, using the results of teacher-reported KRAs in accountability systems gives teachers incentives to adjust the data. Large-scale assessments currently used in K-12 accountability measure students' knowledge and skills directly, and test results are scored anonymously by people with no stake in their outcome. By contrast, most current kindergarten readiness assessments are designed such that scores consist of a reporter's best judgment of a child's knowledge and skills—which are necessarily reported by teachers who know the children and are part of the same community as the programs that previously served those children. Even the most expensive and rigorous system of teacher training on completing teacher-reported assessments may not completely counteract the natural human tendency to bias the

results based on the incentives created by how those results will be used. Similar challenges are well-documented within the context of large-scale implementation of an observational measure of classroom quality.⁸⁹

Efforts are underway to develop new assessment tools that may expand the kind of information that teachers can reliably collect from young children in classrooms.⁹⁰ For example, promising new technology-assisted tools have begun to make it easier for teachers to collect, organize, and score valid and reliable observational data and work samples that demonstrate children's knowledge and skills. In addition, another promising set of assessment tools—many aided by technology such as tablet computers—may help to eliminate some of the training burden and address concerns about incentives by enabling teachers to quickly assess children's skills directly, rather than relying on their own observations and impressions.

One example of technology-based assessment is the Measures of Academic Progress (MAP) for Primary Grades (MPG), an early childhood assessment aimed at kindergarten to grade 2 developed by the Northwest Evaluation Association. The MPG includes a computer-based skills checklist and screening assessment that covers the areas of reading and math and allows for measurement of student growth over time. More than ever before, data is being produced that is comparable across students in a way that teacher-reported kindergarten readiness assessment data is not.

Promising tools like these may eventually enable teachers to collect information about young children's learning with a level of rigor that was once only possible via highly trained and expensive external data collectors and researchers. However, computer-based direct assessments are not able to address some important domains of development (e.g., social-emotional and behavioral skills); using these tests in accountability systems could cause teachers to improperly narrow the focus of their instruction. Moreover, many of these new assessments were designed for formative or research uses and remain untested for use in high-stakes accountability. Thus, as of this publication, the potential validity of either set of technology-assisted tools for accountability purposes remains unknown.

While the results of computer-based assessments administered directly to students are more comparable across students than are the results of teacher-reported assessments, they are not appropriate for inclusion in program accountability. These new assessments do hold significant promise for improving teaching and instruction but also present significant challenges that must be addressed. States should support their development and testing while ensuring that they are used only for appropriate purposes.

c. Proficiency and Growth Measures Are Problematic as Early Learning Accountability Measures

If states decide to proceed with developing KRAs meant for early learning provider accountability, states would then be confronted with the same choice they face in K–12 accountability: Should the use of student assessment scores in school-level accountability focus on proficiency or growth? Proficiency refers to whether a child has met or exceeded specific minimum standards or expectations. In contrast, growth refers to how much a child’s assessment results have improved in a given period of time. In K–12, proficiency was the law of the land for many years, but the ESSA requires states to use both proficiency and growth measures.⁹¹ With early learners, however, both choices are unsatisfactory.

i. USING PROFICIENCY SCORES CREATES INAPPROPRIATE INCENTIVES AND IS DIFFICULT TO ATTRIBUTE

Consistent with how K–12 accountability worked under No Child Left Behind, states may think of KRA accountability based on proficiency, or having all children meet a predetermined level of skill. In this model, states would define a threshold for kindergarten readiness assessment; early learning providers would be given credit for children who achieve that threshold level.

Two states have made provision to use proficiency measures in their early learning accountability systems, Florida and Mississippi.

- > Florida formerly used results from its Florida Kindergarten Readiness Screener to calculate the kindergarten readiness rate for providers in its Voluntary Prekindergarten (VPK) Program, which measured how well a VPK provider prepared 4-year-olds for kindergarten.⁹² When the kindergarten readiness rate was last approved and calculated, the Florida State Board of Education adopted 70% as the minimum acceptable readiness rate for a VPK provider—meaning that 70% of a provider’s children had to be identified as ready based on both a teacher observational tool measuring a child’s skills and learning across multiple developmental domains and assessment in reading.⁹³ VPK providers who did not meet the readiness rate were subject to probation and were required to submit and implement an improvement or professional development plan.⁹⁴

During the 2016 legislative session, a bill was passed directing the state’s Office of Early Learning to not adopt any kindergarten readiness assessment rates for VPK during the 2014–15 and 2015–16 program years. This was based on changes in the kindergarten readiness assessment screening and on increased concerns about the not number and range of assessments in public schools.⁹⁵ The changes in law and policy have made identifying readiness rates for VPK sites an ongoing challenge.⁹⁶

- > Mississippi law requires the state's Department of Education to evaluate the effectiveness of early childhood collaboratives. The department is authorized to adopt a statewide kindergarten readiness screening tool, and if it chooses to do so is required to "adopt a minimum rate of readiness that each prekindergarten provider must meet in order to remain eligible for prekindergarten program funds." All children enrolled in state-funded pre-k must take the kindergarten screener even if they are not attending kindergarten in a public school.⁹⁷ The state KRA, which focuses solely on literacy, is scored on a scale of 300 to 900, and the department has set a kindergarten readiness benchmark score of 498 (which falls in the "Late Emergent Reader" category).⁹⁸ Statewide, 34.6% of children exceeded this benchmark in the fall of 2014.⁹⁹ A program's ratings are adversely affected if more than a third of its children fail to reach the 498 benchmark, or if children fail to show above-average growth.¹⁰⁰

Mississippi has released results of a Pre-Kindergarten Readiness Assessment administered to children at the beginning and end of the school year in 2014–15¹⁰¹ and 2015–16.¹⁰² Statewide almost every provider showed significant improvement from the beginning to the end of each academic year.¹⁰³

There are at least two major problems with this approach:

1. **Inappropriate incentives.** If early learning providers are being held accountable for their children reaching proficiency, that gives them a very strong incentive to seek out children who are likely to be proficient—and to not serve children who are unlikely to be proficient. This "creaming" problem is potentially huge in an early learning context where enrollment is voluntary and programs are not obligated to enroll any particular child, unlike K–12, where neighborhood schools are generally obligated to enroll any child who lives in their attendance area.¹⁰⁴ This means, perversely, that early learning might not serve children with the highest needs, who would most benefit from participating in a high-quality program. Given that in many states early learning has been positioned as a strategy for closing the readiness gap—or at least a strategy for lifting the achievement of children at risk of school failure—using proficiency for accountability would seriously undermine that goal.*

* As three of this paper's authors acknowledged in "Valuing the Early Years in State Accountability Systems Under the Every Student Succeeds Act," the presence of state-level incentives does not always drive local action. Even if state policy gives providers an incentive not to enroll the most at-risk children, many providers would enroll them anyway. Our argument is that it is bad policy for the state to create those disincentives to enrollment, and the provision of service to the most at-risk children should not be dependent on providers fighting their way past those disincentives.

Mississippi's law attempts to address this issue by requiring early learning providers to comply with antidiscrimination requirements. Providers can turn away children based on their usual eligibility guidelines as long as the guidelines themselves do not violate the antidiscrimination requirements.¹⁰⁵ Given the relatively fluid nature of enrollment at many early childhood providers, it is extremely unlikely that this language would be enough to prevent providers from consciously or unconsciously turning away children likely to bring down their scores—and it would likely take zealous oversight from the state to ensure that this is not an issue. While the program is too new for data to be available, it is unlikely that this language in the statute is enough to solve this problem.

- 2. Attribution.** Young children bring a wide range of experiences and competencies at kindergarten entry that reflect their cumulative experiences—experiences at home, in programs, and in their communities.¹⁰⁶ Many children with experiences in early learning settings have been served by multiple programs; the voluntary nature of programs and the nature of early learning funding—including child care policies that can make it difficult for children to remain continuously at the same program—are among the many reasons children switch programs. While continuity in a program is good for children, many children make frequent transitions in and out of programs, particularly in high-mobility families. So states would face the enormously complex challenge of figuring out just how to attribute “success” to the one or more programs that served a child prior to kindergarten entry.

