Contents lists available at ScienceDirect



CHILDREN and YOUTH SERVICES REVIEW

Children and Youth Services Review

journal homepage: www.elsevier.com/locate/childyouth

"Quality" assurance features in state-funded early childhood education: A policy brief



Becca Merrill^{a,*}, Lora Cohen-Vogel^c, Michael Little^b, James Sadler^c, Kenya Lee^d

^a Education Northwest, United States

^b North Carolina State University, United States

^c University of North Carolina at Chapel Hill, United States

^d Notre Dame University, United States

ABSTRACT

Most states in the U.S. now have some form of publicly funded Pre-Kindergarten. This brief asks what they have done to ensure the quality of their programs. It does so by mapping state adoption of three popular program quality assurance features: Early Learning Standards, Kindergarten Entry Assessments, and Tiered Quality Rating and Improvement Systems. We find that all three are now widespread across the American states, despite little research on their effects. We suggest that these quality assurance features need to be subjected to rigorous evaluation tied to their effectiveness in improving student outcomes.

1. Introduction

At least partially in response to research supporting the positive effects of high-quality preschool on a range of student outcomes, most states in the U.S. have implemented some form of publicly funded Pre-Kindergarten (Pre-K) (Barnett et al., 2016; Curran, 2015). As of 2018, 43 states have adopted either universal or targeted programs (Friedman-Krauss et al., 2019). Universal programs are open to all fouryear olds regardless of household income. Targeted programs are means-tested, with eligibility based on income and/or other criteria.

Apart from whether they provide universal or targeted coverage, states also differ in terms of the ways they have encouraged the quality of Pre-K programming. Advocacy groups like the National Institute for Early Education Research (NIEER) have graded states on ten "quality standards" associated with their Pre-K offerings (Friedman-Krauss et al., 2019). On these criteria, states vary dramatically. Of the 43 states with state-funded Pre-K programs in 2018, 8 ranked in the bottom tier (earning 0–4 points out of a possible 10), 17 in the middle tier (5–7 points), and 18 in the top tier (8–10 points) (Friedman-Krauss et al., 2019).

The federal government has also invested in efforts to improve the quality of states' early childhood programs. One major avenue for investment has been the Race to the Top Early Learning Challenge (ELC) announced by the Obama Administration in 2011. This one billiondollar competition, sponsored by the United States Department of Education (ED), in partnership with the Department of Health and

Human Services, awarded federal funds to states based on their adherence to parameters in the competition's guidelines. ELC competitive priorities for funding privileged three assurance features as indicators of quality: Early Learning Standards (ELSs), Kindergarten Entry Assessments (KEAs), and Tiered Quality Rating and Improvement Systems (TQRISs). ELSs describe what specific content students should have learned by the end of a grade (ED, 2013a, 2014a, 2014b; U.S. Department of Education, 2014a; U.S. Department of Education, 2014b; U.S. Department of Education, 2013a; U.S. Department of Education, 2013b; U.S. Department of Education, 2013c; U.S. Department of Education, 2011; U.S. Department of Education, 2011a; U.S. Department of Education, n.d.) KEAs are state assessments that evaluate students' skills and development at kindergarten entry, and TQRISs are systems that rate early care centers on a set of defined benchmarks, which focus on the physical environment, availability of learning manipulatives, and teacher credentials, among others (Mitchell, 2005) (see Table 1).

With incentives from the federal government, state adoption of these three "quality assurance features" is on the rise. In this brief, we show the diffusion of ELSs, KEAs, and TQRISs across the American states. We also review the literature to argue that research on their ability to demarcate quality in early childhood education programs is scant, and - consequently - inconclusive at best; of the three assurance features, TQRISs have garnered the most research and show some promise as a framework for teasing out quality in early childhood education programs. We conclude by suggesting that funding be

* Corresponding author.

https://doi.org/10.1016/j.childyouth.2020.104972

Received 5 December 2019; Received in revised form 27 March 2020; Accepted 29 March 2020 Available online 03 April 2020

0190-7409/ © 2020 Elsevier Ltd. All rights reserved.

E-mail addresses: rebecca.merrill@educationnorthwest.org (B. Merrill), cohenvog@email.unc.edu (L. Cohen-Vogel), mlittle@air.org (M. Little), klee32@nd.edu (K. Lee).

B. Merrill, e	et d	ıl.
---------------	------	-----

Table 1

Definitions of Quality Assurance Features.

Early Learning Standards (ELSs)

Standards that describe what specific content students should learn by the end of a grade. To comply with the federal Early Learning Challenge (ELC) guidelines, standards must include all 5 Essential Domains of School Readiness.

Kindergarten Entry Assessments (KEAs)

Assessments that evaluate students' skills and development at kindergarten entry. To comply with ELC, assessments must include all 5 Essential Domains of School Readiness.

Tiered Quality Rating and Improvement Systems (TQRISs)

A system that rates early care centers on a set of defined benchmarks.

directed to large-scale, rigorous evaluations of ELSs, KEAs, and TQRISs; without them, it is difficult to know the effect, positive or negative, of quality assurance features on students.