This problem could in theory be addressed – but given the complex reality of how children develop and how they are served, whatever solutions are chosen would most likely be arbitrary. And even if a policy framework for attribution can be developed, actually implementing it would require state-level data systems far more advanced than what states have. Indeed, as of 2013, only one state (Pennsylvania) could even provide an unduplicated head count between its pre-k program and child care program.¹⁰⁷ Comprehensive data systems that include a wide range of early childhood programs could have significant positive impacts on states well beyond the program accountability context,¹⁰⁸ but the current status of state efforts suggests it will be many years before states have the kind of robust data systems that could track program entry and exit in a timely manner for accountability attribution purposes.

ii. STATES LACK THE INFRASTRUCTURE TO PROPERLY USE GROWTH SCORES

The ESSA's codification of a requirement that states use growth measures in school accountability was a conscious rejection of No Child Left Behind's focus on using proficiency to measure school effectiveness. One rationale for incorporating growth is that proficiency is often strongly correlated with socio-economic status; thus, it may improperly penalize schools serving high-poverty, high-need populations but doing great work to meet the needs of their students. For example, if a student entered a school in 5th grade reading at a 1st-grade level, and in one year that child improved to reading at a 4th-grade level, a proficiency measure would still show that the school "failed" that child. But using a growth measure would give the school credit for providing three years of improvement in a single year. Growth is seen by many as a better measure of how well a school actually did at helping a child learn and develop.¹⁰⁹

Measuring growth requires two assessments, one at the beginning of the period of accountability and another at the end. For purposes of holding early learning programs accountable KRA would by definition be the assessment that comes after the end of a child's pre-kindergarten experiences – meaning states would have to develop a complete system of assessments to be administered at the "beginning," which would have to be some defined point in the child's pre-kindergarten experience. This would mean:

- > Developing a whole new set of accountability assessments suitable for use at younger ages.
- > Creating a system of reliability training for teacher-reporters and data collectors that is designed to ensure consistent assessment and scoring procedures, which would need to include all early learning settings being held accountable and account for the fact that teachers administering the assessment might have a direct interest in the outcome.
- > Ensuring the assessments are administered every time a child changes programs (itself an enormous undertaking given child mobility rates) and developing a system that defines how much of the change in child assessment results is attributed to the program given the duration of the child's enrollment and other factors in the child's development.
- > Maintaining a state data system that could track all of the assessment information as children enter, exit, and switch programs.

All of these systems would need to be in place to fully implement a growth-oriented assessment-based accountability system in early childhood. States may choose to develop some of these systems for reasons other than using KRAs in accountability—indeed, more robust assessment systems and stronger state data systems might have positive impacts on statewide practice and policy. Until those systems are actually built out, however, a growth-oriented assessment-based accountability system is simply not possible.

School-Level Accountability: Summing Up

In sum, it is not appropriate to use kindergarten readiness assessments for provider-level accountability. We encourage systems to use current kindergarten readiness assessments to hold themselves accountable to providing sufficient supports and resources to programs. But while the temptation to also use kindergarten readiness assessment results as a child outcome metric in school accountability is real, the practical reality cautions against it.

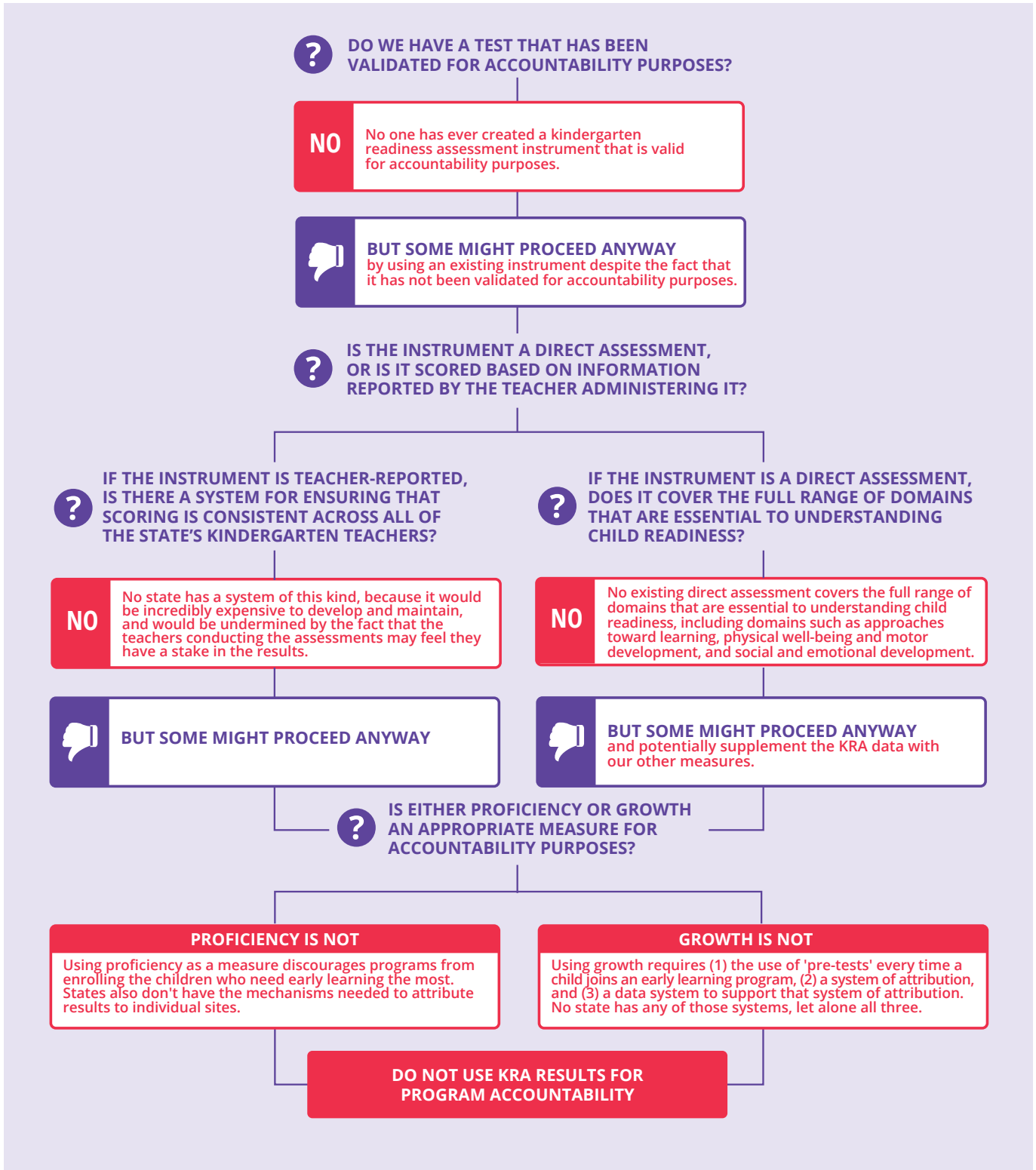
As proponents of using child outcome measures in accountability systems, we acknowledge that it would be easier to accept this conclusion if there were other readily available alternatives. At this time, however, there is no emerging consensus about what metrics of child success can be used appropriately in accountability systems for individual providers and schools prior to 3rd grade. As we have argued previously, we believe that chronic absenteeism is a promising measure of school quality and that more work is needed to develop additional metrics that can be used for this purpose.¹¹⁰ In the meantime, program-level accountability might be more appropriately based on practice-oriented metrics, which provide valuable insight to educators and can support continuous improvement while better metrics of child success are developed.

Figure 1 on the following page is a decision tree that provides a graphical representation of our argument. It is meant to be helpful for policymakers and advocates as a simple visual representation of the argument made here; each of the points made on it is addressed in greater detail in this paper. We acknowledge that all branches of the decision tree lead to the same conclusion—KRA results should not be used for early learning provider accountability—but hope that the graphic will help decision-makers understand why that is the case.



FIGURE 1. KINDERGARTEN READINESS ASSESSMENT (KRA) FOR ACCOUNTABILITY PURPOSES

A decision tree in which many incorrect decisions are needed to end up using KRA results for program accountability.



C. TEACHER EVALUATION

The third potential misuse of KRA results we address here is in teacher evaluation. The central role that assessment results play in state-level accountability for schools was reconfirmed by the ESSA.¹¹¹ But the statute does not require assessment results to be a part of teacher evaluation systems and indeed specifically prohibits the US Secretary of Education from adding that requirement administratively.¹¹² Since the 2009 report “The Widget Effect” called attention to the fact that teacher evaluation systems generally gave positive ratings to all teachers even in low-performing schools,¹¹³ most states and districts have worked to integrate student assessment results into teacher evaluation systems.¹¹⁴ The shift toward including assessment results in teacher evaluation systems accelerated with the US Department of Education’s process for granting waivers from the accountability requirements of No Child Left Behind, which required states to use student growth as a “significant factor” in teacher evaluations.¹¹⁵ However, the passage of the ESSA formally shifts the question of whether to use assessment results in teacher evaluation back to individual states and districts, some of which remain committed to the practice while others will likely change direction with this new flexibility.¹¹⁶

Utilizing student data in teacher evaluation systems is enormously challenging for teachers in grades prior to 3rd grade, when accountability testing begins.¹¹⁷ In states and districts that use assessment results in teacher evaluations, there may be a temptation to use kindergarten readiness results to evaluate kindergarten or preschool teachers. The most basic problem with this approach is that as with school accountability, existing kindergarten readiness assessments have not been validated for this purpose, and therefore the results cannot currently be used in this way. But even if such a validated assessment was created, as with school accountability there would still be a number of practical obstacles to implementation.

1. KRAs Do Not Measure the Growth Produced by Pre-k Teachers

Because KRAs are by definition conducted at the beginning of kindergarten, this one “snapshot” could theoretically be used to measure children’s proficiency at this point in time. But in teacher evaluation, assessment data are typically used to show growth, not proficiency. That is, teachers are evaluated in part on how much growth they were able to produce in students over the course of their year together.¹¹⁸

In theory, KRA results could be used to show growth for pre-k teachers if there were a baseline assessment prior to the pre-k year, although as noted above we are not aware of any state with such an assessment. However, even if states had such an assessment, this approach would still be ill advised because the assessments used to evaluate teachers should be administered during the year the teacher is with the child. Waiting until the

beginning of kindergarten to measure pre-k teaching means that well-documented summer learning loss would result in underestimating how much children learned in pre-k. To make matters worse, the severity of this underestimate is likely to vary systematically by student disadvantage: Research shows that children from disadvantaged backgrounds tend to lose more over the summer than do their more-advantaged peers.¹¹⁹

Moreover, this approach would have the same implementation challenges described in the sections above on school accountability. Even the basic data linkage needed to tie KRA results back to a pre-k teacher is not currently available in many states,¹²⁰ and without that linkage, KRA results cannot be connected to the teacher being evaluated. States are working to strengthen the linkages between early learning and K-12 data but until those systems are developed this essential connection is missing.

Finally, it is important to note that the teaching workforce in pre-k looks very different in many states from the K-12 teaching workforce. Many children who attend preschool do so in private-provider settings or in Head Start, settings that are not necessarily connected in any formal way to the schools where their students ultimately attend kindergarten and that would likely fall outside of state teacher evaluation requirements. Indeed, only 23 states require that preschool teachers have a bachelor's degree.¹²¹ Even assuming that there is an instructional leader in these programs qualified to administer teacher evaluations, there are significant challenges to evaluating pre-k teachers in the same manner as K-12 teachers when they have not been trained and supported in the same way. We hope that in time, early learning teachers will be treated with a level of professionalism commensurate with that of K-12 teachers and that teacher evaluation systems will evolve to reflect that.

2. KRAs Do Not Reliably and Comprehensively Measure the Growth Produced by Kindergarten Teachers

Because KRAs measure readiness at the beginning of the kindergarten year and not growth over the course of that year, KRA results (at least on their own) cannot be used for kindergarten teacher evaluations. Importantly, this problem cannot be solved by simply readministering KRAs at the end of kindergarten. Because KRAs are designed to measure readiness, they are scaled such that well-prepared children will score very high on the assessment in the fall of their kindergarten year. As those children continue to learn new skills during the kindergarten year, they quickly "outgrow" the KRA. That is, they will learn kindergarten-level skills that KRAs are not designed to measure and thus that KRA scores cannot adequately reflect. For example, on a hypothetical scale of 1 to 10, a well-prepared child who enters kindergarten with a score of 9 or 10 is likely to be at 11 (or 15) by the

middle of the kindergarten year, but because the scale only goes up to 10, he or she (and thus his or her teacher) cannot receive credit for those additional skills learned. Statisticians call this measurement problem a "ceiling effect."

Another growth-related challenge in using KRA results as part of an accountability system is that readiness scores are poor predictors of academic achievement in elementary school.¹²² Importantly, this may stem from the very reason that we measure readiness in the first place. As described previously, because KRAs are intended to improve and target instruction, they are inherently a poor predictor of later academic achievement. That is, one goal of KRAs is to help teachers, schools, districts, and communities meet the learning needs of each kindergartener by providing additional support on foundational skills to children who enter kindergarten farther behind. Success of these efforts is thus measured by the extent to which the readiness disparities that existed at the beginning of the year have been reduced or eliminated by the end. Therefore, those children who enter kindergarten at risk may show greater growth on the skills measured by KRAs than their peers who entered kindergarten better prepared, precisely because they scored lower initially.

These challenges are present whether schools are using teacher-reported assessments that rely on teachers' observations and impressions, technology-aided direct assessments that make it easier for teachers (or others) to assess children's knowledge and skills directly, or a combination of both. There are, however, additional barriers to integrating KRAs into teacher evaluations of kindergarten teachers that are unique to each type of assessment.

a. Teacher-Reported KRAs Will Never Be Appropriate for Evaluation Purposes

The problem with using KRAs that rely on teachers' observations and perceptions of children's skills for evaluation purposes is fundamental. If a teacher is conducting an assessment in which he or she has a stake in the outcome, the likelihood of obtaining accurate and reliable data may drop substantially.¹²³ Instituting this kind of evaluation system would give teachers a strong incentive to (consciously or unconsciously) underestimate a child's skills and development at the beginning of the year and then overestimate them at the end of the year, thus artificially inflating estimates of children's learning.

Some states have attempted to account for this issue of using teacher-reported assessments for accountability by avoiding using the assessment scores directly. For example, Delaware and Kentucky have each developed a K-12 teacher evaluation system that uses a rubric-based, collaborative peer and/or supervisor evaluation process to appraise teachers' performance on several metrics, including students'

growth on teacher-designed and/or -administered assessments or “growth goals” (most of the other metrics pertain to observed elements of process quality) in untested grade levels or subjects.¹²⁴ However, more research is needed to determine whether these systems indeed reduce incentives for teachers to overestimate their students’ learning on teacher-scored assessments such as KRAs.

b. Technology-Aided Direct KRAs Are Not Yet Valid for Teacher Evaluation and Will Need to be Used Carefully If They Are

Technology-enabled KRAs that directly assess children’s skills are not yet specifically designed and validated for use in teacher evaluation systems. If this option ever does become available,* then as with any teacher assessment data, it will be important to place these data in proper context. For example, using KRAs that focus solely on reading or math proficiency—arguably domains that are among the easiest to assess directly in young children—could pressure teachers to focus on those areas to the exclusion of other critical areas of development, including social and emotional development.¹²⁵ Moreover, in order to use growth scores, the KRAs would have to be designed specifically to measure growth and then be administered more than once during the kindergarten year. If KRAs that are appropriate for use in kindergarten teacher evaluation are ever created, states choosing to use them should have a thoughtful process of engagement with key stakeholders—particularly the teachers themselves and the principals who will be evaluating them—to discuss how those results can be used to improve the quality of kindergarten education rather than detracting from it.