2. ELC policy adoption across the states

To report on state adoption of the quality assurance features, we gathered data from several sources. Information about state ELSs comes from NIEER, as well as state education websites (New Hampshire Department of Education, 2016; North Dakota Department of Human Services, 2010a; North Dakota Department of Human Services, 2010a; North Dakota Department of Human Services, 2010b; Wyoming Department of Education, 2015; Wyoming Department of Education, 2015; Wyoming Department of Education, 2015; Services and previous work by the Education Commission of the States (2014), Schilder and Carolan (2014), and Stedron and Berger (2014). Last, data on TQRISs come from the Quality Rating and Improvement Systems Compendium, now Quality Compendium, an organization dedicated to online documentation and reporting on TQRISs.

2.1. Comprehensive early learning standards

The ELC called for states to develop "Early Learning and Development Standards" (ELSs). Currently, all 50 states have adopted them.¹ ELSs were considered to meet the ELC guidelines if they aligned with the National Education Goals Panel's (NEGP) Essential Domains of School Readiness, (Domains), established in 1995. The Domains include: (1) language and literacy development, (2) cognition and general knowledge (including early mathematics and early scientific development), (3) approaches toward learning, (4) physical well-being and motor development (including adaptive skills), and (5) social and emotional development (ED, 2011). South Dakota was the last state to meet this definition, revising its ELSs to align with the NEGP's Domains in 2017 (South Dakota Department of Education, 2017).

2.2. Kindergarten entry assessments

Second, KEAs are assessment tools used at the beginning of kindergarten to provide educators with a snapshot of children's school readiness; they are designed to collect information about what each child can do, say and write at the start of school (Ackerman & Lambert, 2020). In 2017, Weisenfeld wrote that states are using a variety of assessment tools in their Kindergarten Entry Assessment systems and that the tools are continually evolving; while the form they take varies, many KEAs require teachers to compile portfolios from students over a designated period early in the kindergarten year. The rationale behind KEAs is that school systems and the educators that teach in them will sharpen their focus on the children not meeting key school readiness standards and work to bring them up to speed (e.g., Little, Cohen-Vogel, Sadler, & Merrill, 2020). Twenty-eight states have adopted KEAs that meet federal guidelines, which, mirroring ELSs guidelines, require that assessments cover all five NEGP Domains (see Fig. 1). A few additional states claim to have a KEA but do not focus their assessment on all five domains; Iowa, Florida, and Mississippi, for example, all measure reading and math ability through their "KEAs," but none of the domains relate to socioemotional learning.

States have adopted KEAs in two ways: through legislation and through regulation (e.g., a policy installed by a state education agency). Whether states took a legislative or regulatory path, they frequently funded pilot studies and scaled implementation of KEAs over time. Of the 28 adopting states, all have state-funded Pre-K programs. Sixteen of these states require KEAs to be administered to all kindergarten students in public settings (Alaska, Colorado, Connecticut, Delaware, Georgia, Illinois, Kentucky, Louisiana, Michigan, New Mexico, North Carolina, Ohio, Oregon, South Carolina, Vermont, and Washington). In the other 12 states, the KEA is not required for all students; in Maryland, for example, individual school districts may choose to administer their KEA to all students or a representative sample of students (Bowie, 2016).

2.3. Tiered quality rating and improvement systems

Third, TQRISs are systems to assess early education programs on a set of defined benchmarks. They are, according to the National Center on Early Childhood Quality Assurance, similar to rating systems for restaurants and hotels, awarding points to programs. Benchmarks vary state to state but are often influenced by the goals and theory of change the state or region adopts and in research about factors that contribute to positive child outcomes (Schilder, Iruka, Dichter, & Mathias, 2015). They typically pertain to staff qualifications, learning environment, and family involvement (see National Center on Child Care Quality Improvement, 2015). Often states use childcare licensing standards as a base and build from those. In addition to awarding points, TQRIS systems include monitoring processes, including site visits, along with financial incentives, including increased childcare subsidy reimbursement rates, refundable tax credits, loans linked to quality ratings, and/ or priority on applications for professional development supports. In Table 2 below, we categorize and detail some common elements in state TQRIS systems (For an in-depth profile of each state's specific TQRIS composition, visit the Quality Compendium website, which features an interactive map of state systems: qriscompendium.org/view-stateprofiles.)

As of 2019, 38 states have adopted statewide TQRISs (see Fig. 2).² Of these 38 states, 33 have a state-funded Pre-K program. Of states with a TQRIS, only one, Colorado, has a blanket regulatory requirement that all Pre-K programs must participate. Seven other states (Arkansas, Maine, Massachusetts, Nevada, Rhode Island, Washington, and Wisconsin) require participation in order to receive public funding (Quality Compendium, n.d.). Only Nevada differentiates participate in that states' TQRIS to receive public funding, but home-based programs can participate in the TQRIS program on a voluntary basis (Quality Compendium, n.d.).

ELC guidelines do not state what features must be present to constitute a TQRIS. We consider states to have a TQRIS if a state legislature or state board of education has formally adopted a TQRIS, regardless of

¹ NIEER provides annual lists of states with Pre-K programs that require use of ELSs. For states without state-funded Pre-K programs, and thus unavailable in the NIEER ratings, we visited state education agency websites to determine if the state required use of ELSs that align with ELC guidelines, explained in the following sections. Link to state ELSs guidelines: <u>South Dakota North Dakota New Hampshire.</u>

² Hawaii piloted a statewide system in 2014 but reallocated the funds elsewhere after the reauthorization of the Child Care Development Block Grant in November of 2014.



Fig. 1. State Adoption of KEAs. A map of the United States indicating adoption of a KEA by state. The 22 states that have not yet adopted KEAs that address the Essential Domains of School Readiness: Alabama, Colorado, Idaho, Iowa, Florida, Kansas, Mississippi, Missouri, Montana, Nebraska, Nevada, New Hampshire, New York, North Dakota, Oklahoma, Rhode Island, South Dakota, Tennessee, Texas, Utah, Virginia, West Virginia, Wisconsin and Wyoming.

Table 2				
Overview	of co	mmon	TORIS	elements.

TQRIS Category	Rating Elements	State Example
Learning Environment and Curriculum	 Staff:child ratio Standards and curriculum used Early Childhood Environment Rating Scale (interaction, learning, and language subscales) CLASS Observation 	The Louisiana Department of Education requires all publicly funded early childhood programs to receive two CLASS observations each year, which is the sole measure in the state TRQIS
Qualified Professionals	 Certification levels of teachers Degree levels of teachers Professional development credits 	Colorado Shines, Colorado's TQRIS system awards points to programs based on the number of hours in the past year teachers have spent with coaches or in certified professional development
Program Safety and Administrative Practices	 Early Childhood Environment Rating Scale (space, care, and program structure subscales) Family and community engagement practices 	In North Carolina, all programs participating in the TQRIS receive the Environmental Rating Scale, and points are awarded based on cut points of the scale score result



Fig. 2. State Adoption of TQRIS. The 12 states that have not yet adopted TQRISs: Alabama, Alaska, Connecticut, Florida, Hawaii, Kansas, Louisiana, Missouri, South Dakota, Texas, West Virginia, and Wyoming.

the level of participation among early childhood education program providers. For example, Florida is not included in the definition because the state has not voted to adopt a TQRIS, even though four individual counties in Florida have developed their own systems (Quality Compendium, n.d.).

3. "Quality" assurance features in Pre-K Programs: What do we Know?

Having shown that the adoption of ELSs, KEAs, and TQRIS is widespread across the United States, we turn next to reviewing what the literature shows regarding the impact of each on student outcomes and find scarce evidence.

3.1. Comprehensive early learning standards

The theory behind ELSs (and content standards generally) is that student learning will be improved when teachers align their instruction with detailed, developmentally appropriate content standards (Clune, 1993; Cohen-Vogel, Sadler, Little, & Merrill, 2020; Polikoff, 2012; Smith & O'Day, 1990; Wachen, Harrison, & Cohen-Vogel, 2017). However, there is no direct evidence on the effectiveness of ELSs in terms of student outcomes. While educational standards are relatively new in the world of early childhood education (Scott-Little, Kagan, & Frelow, 2006), standards and standards-based reform have been a prominent strategy in K-12 education for over 25 years (Cohen-Vogel & Rutledge, 2009; Smith & O'Day, 1990). In the K-12 research, there is evidence that content standards do influence teachers' instruction, but there is no causal evidence regarding their effects on student outcomes (Polikoff, 2012). The argument for standards in education, then, is largely theoretical and more empirical research is needed.

Despite the lack of evidence on the effects of educational standards in either K-12 or Pre-K, there has been work to define what should be included in comprehensive ELSs. A consensus, signaled by the ED's adoption of the Domains, has emerged that children's school readiness should be defined in a way that is consistent with the work of the NEGP's Domains (U.S. DOE, 2011a). Supplemental work has emerged that details ways that early childhood educators can effectively align their instruction to ELSs (Flores, Curby, Coleman, & Melo, 2016). Despite generally positive views of ELSs among Pre-K teachers (DeBruin-Parecki & Slutzky, 2016), others have also raised concerns about the proliferation of ELSs. A recent paper by Mueller and File (2019), for example, argues that ELSs have significantly influenced the development of commercial curriculum packages in early education, which they argue may distract from a whole-child perspective in Pre-K education; another by Graue, Ryan, Nocera, Northey, and Wilinski (2017) argue that the push for early learning standards pulls Pre-K education into the "K-12 orbit" with its associated norms of accountability, assessments, and benchmarks, which the authors see as inappropriate in early childhood.

3.2. Kindergarten entry assessments

The theory of action behind KEAs is grounded in the fact that children enter kindergarten with drastic variation in skills (Bassok & Latham, 2017). KEAs enable teachers to capture an early picture of this variation—across a range of developmental domains—so they may individualize and optimize their instructional approach. Arguably, such an assessment is necessary in kindergarten, in particular, where students are often new to the school and teachers have little to no information about the skills of incoming students. This contrasts with other grades where a portfolio of data can be developed on individual students and easily shared from one grade to the next. However, despite widespread adoption of KEAs across the states, there is not much empirical research on them, and no research has looked at whether the implementation of KEAs has led to improved student outcomes in kindergarten and beyond.