3. Teacher Evaluation: Summing Up

Kindergarten readiness assessments have a valuable role to play in teacher professional development. Indeed, KRA results can form the basis of important conversations between principals and teachers about how children are progressing.¹²⁶ However, at the time of this publication, KRA results are not validated for use as a formal component of the scoring of teacher evaluations. Furthermore, if and when they are, it will be a very heavy lift for most states to create the systems needed to use KRAs appropriately in teacher evaluation systems.

* We express no opinion here as to whether this is possible or advisable.

CONCLUSION

We are strong supporters of kindergarten readiness assessments and of early learning accountability. But we do not believe it is appropriate for kindergarten readiness assessments to be a part of school- or teacher-level early learning accountability.

The results of KRAs are sufficiently precise and reliable enough to inform how teachers and school leaders work with children. Indeed, KRAs are a critical strategy for improving the quality of teaching and other supports in early learning and are also valuable as a system-wide measure of progress. However, they are not validated for external accountability purposes. School-level accountability for early learning should focus on measures of practice quality and valid child outcome metrics, and the field should continue to work toward improved metrics that will fill the gaps that fuel temptation to misuse the results of kindergarten readiness assessment.

In the meantime, KRAs that are implemented effectively and used for appropriate purposes can continue to provide policymakers, school leaders, teachers, and parents with valuable information about young children's learning and development. At the local level, they can provide educators with information they can use to improve instruction in their own classrooms and build relationships across educational settings. At a system level, they can provide valuable data about how children are doing when they enter kindergarten, allowing for more thoughtful system-wide investment and regional and local targeting of resources. States should continue to implement them in ways that help achieve important educational goals. Policymakers should facilitate that work by continuing to support and fund kindergarten readiness assessment efforts and by ensuring that the results of those assessments are used only as intended.

The authors are grateful to their many colleagues at the Ounce of Prevention Fund who reviewed and commented on drafts of this paper. The authors are also thankful for input from their external reviewers: Laura Bornfreund, Rachel Canter, Gerry Cobb, Lori Connors-Tadros, Shayna Cook, Libby Doggett, Chrisanne Gayl, Rolf Grafwallner, Lauren Hogan, Theresa Hawley, Robert Hull, Erica Lee, Kelly Maxwell, Rick Mockler, Tom Schultz, Catherine Scott-Little, Kyle Snow, Sarah Trumble, Yasmina Vinci, Albert Wat, Anna White, and Hiro Yoshikawa. The views expressed in this paper are those of the authors and do not necessarily represent those of the reviewers.

Join the Conversation at www.theOunce.org/policyconversation. Ounce Policy Conversations are published by the Ounce of Prevention Fund to stimulate and facilitate discussion on topics of interest in early learning policy. We welcome feedback on this paper, and may address that feedback in subsequent versions of this paper or in other papers. If you have comments or questions, please visit our webpage at www.theOunce.org/policyconversations or contact Elliot Regenstein at eregenstein@ounceofprevention.org

ENDNOTES

- 1 The White House Office of the Press Secretary. (December 10, 2014). "Fact Sheet: Invest in US: The White House Summit on Early Childhood Education." <https://www.whitehouse.gov/the-press-office/2014/12/10/fact-sheet-invest-us-white-house-summit-early-childhood-education> (accessed January 21, 2017); Daily, S. (October 2014). "Initiatives From Preschool to Third Grade: A Policymaker's Guide." Education Commission of the States. <http://www.ecs.org/docs/early-learning-primer.pdf>; National Governors Association. (n.d.). "Early Learning from Birth through Third Grade." <http://www.nga.org/cms/home/nga-center-for-best-practices/center-divisions/center-issues/page-edu-issues/early-learning-from-birth-throug.html>; National Conference of State Legislatures. (January 2, 2015). "Early Care and Education 2014 Legislative Action." <http://www.ncsl.org/research/human-services/early-care-and-education-2014-legislative-action.aspx>.
- 2 Schweinhart, L., et al. (2005). "The High Scope Perry Preschool Study Through Age 40: Summary, Conclusions, and Frequently Asked Questions." High/Scope Educational Research Foundation. http://www.highscope.org/file/Research/PerryProject/specialsummary_rev2011_02_2.pdf; Isaacs, J.B., and Roessel, E. (2008). "Impacts of Early Childhood Programs." Brookings Institute. <http://www.brookings.edu/research/papers/2008/09/earlyprograms-isaacs>; Barnett, W.S. (2008). "Preschool Education and Its Lasting Effects." National Institute for Early Education Research. <http://nieer.org/wp-content/uploads/2016/08/PreschoolLastingEffects.pdf>; Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M.R., Espinosa, L.M., Gormley, W.T., ... Zaslow, M.J. (October 11, 2013). "Investing in Our Future: The Evidence Base on Preschool Education." Policy brief. Society for Research in Child Development and Foundation for Child Development. <http://fcd-us.org/resources/evidence-base-preschool>; Regenstein, E., Marable, B., and Britten, J. (2016). "Starting at Five Is Too Late: Early Childhood Education and Upward Mobility." In Petrilli, M.J. (2016). *Education for Upward Mobility*. Lanham, MD: Rowman & Littlefield.
- 3 Daily. "Initiatives"; Meisels, S. (2006). "Accountability in Early Childhood: No Easy Answers." Herr Research Center for Children and Social Policy, Erikson Institute. <http://www.erikson.edu/research/accountability-in-early-childhood-no-easy-answers/>; Regenstein, E., Romero-Jurado, R., Cohen, J., and Segal, A. (September 22, 2014). "Changing the Metrics of Turnaround to Encourage Early Learning Strategies." Policy Conversation No. 4, Version 1.0. Ounce of Prevention Fund and Mass Insight Education. 21. <http://www.theounce.org/what-we-do/policy/policy-conversations>; Regenstein, E. and Romero-Jurado, R. (June 2014). "A Framework for Rethinking State Education Accountability and Support From Birth Through High School." Policy Conversation No. 3, Version 1.0. Ounce of Prevention Fund. <http://www.theounce.org/what-we-do/policy/policy-conversations>.
- 4 The ACT and SAT exams used in some state accountability systems were not designed primarily for school accountability but instead for student admission into college. States can use these exams in their accountability systems if they pass a peer review process through the US Department of Education. Gewertz, C. (January 4, 2016). "Will States Swap Standards-Based Tests for SAT, ACT?" *Education Week*. <http://www.edweek.org/ew/articles/2016/01/06/will-states-swap-standards-based-tests-for-sat.html>.
- 5 Connors-Tadros, L. (March 13, 2013). "Overview of State Kindergarten Readiness Definitions." Center on Enhancing Early Learning Outcomes (CEELO) presentation to the Arkansas Kindergarten Readiness Indicators Committee. <http://ceelo.org/wp-content/uploads/2013/08/CEELO-AR-SchoolReadiness.pdf>; Texas Early Learning Council. (September 2011). "Defining School Readiness. National Trends in School Readiness Definitions." <http://earlylearningtexas.org/media/10138/trends%20in%20school%20readiness%20final%2011-1.pdf>; Rhode Island KIDS COUNT. (February 2005). "Getting Ready: The School Readiness Indicators Initiative: A 17 State Partnership. Executive Summary." <http://www.rikidscount.org/Portals/0/Uploads/Documents/Early%20Learning/Getting%20Ready/Executive%20Summary.pdf>. In addition to the noted resources, a review of individual state early learning websites and affiliated reports was conducted to assess the prevalence and type of individual state definitions of school readiness.