One segment of this small literature focuses on the psychometric aspects (e.g., validity and reliability) involved in assessing young children (Ackerman, 2018; Dahlke et al., 2017; Epstein, Schweinhart, DeBruin-Parecki, & Robin, 2004; Scott-Little & Niemeyer, 2001). The findings from this line of research vary based on the form and scope of the specific KEA studied; in general, KEAs are shown to have relatively strong construct and concurrent validity, but relatively weak reliability. Low reliability is not surprising given these assessments are often teacher-administered, observational assessments (Dahlke et al., 2017). We could find no studies that looked at whether and to what extent the implementation of KEAs influenced student outcomes, which is surprising given their prevalence. Despite a lack of research on KEAs regarding student outcomes, a set of studies by one group of researchers suggest that data from early assessments may be driving some highstakes decisions about programs and children, such as determining eligibility for kindergarten and determining children's classroom placements (Curran, Little, Cohen-Vogel, & Domina, 2018; Little & Cohen-Vogel, 2017; Little, Cohen-Vogel, & Curran, 2016).

A final segment of the scant literature on KEAs focuses on their implementation in specific state contexts, including North Carolina (Ferrara & Lambert, 2016; Little et al., 2020); Alaska (Harvey & Ohle, 2018); Ohio (Schachter, Strang, & Piasta, 2017); and Maryland, Oregon, Pennsylvania, and Washington (Golan, Woodbridge, & Davies-Mercier, 2016). These studies—while not including program outcomes—reveal common challenges with the implementation of KEAs, including lack of teacher buy-in, lack of teacher time, and challenges with the technological platforms on which the assessments operate. Implementation facilitators include the whole-child focus of some KEAs, and staged approaches wherein teacher feedback is incorporated throughout the implementation process.

3.3. Tiered quality rating and improvement systems

Research on TQRISs has burgeoned in recent years, though much remains to be understood. TQRISs are designed to detect varying levels of quality in early education programs. They are not designed to be interventions themselves. Thus, the research on TQRISs has not evaluated the overall effectiveness of these systems on students' early learning outcomes. That is, no research has examined if a state's adoption of a TQRIS *system* led to changes in the outcomes of students served by programs within that system. Rather, the research has largely focused on the extent to which components within TQRISs promote program quality either by identifying and tracking program quality characteristics that are predictive of student outcomes or being able to meaningfully detect differences in program quality.

One arm of research examines if the components (i.e., teacher-child ratios, teacher credentials) that TQRIS systems seek to incentivize are actually predictive of improved program quality for student outcomes, not if the entire TQRIS system is predictive of student outcomes (Yazejian & Iruka, 2015). The results from this line of research are decidedly mixed. Structural measures of quality-including teacher preparation and child-teacher ratios, for example-are not generally predictive of student outcomes. However, process measures of quality-including observations of instruction and/or teacher-child interactions (i.e., CLASS scores)-are generally more predictive of student outcomes (Elicker & McConnell, 2011; Karoly, Zellman, & Perlman, 2013; Zellman, Perlman, Le, & Setodji, 2008). That said, a recent research study examined the inter-rater reliability of CLASS ratings in Louisiana's TQRIS and found that scores rated by local raters and researchers were only modestly correlated. The researchers found that local raters provided systematically higher scores-raising questions about the practical reliability of these measures when used at scale (Vitiello, Bassok, Hamre, Player, & Williford, 2018). For CLASS ratings to be useful, valid measures used as part of a TQRIS system, states will need to attend to effectively training raters.

Most research on TQRISs has focused on how these systems work to inform program quality improvement through financial incentives (i.e., higher reimbursement rates for higher rated programs), support for increasing teacher credentials, and on-site technical assistance (TA) (Yazejian & Iruka, 2015). Research on financial incentives is inchoate, but the existing evidence suggests that financial incentives, such as levels of program reimbursement, can facilitate program improvement, where program improvement is defined as a program moving up tiers within the TQRIS (e.g., moving from a 3-star to a 4-star program) (Gormley & Lucas, 2000; Mitchell, 2012; Schulman, Matthews, Blank, & Ewen, 2012). Research on the effects of scholarships and other supports for increasing teacher credentials is limited because of the long-range nature of education and degree attainment. However, two studies find a positive association between the proportion of staff in a program participating in the T.E.A.C.H. model, a teacher education program, and quality ratings (Cassidy, Buell, Pugh-Hoese, & Russell, 1995; Childcare Services Association (2003) (2003), 2003). Research generally suggests that on-site TA is associated with improvements in program quality, as measured by classroom observations (e.g., CLASS and ECERS) (Isner et al., 2011; Karoly, Schwartz, Setodji, & Haas, 2016; Snell, Forston, Stanton-Chapman, & Walker, 2013). Elements of successful professional development include sustained intervention, collaboration, grounded in practice, and links to information about standards and children's growth (Diamond & Powell, 2011; Weiland & Yoshikawa, 2013).

A core assumption underlying the model of change for TQRISs is that the public nature of rating scales will inform consumers and generate incentives for programs to improve. A paper from Bassok, Markowitz, Player, and Zagardo (2018) directly tested the extent to which parental evaluations of Pre-K programs aligned with TQRIS rating standards, finding that parents' evaluations of the program their child attends were not systematically related with the TQRIS rating standards, including teacher education, class size, hours, cost, or measures of average classroom learning gains.

In summary, the evidence on three quality assurance features incentivized by the ELC competitive priorities and popular among states is scant. The evidence on ELSs, similar to standards in K-12 education, is largely theoretical and lacks empirical support. With respect to KEAs, there is essentially no evidence on their effects on student outcomes during the kindergarten year and beyond, though researchers have worked to define the appropriate content and use of such assessments. Finally, while there is evidence that some components within TQRISs may support program improvement (e.g., financial incentives), there is also little evidence about the efficacy of other components (e.g., structural factors), and no evidence on these systems' overall effectiveness in terms of student outcomes. Given the scarcity of empirical research to support TQRIS and especially ELSs and KEAs, we question the wisdom of their widespread adoption and the investment of public monies without large scale evaluation.

4. Discussion

As we have shown, the adoption of ELSs, KEAs, and TQRISs is now widespread across the American states. All 50 states have ELSs; 28 have KEAs; and 38 states have adopted TQRISs. Despite widespread adoption, there is relatively little research into their effects on students. While hardly the first understudied policies to be adopted by state legislatures, the popularity and propagation of ELSs, KEAs, and TQRISs should nevertheless incent policymakers at both the state and federal levels to invest in large scale evaluations or risk what has been described as the faddish churn of ideas that filter in and out of schools, disrupting teaching and learning and potentially wasting state and federal dollars (Cuban, 1990; Tyack & Cuban, 1995; Cohen-Vogel & Mehta, 2017).

Specifically, large-scale, rigorous evaluations can simultaneously vet the quality assurance features to assess their ability to improve student outcomes and, at the same time, ensure public dollars are allocated towards programs that are supported by research. On this front, commonalities in ELSs, KEAs, and TQRISs resulting from states' adherence to ELC guidelines present opportunities to design large, cross-state evaluations with potentially high saliency for policymakers across the nation. Evaluations should be designed to answer: What is the effect of implementing each quality assurance feature on academic and non-academic outcomes for students in the short and long term? They should be careful to use rigorous methods to consider children's experiences before, during, and after the pre-k year, lest we miss information that can help build understanding about how to design and scale pre-k programming to get the best results (Phillips, Lipsey, Dodge, Haskins, Bassok, Burchinal, & Weiland, 2017). Evaluations should also take care to include implementation metrics that measure both spread and depth (e.g., number of KEA assessors trained; quality of the training), in order to assure that any detected effect (or lack thereof) on student outcomes is in fact due to the program. Finally, in determining the student outcomes of interest, researchers should work in partnership with policymakers and practice experts. Doing so will help ensure that the results of the evaluations are actionable. Partnering in this way helps push researchers to focus in on the needs of practitioners and generate findings that are most likely to be used and incorporated into changes to both policy and practice (Coburn & Penuel, 2016).

While waiting for more definitive guidance from large scale evaluations on how to design and re-engineer early learning programs, our review of the limited literature suggests that policymakers might work to embed and improve observational measures of classroom instruction and on-site technical assistance for Pre-K teachers into TQRISs (Elicker & McConnell, 2011; Isner et al., 2011; Karoly et al., 2016; Karoly et al., 2013; Snell et al., 2013; Zellman et al., 2008). The limited literature on TQRISs implies that technical assistance and other forms of teacher professional development should focus on sustained intervention and collaboration, be grounded in practice, and linked to information about standards and children's growth (Diamond & Powell, 2011; Weiland & Yoshikawa, 2013).

CRediT authorship contribution statement

Becca Merrill: Investigation, Resources, Writing - original draft, Writing - review & editing, Supervision, Project administration. **Lora Cohen-Vogel:** Conceptualization, Resources, Data curation, Writing original draft, Writing - review & editing, Supervision, Funding acquisition. **Michael Little:** Writing - original draft, Writing - review & editing. **James Sadler:** Visualization, Data curation, Writing - original draft, Writing - review & editing. **Kenya Lee:** Writing - original draft, Writing - review & editing, Investigation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The research reported here was supported in part by the Institute of Education Sciences, U.S. Department of Education, through Grant R305N160022 to the University of North Carolina at Chapel Hill. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.childyouth.2020.104972.