- 6 Early Childhood Education Research Alliance. (April 24, 2013). "Implementing Early Learning Standards: Lessons From Research and Practice." Regional Educational Laboratory Northeast and Islands, Bridge Webinar. <http://www.relnei.org/wp-content/uploads/2013/06/ECEA-4-24-13-QA.pdf>; Daily, S., Burkhauser, M., and Halle, T. (June 17, 2010). "A Review of School Readiness Practices in the States: Early Learning Guidelines and Assessments." *Child Trends: Early Childhood Highlights*, Vol. 1, Issue 3. <http://www.childtrends.org/wp-content/uploads/2013/05/2010-14-SchoolReadinessStates.pdf>. Education Commission of the States. (March 2014). "50-State Comparison: Kindergarten Standards—General Info." <http://ecs.force.com/mbdata/mbquestRT?rep=Kq1410>.
- 7 Shepard, L., Kagan, S.L., and Wertz, E. (February 1998). "Principles and Recommendations for Early Childhood Assessments." Goal 1 Early Childhood Assessments Resource Group, National Education Goals Panel. <http://govinfo.library.unt.edu/negp/reports/prinrec.pdf>. For a summary of skills measured in these domains, see Golan, S., et al. "Case Studies of the Early Implementation of Kindergarten Entry Assessments." US Department of Education. 4. <https://www2.ed.gov/rschstat/eval/disadv/kindergarten-entry-assessments/report.pdf>.
- 8 Meisels. "Accountability"; National Association for the Education of Young Children (NAEYC). (n.d.). "12 Principles of Child Development and Learning That Inform Practice." <http://www.naeyc.org/dap/12-principles-of-child-development>; NAEYC. (2009). "Developmentally Appropriate Practice in Early Childhood Programs: Serving Children From Birth to Age 8." Position Statement of NAEYC. <http://www.naeyc.org/files/naeyc/file/positions/PSDAP.pdf>.
- 9 Snow, K. (December 2011). "Developing Kindergarten Readiness and Other Large-Scale Assessment Systems: Necessary Considerations in the Assessment of Young Children." NAEYC Center for Applied Research. http://www.naeyc.org/files/naeyc/file/research/Assessment_Systems.pdf; Kamler, C., Moiduddin, E., and Malone, L. (April 2014). "Using Multiple Child Assessments to Inform Practice in Early Childhood Programs: Lessons From Milpitas Unified School District." *Mathematica Policy Research*. http://www.mathematica-mpr.com/~media/publications/PDFs/education/Multiple_assessments_ib.pdf.
- 10 Kentucky Governor's Office of Early Childhood. (n.d.). "Preparing for Kindergarten." <http://kidsnow.ky.gov/School%20Readiness/Pages/default.aspx>; Georgia Department of Early Care and Learning. (n.d.). "School Readiness." <http://www.dec.ga.gov/prek/schoolreadiness.aspx>; Virginia Performs. (n.d.). <http://vaperforms.virginia.gov/indicators/education/schoolReadiness.php>.
- 11 Education Commission of the States. "50-State Comparison"; Regenstein, E. (2013) "Considering a Multistate Approach to Early Learning Standards." *Policy Conversation No. 2, Version 1.0*. Ounce of Prevention Fund. 5-6. <http://www.theounce.org/what-we-do/policy/policy-conversations>; Scott-Little, Catherine. "Early Learning Standards: Variations Across the States and Issues to Consider." 25. Presented at Listening and Learning About Early Learning Meeting, May 11, 2010. <http://www2.ed.gov/about/inits/ed/earlylearning/scottlittledock.pdf>.
- 12 Golan. "Case Studies."
- 13 National Research Council. (2008). *Early Childhood Assessment: Why, What, and How*. Committee on Developmental Outcomes and Assessments for Young Children, Snow, C.E., and Van Hemel, S.B., eds. Board on Children, Youth, and Families, Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press. <http://www.nap.edu/catalog/12446/early-childhood-assessment-why-what-and-how#toc>; Golan, S., Petersen, D., and Spiker, D. (December 8, 2008). "Kindergarten Assessment Process Planning Report." SRI International. http://www.k12.wa.us/WaKIDS/pubdocs/KindergartenAssessment_SRIreport.pdf; Hanover Research. (December 2013). "Kindergarten Entry Assessments: Practices and Policies." <http://www.hanoverresearch.com/media/Kindergarten-Entry-Assessments-Practices-and-Policies.pdf>.

- 14 Schultz, T. (July 2011). "Moving Forward With Kindergarten Readiness Assessment Efforts." Council of Chief State School Officers. http://www.ccsso.org/Resources/Publications/Moving_Forward_with_Kindergarten_Readiness_Assessment_Efforts.html; NAEYC and National Association for Early Childhood Specialists in State Departments of Education (NAECS/SDE). (November 2003). "Early Childhood Curriculum, Assessment, and Program Evaluation: Building an Effective, Accountable System in Programs for Children Birth Through Age 8." Joint position statement, NAEYC and NAECS/SDE. <http://www.naeyc.org/files/naeyc/file/positions/pscape.pdf>; Snow. "Developing Kindergarten Readiness"; Scott-Little, C., and Niemeyer, J. (2001). "Assessing Kindergarten Children: What School Systems Need to Know." SERVE. http://earlysuccess.org/sites/default/files/website_files/files/2012-10-Inter-State-Discussion-Time-Doc-2.M.Boyajian.pdf.
- 15 Federal Register, Vol. 78, No. 169, 53996. <https://www.gpo.gov/fdsys/pkg/FR-2013-08-30/pdf/2013-21139.pdf> p. 6.
- 16 Schilder, D., and Dahlin, M. (June 2014). "Considerations for 'Rebranding' Kindergarten Readiness Assessments." CEELO. http://ceelo.org/wp-content/uploads/2014/06/ceelo_fast_fact_kea_branding.pdf.
- 17 Daily. "Review of School Readiness Practices."
- 18 NAEYC. "Early Childhood Curriculum."
- 19 Espinosa, L.M. (May 2010). "Assessing Young Dual Language Learners: Challenges and Opportunities." <https://www2.ed.gov/about/initiatives/earlylearning/espinosadeck.pdf>; Snow. "Developing Kindergarten Readiness"; Barrueco, S., and Lopez, M.L. (n.d.). "Assessing Dual Language Learners: Critically Examining Our Measures." http://www.futureofchildren.org/futureofchildren/events/enhancing-practice-with-e/barrueco_lopez.pdf.
- 20 Snow. "Developing Kindergarten Readiness"; US Department of Health and Human Services and US Department of Education. (June 2016). "Policy Statement on Supporting the Development of Children Who Are Dual Language Learners in Early Childhood Programs." https://www.acf.hhs.gov/sites/default/files/eecd/dll_policy_statement_final.pdf; Golan. "Kindergarten Assessment."
- 21 US Department of Education and US Department of Health and Human Services. (July 2011). "Race to the Top—Early Learning Challenge." <http://www.ed.gov/sites/default/files/rtt-elc-draft-execsumm-070111.pdf>.
- 22 Weisenfeld, G. (October 2016). "Using Teaching Strategies GOLD® Within a Kindergarten Entry Assessment System." CEELO FastFact. CEELO. http://ceelo.org/wp-content/uploads/2016/10/ceelo_fast_fact_gold_as_KEA_10_16.pdf; Weisenfeld, G. (January 2017). "Kindergarten Entry Assessments: Common Measures and Consortia Efforts." CEELO FastFact. CEELO. http://ceelo.org/wp-content/uploads/2017/01/ceelo_fast_fact_kea_state_scan_2017_01_for_web.pdf.
- 23 Schimke, A. (August 13, 2014). "New Kindergarten Assessments Voluntary This Year, Mandatory Next Year." Chalkbeat. <http://www.chalkbeat.org/posts/co/2014/08/13/new-kindergarten-assessments-voluntary-this-year-mandatory-next-year/>; Vermont Agency of Education. (n.d.) "Kindergarten Readiness." <http://education.vermont.gov/sites/aoe/files/documents/edu-early-education-ready-for-kindergarten-report.pdf>. (For the 2015-2016 school year, Vermont made modifications to its kindergarten readiness assessment which included adopting a new measurement instrument, changing scoring methods and criteria for "readiness", and making revisions to survey questions across domains.); Scott-Little, C., and Maxwell, K. (n.d.). "Kindergarten Entry Assessments in Early Learning Challenge Grants: Charting a Course for a Successful Application You Can Live With." Webinar presentation for Early Learning Challenge Collaborative. http://www.elccollaborative.org/illinois/doc_download/155-kea-in-early-learning-challenge-grants-webinar-presentation.html; Connors-Tadros, L. (February 2014). "Information and Resources on Developing State Policy on Kindergarten Entry Assessment (KEA)." CEELO. http://ceelo.org/wp-content/uploads/2014/02/KEA_Fast_Fact_Feb_11_2014_2.pdf; CEELO. (n.d.). "State by State—Kindergarten Assessment." <http://ceelo.org/state-information/state-map/>.