References

- Ackerman, D. J. (2018). Real world compromises: Policy and practice impacts of kindergarten entry assessment-related validity and reliability challenges. ETS Research Report Series, 2018(1), 1–35.
- Ackerman, D., & Lambert, R. (2020). Kindergarten Entry Assessments: Policies. Practices: Potential Pitfalls, and Psychometrics. Early Education and Development30.
- Barnett, W. S., Friedman-Krauss, A. H., Gomez, R. E., Horowitz, M., Weisenfeld, G. G., & Squires, J. H. (2016). The state of preschool 2015: State preschool yearbook. New Brunswick, NJ: National Institute for Early Education Research.
- Bassok, D., Markowitz, A. J., Player, D., & Zagardo, M. (2018). Are parents' ratings and satisfaction with preschools related to program features? *AERA Open*, 4(1).
- Bassok, D., & Latham, S. (2017). Kids today: The rise in children's academic skills at kindergarten entry. *Educational Researcher*, 46(1), 7–20.
- Bowie, L. (2016, December 5). Less than half of Maryland kindergartners ready for school. The Baltimore Sun. http://www.baltimoresun.com/bs-md-kindergarten-readiness-20161205-story.html.
- Cassidy, D. I., Buell, M. I., Pugh-Hoese, S., & Russell, S. (1995). The effect of education on childcare teachers' beliefs and classroom quality: Year one evaluation of the TEACH early childhood associate degree scholarship program. *Early Childhood Research Quarterly*, 10(2), 171–183.
- Childcare Services Association (2003). Working in Child Care in North Carolina: The North Carolina Child Care Survey of 2003. Retrieved from http://www.childcareservices.org/_downloads/NC2003wfreport.pdf.
- Clune, W. H. (1993). The best path to systemic educational policy: Standard/centralized or differentiated/decentralized? *Educational Evaluation and Policy Analysis*, 15(3), 233–254.
- Coburn, C. E., & Penuel, W. R. (2016). Research–practice partnerships in education: Outcomes, dynamics, and open questions. *Educational Researcher*, 45(1), 48–54. https://doi.org/10.3102/0013189X16631750.
- Cohen-Vogel, L., & Mehta, J. (2017). Why reform sometimes succeeds: Understanding the conditions that produce reforms that last. *American Educational Research Journal*, 54(4), 644–690.
- Cohen-Vogel, L., & Rutledge, S. (2009). The pushes and pulls of new localism: Schoollevel instructional arrangements, instructional resources, and family-community partnerships. The new localism in American education: The yearbook of the National Society for the Study of Education, 108, 70–103.
- Cohen-Vogel, L., Sadler, J., Little, M., & Merrill, B. (2020). (Mis)Alignment of instructional policy supports in Pre-K and Kindergarten: Evidence from rural districts in North Carolina. *Early Childhood Research Quarterly*. https://doi.org/10.1016/j.ecresq. 2019.11.001.
- Cuban, L. (1990). Reforming again, again, and again. Educational Researcher, 19(1), 3-13.
- Curran, F. C. (2015). Expanding downward: Innovation, diffusion, and state policy adoptions of universal preschool. *Education Policy Analysis Archives*, 23, 36. https:// doi.org/10.14507/epaa.v23.1688.
- Curran, F. C., Little, M. H., Cohen-Vogel, L., & Domina, T. (2018). School readiness assessments for class placements and academic sorting in kindergarten. *Educational Policy*. https://doi.org/10.1177/0895904818802109.
- Dahlke, K., Yang, R., Martínez, C., Chavez, S., Martin, A., Hawkinson, L., ... Carle, J. (2017). Scientific evidence for the validity of the New Mexico Kindergarten Observation Tool (REL 2018–281). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest.
- DeBruin-Parecki, A., & Slutzky, C. (2016). Exploring pre-k age 4 learning standards and their role in early childhood education: Research and policy implications. *ETS Research Report Series*, 2016(1), 1–52.
- Diamond, K. E., & Powell, D. R. (2011). An iterative approach to the development of a professional development intervention for Head Start teachers. *Journal of Early Intervention*, 33(1), 75–93.
- Education Commission of the States (2014). 50 State Comparison: Kindergarten Entrance Assessments. http://ecs.force.com/mbdata/mbquestRT?rep = Kq1407.
- Elicker, J. D., & McConnell, N. L. (2011). Interactive learning in the classroom: Is student response method related to performance? *Teaching of Psychology*, 38(3), 147–150.
- Epstein, A. S., Schweinhart, L. J., DeBruin-Parecki, A., & Robin, K. B. (2004). Preschool assessment: A guide to developing a balanced approach. *Preschool Policy Matters*, 7, 1–2
- Ferrara, A. M., & Lambert, R. G. (2016). Findings from the 2015 statewide implementation of the North Carolina K–3 formative assessment process: Kindergarten entry assessment. Charlotte: University of North Carolina at Charlotte, Center for Educational Measurement and Evaluation. Retrieved from https://ceme.uncc.edu/ sites/ceme.uncc.edu/files/media/Findings%20from%20the%202015%20North %20Carolina%20K-3%20FAP%20KEA% 20Implementation.pdf.
- Flores, R. L., Curby, T. W., Coleman, H., & Melo, K. (2016). Using early learning standards to provide high-quality education for all children: The early learning guidelines toolkit. *Theory into Practice*, 55(2), 145–152.
- Friedman-Krauss, A. H., Barnett, W. S., Weisenfeld, G. G., Kasmin, R., DiCrecchio, N., & Horowitz, M. (2019). *The State of Preschool 2018: State Preschool Yearbook*. New Brunswick, NJ: National Institute for Early Education Research.
- Golan, S., Woodbridge, M., & Davies-Mercier, B. (2016). Case studies of the early implementation of kindergarten entry assessments. Washington, DC: U.S. Department of Education.
- Gormley, W. T., & Lucas, J. K. (2000). Money, accreditation and childcare center quality. Foundation for Child Development.
- Graue, M. E., Ryan, S., Nocera, A., Northey, K., & Wilinski, B. (2017). Pulling preK into a K-12 orbit: The evolution of preK in the age of standards. *Early Years*, 37(1), 108–122.