- 24 Center on Standards & Assessment Implementation. (n.d.). "State of the States—Pre-K/K Assessment." <http://www.csai-online.org/sos>; Utah State Board of Education. (n.d.). "Assessment and Accountability: Kindergarten." <http://www.schools.utah.gov/assessment/Kindergarten.aspx>; Education Commission of the States. (March 2014). "Kindergarten Entrance Assessments." <http://ecs.force.com/mbdata/mbquestRT?rep=Kq1407>; Utah Administrative Code R277-489 (<http://www.rules.utah.gov/publicat/code/r277/r277-489.htm>).
- 25 Jacobsen, M. (November 6, 2015). "New Assessment Possible for Utah Kindergartners." KSL News. <http://www.ksl.com/?sid=37248953&nid=148>.
- 26 Utah Education Association. (March 2016). "2016 UAE Legislative Summary." http://www.myuea.org/issues_action/uea_under_the_dome/under_the_dome_archive/2016_archive/2016_legislative_summary.aspx.
- 27 Golan. "Case Studies."
- 28 Ibid., xi-xv.
- 29 Connors-Tadros. "Information and Resources"; Center on Standards & Assessment Implementation. "State of the States."
- 30 Ohio Department of Education. (n.d.). "Ohio's Kindergarten Readiness Assessment." <http://education.ohio.gov/Topics/Early-Learning/Kindergarten/Ohios-Kindergarten-Readiness-Assessment>.
- 31 Stakeholders want to be consulted in the development process, and in some instances local officials perceive statewide KRAs as top-down mandates when they were not included in its development. Golan. "Case Studies." 36.
- 32 Barnett, W.S., Riley-Ayers, S., and Francis, J. (November 2015). "Measuring Child Outcomes in the Early Years." CEELO Policy Brief. CEELO. 6-7. http://ceelo.org/wp-content/uploads/2015/11/ceelo_policy_brief_assessment_final_web_2015_11_11.pdf.
- 33 Ibid. 5-6.
- 34 Snow. "Developing Kindergarten Readiness."
- 35 Desired Results Training and Technical Assistance Project Team. (2012). "Desired Results Developmental Profile-School Readiness Validation Studies Summary." http://drdpk.org/docs/DRDP-SR_ValidationStudiesSummary.pdf.
- 36 University of California—Berkeley Evaluation and Assessment Research Center and WestEdCenter for Child and Family Studies. (December, 2015). "Preliminary KIDS Analysis: Three Subscales Using a Subset of Measures Across Domains." Unpublished memorandum.
- 37 Maryland State Department of Education. (2015). "Report on the Kindergarten Readiness Assessment (KRA)." Joint Chairmen's Report (HB70). <http://marylandpublicschools.org/about/Documents/Gov-Rel/R4K-LegislatureReporl.pdf>.
- 38 Golan. "Kindergarten Assessment."
- 39 Shields, K.A., Cook, K.D., and Greller, S. (2016) "How Kindergarten Entry Assessments Are Used in Public Schools and How They Correlate With Spring Assessments" (REL 2017-182). Washington, DC: US Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. 7-8.
- 40 NGA Task Force on School Readiness. (2005) "Building the Foundation for Bright Futures." 19-21. <http://www.nga.org/files/live/sites/NGA/files/pdf/0501TASKFORCEREADINESS.pdf>.

- 41 The Ounce of Prevention Fund and Illinois Association of School Boards. (2015) "Early Learning User's Guide for Illinois School Boards." 13. <http://www.iasb.com/earlylearning/>.
- 42 Golan. "Case Studies." 41-42.
- 43 National Center on Quality Teaching and Learning. (Winter 2014). "Transition to Kindergarten: Transition and Alignment Summit Guide." <https://eclkc.ohs.acf.hhs.gov/hslc/tta-system/teaching/docs/transition-summit-guide.pdf>.
- 44 Oregon Early Learning Division. (n.d.). "Kindergarten Assessment." <https://oregonearlylearning.com/kinrgarten-assessment/>; Irwin, M. (September 1, 2015). "The Importance of Kindergarten Assessment." Oregonian. http://www.oregonlive.com/opinion/index.ssf/2015/09/the_importance_of_kindergarten.html.
- 45 Golan. "Case Studies."
- 46 Administration for Children and Families, Early Childhood Learning and Knowledge Center. (January 14, 2015). "Steps to School Readiness." <https://eclkc.ohs.acf.hhs.gov/hslc/hs/sr/steps/steps-school-readiness.html>.
- 47 Schultz. "Moving Forward"; National Research Council. Early Childhood Assessment.
- 48 Green, E.L., and Bowie, L. (May 19, 2015). "New Assessments Show Half of Maryland's Students Ready for Kindergarten." Baltimore Sun. <http://www.baltimoresun.com/news/maryland/bs-md-kindergarten-readiness-20150519-story.html>.
- 49 Bruner, C., Agnamba, L.A., Calderon, M., and Anderson Simons, K. (2013). "Families Know Best." BUILD Initiative. <http://www.buildinitiative.org/Portals/0/Uploads/Documents/FamilyKnowsBest.pdf>.
- 50 Snow. "Developing Kindergarten Readiness."
- 51 Oregon Early Learning Division. "Kindergarten Assessment"; Oregon Social Learning Center. (n.d.). "Kids in Transition to School (KITS)." <http://www.oslc.org/projects/kids-transition-school-kits/>.
- 52 Golan. "Case Studies." 58-59, 69.
- 53 Ibid. 57-58, 69.
- 54 Shields. "How Kindergarten Entry Assessments Are Used." 8-9.
- 55 Golan. "Kindergarten Assessment."
- 56 Guss, S.S., Sweet-Darter, M., Mangus, B., and Stein, A. (November 2015). "Measuring Data Utilization: A Report Examining Data Utilization Applicable to Educare Schools." <http://www.ou.edu/content/dam/Education/documents/ECEI/Educare%20Data%20Utilization%20Literature%20Review%20revised%206-3.pdf>.
- 57 Educare Learning Network Research & Evaluation Committee. (April 2014). "A National Research Agenda for Early Education." Educare Learning Network. <http://www.educareschools.org/wp-content/uploads/2015/09/National-Research-Agenda.pdf>.
- 58 Some examples include three of the states highlighted in the Department of Education's 2016 case studies report. Golan. "Case Studies." 53-54.
- 59 Annie E. Casey Foundation. (2010). "Early Warning!: Why Reading by the End of Third Grade Matters." A KIDS COUNT Special Report from the Annie E. Casey Foundation. <http://www.aecf.org/resources/early-warning-why-reading-by-the-end-of-third-grade-matters/>; Hernandez, D.J. (2011). "Double Jeopardy: How Third Grade Reading Skills and Poverty Influence High School Graduation." Annie E. Casey Foundation. <http://www.aecf.org/resources/double-jeopardy/>.

- 60 Hart, B., and Risley, T.R. (1995). *Meaningful Differences in the Everyday Experiences of Young Children*. Baltimore, MD: Brookes Publishing. <http://products.brookespublishing.com/Meaningful-Differences-in-the-Everyday-Experience-of-Young-American-Children-P14.aspx>; Dealy, K., Pacchiano, D., and Shimpi, P. (2007). "The Language of Babies, Toddlers, and Preschoolers: Connecting Research to Practice." Ounce of Prevention Fund. <http://www.theounce.org/pubs/LanguageofBabies.pdf?v=1>; Halle, T., et al. (2009). "Disparities in Early Learning and Development: Lessons From the Early Childhood Longitudinal Study–Birth Cohort (ECLS-B)." *Child Trends*. <http://www.childtrends.org/?publications=disparities-in-early-learning-and-development-lessons-from-the-early-childhood-longitudinal-study-birth-cohort-ecls-b>; Loeb, S., and Bassok, D. (2007). "Early Childhood and the Achievement Gap." In Ladd, H.F., and Fiske, E.B., eds. (2007). *Handbook of Research in Education Finance and Policy*. New York: Routledge Press. <https://cepa.stanford.edu/sites/default/files/Early%20Childhood%20and%20the%20Achievement%20Gap.pdf>.
- 61 Maryland State Department of Education. "Maryland Model for School Readiness (MMSR)." (n.d.). http://marylandpublicschools.org/msde/divisions/child_care/early_learning/MMSR.htm. Maryland's efforts to utilize statewide KRA data have also been highlighted by the US Department of Education. Golan. "Case Studies." 13-14.
- 62 National Institute for Early Education Research. (2015). Maryland State Profile. http://nieer.org/wp-content/uploads/2015/05/Maryland_2014_0.pdf.
- 63 Advance Illinois. (November 2014). "The State We're In: 2014. A Report Card on Public Education in Illinois." http://www.advanceillinois.org/wp-content/uploads/2014/11/FINAL_Report_The_State_Were_In_2014.pdf.
- 64 Burnette II, D. (April 13, 2016). "Maryland Retreats From New Kindergarten Readiness Testing." *Education Week. Curriculum Matters* blog. http://blogs.edweek.org/edweek/curriculum/2016/04/maryland_retreats_from_new_kindergarten_readiness_testing.html.
- 65 Gewertz, C. (December 16, 2014). "Maryland Teachers Union Calls for Suspension of Kindergarten Test." *Education Week. Curriculum Matters* blog. http://blogs.edweek.org/edweek/curriculum/2014/12/maryland_teachers_union_calls_.html.
- 66 Marshall, B. (May 18, 2016). "Washington County Ranks Last in Kindergarten Readiness. Washington County Officials Create Task Force to Focus on Kindergarten Readiness." *WHAG News*. <http://www.your4state.com/news/maryland/washington-county-ranks-last-in-kindergarten-readiness>.
- 67 Golan. "Case Studies." 19, 29-30.
- 68 Oregon Department of Education. (January 2015). "Oregon Kindergarten Assessment: Report Overview." 5. http://www.ode.state.or.us/gradelevel/pre_k/kasdooverviewfinal.pdf.
- 69 Education Commission of the States. "Kindergarten Entrance Assessments." <http://ecs.force.com/mbdata/mbquestRT?rep=Kq1407>.
- 70 Golan. "Case Studies." 49-50.
- 71 National Association of Early Childhood Specialists in State Departments of Education (2000). "STILL Unacceptable Trends in Kindergarten Entry and Placement." Principle 5. This position was endorsed in March 2001 by NAEYC.
- 72 Golan. "Case Studies." 57.
- 73 Shields. "How Kindergarten Entry Assessments Are Used." 8.
- 74 Golan. "Case Studies." 57.
- 75 For information about the current status of state QRIS, see <http://qriscompendium.org/> (including information about growth trends: <http://qriscompendium.org/top-ten/question-1/>); Regenstein. "Framework for Rethinking." 15-19.

- 76 Connors, M.C., and Morris, P.A. (2015). "Comparing State Policy Approaches to Early Care and Education Quality: A Multidimensional Assessment of Quality Rating and Improvement Systems and Child Care Licensing Regulations." *Early Childhood Research Quarterly*, Vol. 30, 266-279.
- 77 US Department of Health and Human Services (2011). "Policies and Procedures for Designation Renewal of Head Start and Early Head Start Grantees, 45CFR Part 1307." <https://www.gpo.gov/fdsys/pkg/CFR-2014-title45-vol4/pdf/CFR-2014-title45-vol4-sec1306-36.pdf>.
- 78 Every Student Succeeds Act, Sec. 1111(c)(4)(B)(v); Regenstein, E., Connors, M., and Romero-Jurado, R. (February 2016). "Valuing the Early Years in State Accountability Systems Under the Every Student Succeeds Act," 6. Ounce of Prevention Fund. Policy Conversation No. 5, Version 1.0. <http://www.theounce.org/what-we-do/policy/policy-conversations>.
- 79 Regenstein. "Valuing the Early Years."
- 80 Regenstein. "Framework for Rethinking." 6-7, 9-10, 24-27.
- 81 Burchinal, M., Zaslow, M., and Tarullo, L., eds. (2016). "Quality Thresholds, Features, and Dosage in Early Care and Education: Secondary Data Analyses of Child Outcomes." *Monographs of the Society for Research on Child Development*, Vol. 81, No. 2.
- 82 Yoshikawa, H., et al. (October 2013). "Investing in Our Future: The Evidence Base on Preschool Education." Policy brief. Society for Research in Child Development and Foundation for Child Development. <http://fcd-us.org/resources/evidence-base-preschool>.
- 83 Regenstein. "Framework for Rethinking." 12.
- 84 Shepard. "Principles and Recommendations."
- 85 Ibid.
- 86 National Early Childhood Accountability Task Force. (2007). "Taking Stock: Assessing and Improving Early Childhood Learning and Program Quality." <http://policyforchildren.org/wp-content/uploads/2013/07/Taking-Stock.pdf>; National Research Council. *Early Childhood Assessment*.
- 87 Mashburn, A. J., and Henry, G. T. (2004). "Assessing School Readiness: Validity and Bias in Preschool and Kindergarten Teachers' Ratings." *Educational Measurement: Issues and Practice*, Vol. 23, No. 4, 16-30.
- 88 States have implemented systems to promote consistent data collection, including through having teachers complete an on-line assessment of students using video clips and having their scores compared to expert ratings. Golan. "Case Studies." 43. These systems have experienced some inconsistencies in measurement for a variety of reasons, including the timing of when assessments were conducted. Ibid. 65.
- 89 Cash, A. H., Hamre, B. K., Pianta, R. C., and Myers, S. S. (2012). "Rater Calibration When Observational Assessment Occurs at Large Scale: Degree of Calibration and Characteristics of Raters Associated With Calibration." *Early Childhood Research Quarterly*, Vol. 27, No. 3, 529-542.
- 90 Children's Learning Institute and University of Texas Health Science Center at Houston. (2016). CLI Engage. <https://cliengage.org/public/tools/>; NORC at the University of Chicago. (November 2013). "Getting on Track Early for School Success: Project Overview." <http://norc.org/PDFs/Getting%20on%20Track%20Early%20for%20School%20Success/Project%20Overview-%20Getting%20on%20Track%20for%20Early%20School%20Success%20.pdf>; Carlson, S.M., and Zelazo, P.D. (2014). "Minnesota Executive Function Scale." Reflection Sciences, LLC.
- 91 Every Student Succeeds Act, Sec. 1111(c)(4)(B)(ii)(I).