- Harvey, H., & Ohle, K. (2018). What's the purpose? Educators' perceptions and use of a state-mandated kindergarten entry assessment. *Education Policy Analysis Archives*, 26(142), https://doi.org/10.14507/epaa.26.3877.
- Isner, T., Tout, K., Zaslow, M., Soli, M., Quinn, K., Rothenberg, L., & Burkhauser, M. (2011). Coaching in Early Care and Education Programs and Quality Rating and Improvement Systems (QRIS): Identifying Promising Features. Bethesda, MD: Child Trends.
- Karoly, L. A., Schwartz, H. L., Setodji, C. M., & Haas, A. C. (2016). Evaluation of Delaware Stars for Early Success: Final Report. Santa Monica, CA: RAND Corporation.
- Karoly, L. A., Zellman, G. L., & Perlman, M. (2013). Understanding variation in classroom quality within early childhood centers: Evidence from Colorado's quality rating and improvement system. *Early Childhood Research Quarterly*, 28(4), 645–657.
- Little, M. H., & Cohen-Vogel, L. (2017). Ready for school? Assessing America's kindergarteners. Teachers College Record. https://www.tcrecord.org ID Number: 21959.
- Little, M. H., Cohen-Vogel, L., & Curran, F. C. (2016). Facilitating the transition to kindergarten: What ECLS-K data tell us about school practices then and now. AERA Open, 2(3), https://doi.org/10.1177/2332858416655766.
- Little, M., Cohen-Vogel, L., Sadler, J., & Merrill, B. (2020). Moving Kindergarten Entry Assessments from Policy to Practice Evidence from North Carolina. *Early Education* and Development, 1–20.
- Mitchell, A.W. (2012). Considerations for an efficient, inclusive and implementable Quality Rating and Improvement System. The BUILD Initiative, QRIS National Learning Network. http://www.qrisnetwork.org.
- Mitchell, A. W. (2005). Stair steps to quality: A guide for states and communities developing quality rating systems for early care education. Alliance for Early Childhood Finance. http://www.earlychildhoodfinance.org/downloads/2005/MitchStairSteps_ 2005.pdf.
- Mueller, J. J., & File, N. (2019). Standards, Correlations, and Questions: Examining the Impact of the Accountability Regime on Early Childhood Curriculum. In Curriculum in Early Childhood Education (pp. 34-50). Routledge.
- National Center on Child Care Quality Improvement (2015). QRIS Standards, Levels, Rating Systems (No. 730). United States Office of Health and Human Services, Office of Child Care. Retrieved on April 18, 2018 from https://qrisguide.acf.hhs.gov/files/ QRIS_Levels_Rating.pdf.
- National Institute for Early Education Research (2020) State of Preschool Yearbooks. http://nieer.org/state-preschool-yearbooks.
- New Hampshire Department of Education (2016). New Hampshire Early Learning Standards. Retrieved from https://www.education.nh.gov/instruction/curriculum/ early learning.htm.
- North Dakota Department of Human Services (2010). North Dakota Early Learning Guidelines North Dakota Department of Human Services: Birth to Three. Retrieved from https://www.nd.gov/dhs/info/pubs/docs/cfs/2009-08-earlylearning-birth-3. pdf.
- North Dakota Department of Human Services (2010). North Dakota Early Learning Guidelines North Dakota Department of Human Services: Birth to Five. Retrieved from http://www.nd.gov/dhs/info/pubs/docs/cfs/nd-early-learning-guidelines-forages-3-thru-5.pdf.
- Phillips, D., Lipsey, M. W., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M. R., & Weiland, C. (2017). Puzzling it out: The current state of scientific knowledge on prekindergarten effects-A consensus statement. https://www.brookings.edu/wpcontent/uploads/2017/04/consensus-statement final.pdf.
- Polikoff, M. S. (2012). Instructional alignment under No Child Left Behind. American Journal of Education. 118(3), 341–368.
- Quality Compendium (n.d.). View state profiles: View full state profile. Retrieved from https://qualitycompendium.org/create-a-report.
- Quality Compendium (n.d.). About: History. https://qualitycompendium.org/about.
- Schachter, R., Strang, T., & Piasta, S. (Summer 2017). Teachers' perspectives on the Kindergarten Readiness Assessment in year 2: Easier to administer but what role can it play in instruction? Columbus, OH: Crane Center for Early Childhood Research and Policy, The Ohio State University.
- Schilder, D., & Carolan, M. (2014). State of the States Policy Snapshot: State early childhood assessment policies. Center on Enhancing Early Learning Outcomes. Retrieved from http://ceelo.org/wpcontent/uploads/2014/03/CEELO_policy_ snapshot_child_assessment_march_2014.pdf.
- Schilder, D., Iruka, I., Dichter, H., & Mathias, D. (2015). Quality Rating and Improvement Systems: Stakeholder Theories of Change and Models of Practice. version https:// qrisnetwork.org/sites/all/files/resources/2016-02-10%2009%3A21/QRIS%203.0% 20Report%20V11%202016.2.5%20FINAL.pdf. (Accessed 3 April 2020).
- Scott-Little, C., Kagan, S. L., & Frelow, V. S. (2006). Conceptualization of readiness and the content of early learning standards: The intersection of policy and research? *Early Childhood Research Quarterly*, 21(2), 153–173.
- Scott-Little, C. & Niemeyer, J. (2001). Assessing Kindergarten Children: A Compendium of Assessment Instruments. Southeastern Regional Vision for Education. Office of Educational Research and Improvement, United States Department of Education.
- Schulman, K., Matthews, H., Blank, H., & Ewen, D. (2012). A Count for Quality: Child Care Center Directors on Rating and Improvement Systems. Center for Law and Social Policy, Inc. (CLASP).
- Smith, M. S., & O'Day, J. (1990). Systemic school reform. Journal of Education Policy, 5(5), 233–267.
- Snell, M. E., Forston, L. D., Stanton-Chapman, T. L., & Walker, V. L. (2013). A review of 20 years of research on professional development interventions for preschool teachers and staff. *Early Child Development and Care*, 183(7), 857–873.
- South Dakota Department of Education (2017). South Dakota Early Learning Guidelines. Retrieved from https://dss.sd.gov/docs/childcare/early_learning_guidelines.pdf.
- Stedron, J., & Berger, A. (2014). NCSL technical report: State approaches to school readiness assessment. 2014 Update. In National Conference of State Legislators.