- 92 Florida Office of Early Learning. (n.d.). "Frequently Asked Questions," questions 8-12. <https://vpk.fldoe.org/infopages/faq.aspx>.
- 93 Florida Office of Early Learning. (October 2014). "Florida Kindergarten Readiness Screener (FLKRS) and Voluntary Prekindergarten (VPK) Education Program Accountability: An Overview." http://www.floridaearlylearning.com/sites/www/Uploads/files/Providers/VPK/FLKRS_VPK_Accountability_Overview_11-07-14.pdf.
- 94 Florida Office of Early Learning. (n.d.). "Frequently Asked Questions," questions 13-37
- 95 Nickinson, S. (March 8, 2016). "Kindergarten Readiness Rates Stalled Another Year." Studer Community Institute. <http://studer.org/2016/03/kindergarten-readiness-rates-stalled-another-year/>.
- 96 Nickinson, S. (March 19, 2016). "VPK Providers Face Funding, Testing Challenges." Pensacola News Journal. <http://www.pnj.com/story/news/2016/03/19/vpk-providers-face-funding-testing-challenges/81984078/>.
- 97 Mississippi Code § 37-21-51(3)(c)(iv). (2015).
- 98 Wright, C.M., and Benton, K.S. (March 2016). "Effectiveness Evaluation for Early Learning Collaboratives and Other Pre-k Classrooms." Mississippi Department of Education. 8. <http://www.mde.k12.ms.us/docs/elementary-education-and-reading-library/pre-k-classroom-effectiveness-evaluation-plan-2016-06.pdf?sfvrsn=2>.
- 99 Mississippi Department of Education. (2015). "MKAS2 Kindergarten Readiness Assessment Update." <http://www.mde.k12.ms.us/docs/student-assessment/kra-sbe-pres0-2015-05-21.pdf?sfvrsn=2>.
- 100 Wright. "Effectiveness Evaluation." 9.
- 101 Mississippi First. (January 2017). "The State of Pre-K (2014-15)." 42-45. http://1iq0332x28t34od07uajkv11.wpengine.netdna-cdn.com/wp-content/uploads/2017/01/The-State-of-Pre-K-Analysis_Final-3.pdf.
- 102 Wright, C.M. (June 2016). "Pre-Kindergarten Assessment Results for Early Learning Collaboratives and Other Four-Year-Old Classrooms." Mississippi Department of Education. http://mdereports.mdek12.org/pdf/a/2016/Pre-K-SBE-REPORT_FINAL.pdf.
- 103 Mississippi First. "State of Pre-K." 42-45; Wright. "Pre-Kindergarten Assessment Results."
- 104 Snow. "Developing Kindergarten Readiness"; Schultz. "Moving Forward." Further, a study by Yale University professor Walter Gilliam found that "preschoolers were expelled at three times the rate of students in K-12 settings. Such a process, though, was rarely called 'expulsion' by school personnel, he said. Instead, families were often just told that a preschool was not the right fit and that they should look elsewhere." Samuels, C. (March 31, 2014). "Pre-k Suspension Data Prompt Focus on Intervention." Education Week <http://www.edweek.org/ew/articles/2014/04/02/27ocrprek.h33.html>; Gilliam, W.S. (January 2008). "Implementing Policies to Reduce the Likelihood of Preschool Expulsion." FCD Policy Brief Advancing PK-3, no. 7, <http://fcd-us.org/sites/default/files/ExpulsionBriefImplementingPolicies.pdf>.
- 105 MS Code § 37-21-51(3)(d)(xiv). (2015).
- 106 Chen, J.Q., and McNamee, G.D. (June 8, 2007). Bridging: Assessment for Teaching and Learning in Early Childhood Classrooms, PreK-3. Thousand Oaks, CA: Corwin Press.
- 107 Early Childhood Data Collaborative. (February 2013). "2013 State of States' Early Childhood Data Systems." <http://www.ecedata.org/files/2013%20State%20of%20States'%20Early%20Childhood%20Data%20Systems.pdf>.
- 108 Data Quality Campaign and the Early Childhood Data Collaborative. (January 20, 2016) "Roadmap for Early Childhood and K-12 Data Linkages." The Early Childhood Data Collaborative. <http://dataqualitycampaign.org/resource/roadmap-early-childhood-k-12-data-linkages/>.

- 109 For an interesting point-counterpoint on the relative benefits of growth and proficiency, see Petrilli, M. (August 14, 2013). "The Problem With Proficiency." Thomas B. Fordham Institute. Flypaper blog. <https://edexcellence.net/commentary/education-gadfly-weekly/2013/august-15/the-problem-with-proficiency.html>, and Finn, C., Jr. (August 19, 2013). "Let's Hear It for Proficiency." Thomas B. Fordham Institute. Flypaper blog. <http://www.edexcellence.net/commentary/education-gadfly-daily/flypaper/2013/lets-hear-it-for-proficiency.html>.
- 110 Regenstein. "Changing the Metrics." 21; National Early Childhood Accountability Task Force. "Taking Stock."
- 111 Every Student Succeeds Act, Sec. 1111(c)(4)(B)(i)-(iv).
- 112 Ibid. Sec.1111(e)(1)(B)(iii)(IX).
- 113 Weissberg, D., Sexton, S., Mulhern, J., and Keeling, D. (2009). "The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness." The New Teacher Project.
- 114 Doherty, K.M., and Jacobs, S. (2015). "State of the States 2015: Evaluating Teaching, Leading and Learning." 2, 4, 29-32. National Council on Teacher Quality (NCTQ). <http://www.nctq.org/dmsView/StateofStates2015>; Toch, T. (2016). "Grading the Graders: A Report on Teacher Evaluation Reform in Public Education." Center on the Future of American Education at Georgetown University. <https://georgetown.app.box.com/s/f47qnfh63wfxhqu88pu5r0y0tkbo6bk>.
- 115 US Department of Education. "ESEA Flexibility Policy Document." (May 12, 2016). <http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html>; Sawchuck, S. (January 5, 2016). "ESSA Loosens Reins on Teacher Evaluations, Qualifications." Education Week. <http://www.edweek.org/ew/articles/2016/01/06/essa-loosens-reins-on-teacher-evaluations-qualifications.html>.
- 116 Sawchuck. "ESSA Loosens Reins."
- 117 Bornfreund, L. (May 2013). "An Ocean of Unknowns: Risks and Opportunities in Using Student Achievement Data to Evaluate PreK-3rd Grade Teachers." New America Foundation. <https://static.newamerica.org/attachments/2319-an-ocean-of-unknowns/An%20Ocean%20of%20Unknowns%20New%20America.4c931add59534b8680b2798e3d68ca99.pdf>.
- 118 NCTQ. "State of the States 2015." 1-12, Appendices B and H.
- 119 Alexander, K., Entwisle, D., and Olson, L. (2007). "Lasting Consequences of the Summer Learning Gap." American Sociological Review, Vol. 72, 167-180. http://www.ewa.org/sites/main/files/file-attachments/summer_learning_gap-2.pdf; Campaign for Grade-Level Reading. (n.d.). "Summer Learning Loss." <http://gradelevelreading.net/our-work/summer-learning-loss>.
- 120 As of the most recent update on state early childhood data systems in 2013, only 30 states had the ability to link early childhood data to their K-12 system. Early Childhood Data Collaborative. "The State of States' Early Childhood Data Systems." 11-12. It is not clear from the collaborative's report whether all 30 of those states could show linkages between individual kindergartners and their pre-k teachers.
- 121 National Institute for Early Education Research. (2016). "The State of Preschool 2015." 17. http://nieer.org/wp-content/uploads/2016/05/Yearbook_2015_rev1.pdf. The yearbook lists 33 preschool programs that require bachelor-certified teachers, but some states have multiple programs; in some states, all of the listed programs require bachelor-certified teachers (including Kansas, Louisiana, and New Jersey) and in other states, not all do (including Iowa, Pennsylvania, South Carolina, and Wisconsin).
- 122 La Paro, K. M., and Pianta, R. C. (2000). "Predicting Children's Competence in the Early School Years: A Meta-Analytic Review." Review of Educational Research, Vol. 70, No. 4, 443-484.
- 123 Campbell, D.T. (1979). "Assessing the Impact of Planned Social Change." Evaluation and Program Planning, Vol. 2, No. 1, 67-90.; Duckworth, A.L., and Yeager, D.S. (2015). "Measurement Matters Assessing Personal Qualities Other Than Cognitive Ability for Educational Purposes." Educational Researcher, Vol. 44, No. 4, 237-251.

- 124 Delaware Department of Education. (July 2016). "Delaware Performance Appraisal System-Two (DPAS-II) Guide (Revised) for Teachers." <http://www.doe.k12.de.us/domain/375>; Kentucky Department of Education. (2016). "Professional Growth and Effectiveness System (PGES)." <http://education.ky.gov/teachers/pges/pages/pges.aspx>.
- 125 Bornfreund. "Ocean of Unknowns." 9.
- 126 Shepard. "Principles and Recommendations"; Hanover Research. "Kindergarten Entry Assessments"; Helsel, F., and Krasnoff, B. (August 11, 2015). "Six Ways Principals Can Support the Transition to Kindergarten." Education Northwest. Northwest Matters blog. <http://educationnorthwest.org/northwest-matters/six-ways-principals-can-support-transition-kindergarten>.