Retrieved from http://www.ncsl.org/Portals/1/Documents/educ/NCSL_Readiness_ Assessment_2014_Update_Report_Chart.pdf.

- Tyack, D. B., & Cuban, L. (1995). *Tinkering toward utopia*. Harvard University Press. U.S. Department of Education (2014a). Programs: Race to the Top Early Learning Challenge Phase 1. Retrieved from http://www2.ed.gov/programs/racetothetopearlylearningchallenge/awards-phase-1.html.
- U.S. Department of Education (2014b). Programs: Race to the Top Early Learning Challenge Phase 2. Retrieved from http://www2.ed.gov/programs/racetothetopearlylearningchallenge/awards-phase-2.html.
- U.S. Department of Education (2013a). Programs: Race to the Top Early Learning Challenge: Applicant Information. Retrieved from http://www2.ed.gov/programs/ racetothetop-earlylearningchallenge/applicant.html.
- U.S. Department of Education (2013b). Programs: Race to the Top Early Learning Challenge: Applicant Information Phase 1. http://www2.ed.gov/programs/ racetothetop-earlylearningchallenge/applicant-phase-1.html.
- U.S. Department of Education (2013c). Programs: Race to the Top Early Learning Challenge: Applicant Information Phase 2. http://www2.ed.gov/programs/ racetothetop-earlylearningchallenge/applicant-phase-2.html.
- U.S. Department of Education (2011). Race to the Top-Early Learning Challenge Application. https://www2.ed.gov/programs/racetothetopearlylearningchallenge/ applicant-phase-1.html.
- U.S. Department of Education (2011a). Race to the Top- Early Learning Challenge NIA Draft Executive Summary. https://www.ed.gov/sites/default/files/rtt-elc-draftexecsumm-070111.pdf.

- U.S. Department of Education (nd). Race to the Top-Early Learning Challenge Program: Definitions. https://www.ed.gov/early-learning/elc-draft-summary/definitions.
- Vitiello, V. E., Bassok, D., Hamre, B. K., Player, D., & Williford, A. P. (2018). Measuring the quality of teacher-child interactions at scale: Comparing research-based and state observation approaches. *Early Childhood Research Quarterly*, 44, 161–169.
- Wachen, J., Harrison, C., & Cohen-Vogel, L. (2017). Data use as instructional reform: Exploring educators' reports of classroom practice. *Leadership and Policy in Schools*, 1–30.
- Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development*, 84(6), 2112–2130.
- Weisenfeld, G. G. (2017). Assessment tools used in kindergarten entry assessments (KEAs): State scan. New Brunswick, NJ: Center on Enhancing Early Learning Outcomes.
- Wyoming Department of Education (2015). Wyoming Early Learning Guidelines. http:// wyqualitycounts.org/wp-content/uploads/2015/01/Early-Learning-Guidelines.pdf.
- Wyoming Department of Education (2014). Wyoming Early Learning Foundations. Retrieved from https://edu.wyoming.gov/downloads/early-childhood/2014/14align-0009-early-learning-foundations-spreads-1.pdf.
- Yazejian, N., & Iruka, I. U. (2015). Associations among tiered quality rating and improvement system supports and quality improvement. *Early Childhood Research Quarterly*, 30, 255–265.
- Zellman, G. L., Perlman, M., Le, V. N., & Setodji, C. M. (2008). Assessing the validity of the Qualistar early learning quality rating and improvement system as a tool for improving child-care quality. Rand Corporation.