Classroom Quality in the Age of Accountability: Using a Comprehensive Multidimensional Rasch Approach to Investigate the Validity of the Early Childhood Environment Rating Scale-Revised

A Dissertation submitted by

Brandon Foster

in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Child Study and Human Development

Tufts University
Eliot-Pearson Department of Child Study and Human Development

May 2017

Committee: Drs. Christine McWayne (Chair),
Tama Leventhal, Shelagh Peoples, and Eric Dearing

PREVIEW

ProQuest Number: 10282401

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 10282401

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

PREVIEW

Abstract

The Early Childhood Environment Rating Scale–Revised (ECERS-R) is the most widely used measure of classroom quality, and has been implemented into numerous states' Quality Rating and Improvement Systems (QRISs). However, the ECERS-R was never designed for implementation into these systems or to be used for the purposes of accountability. As such, its validity for determining benchmarks on key indicators of classroom quality and suitability for identifying low or high performing preschools has yet to be established. Findings from a handful of validity studies which utilized classical test theory (CTT) have highlighted information pertaining to the factor structure of the measure. However, the CTT approach provides very limited diagnostic information about the functioning of the ECERS-R (e.g., Fan, 1998; Kieffer, 1998; Macdonald & Paunonen, 2002; Traub, 1997; Schumacker & Smith, 2007), and studies investigating the psychometric properties of the ECERS-R using more rigorous methods from item-response theory have hardly emerged (Gordon et al., 2013; 2015). This dissertation provides the first comprehensive account of the psychometric properties of the ECERS-R, utilizing the ECLS-B dataset, a nationally representative sample of early childhood environments. In particular, three proposed dimensional specifications of the measure that have emerged from the literature were examined. Psychometric analyses were carried out using Andrich Rating Scale Models and Multidimensional Random Coefficient Multinomial Logit (MRCML) models. To organize this information, this dissertation adopted Wolfe and Smith's (2007) Rasch validity framework, which ensured that analyses provided psychometric information about each component of Messick's seminal (1989) construct validity framework. Results illuminate the psychometric trade-offs that occur when different specifications of the ECERS-R are utilized to measure classroom quality. Additional information, which builds off the work of Gordon et al. (2013; 2015b), is provided to guide policymakers in the use of total scores from the measure in policy applications. Further, a promising new dimensional structure of the ECERS-R is posited and confirmed. Results from these analyses suggested that the ECERS-R functioned as a risk assessment, capable of precisely measuring classrooms with low levels of classroom quality, but functioned poorly for classrooms with average or above average levels of classroom quality. These results raise serious concerns about the use of the ECERS-R in most states' early-childhood assessment systems, and call into question some basic assumptions for how this measure is assumed to have performed in the prior research studies.

PREVIEW

Acknowledgements

The culmination of this milestone would not have been possible without the support of many individuals.

I would like to first thank my dissertation committee. Drs. Christine McWayne, Tama Leventhal, and Shelagh Peoples have each provided helpful feedback and have been great mentors.

To my advisor, Dr. Christine McWayne, for her unwavering support and encouragement over the years. She has helped me realize professional potentials that I could have never imagined, and always kept me focused on the big picture. Her passion for this work continues to inspire me.

To my EP Ph.D. cohort (Cricket, Lisette, Maggie and Judith), for their friendship and laughter. I was not expecting to make such amazing friends while here, and I am so grateful for the support they have provided me.

To labmates (Sunah & Lok), for their kindness and generosity.

To my sister, Kayla Foster, for her love and support, which always came just in the nick of time. She has never let me forget to smile throughout this process, as well as in life.

To my parents, Theresa and Ron Foster, who answered every call and assuaged every doubt. It was their hard work and continued sacrifices when I was growing up that enabled me to reach beyond what I thought was possible. This accomplishment is as much theirs as it is mine. I know they are proud, and that fills my heart with joy.

And finally, to my wonderful wife, Julie Thomason, who has been there with me since the beginning of this journey. My gratitude for her sacrifices is unmeasurable. Her honesty and tough love helped me to dig deeper, work harder, and persist. I love you and could not imagine a better partner in life.

Thank you.

PREVIEW

Table of Contents

List of Tables 6

List of Figures 8

Introduction 9

Guiding Theoretical Frameworks for Understanding Preschool Classroom Quality 13

Literature Review 18
  Current Policy Context for Expanding Access to High Quality Early Childhood Education in the U.S 18
  Summary of the Context for Investigating the Validity of Measures of Preschool Quality 23
  Definitions of Early Childhood Classroom Quality 24
  Measuring Classroom Quality 28
  Relationships Between Measures of ECE Quality and Child Outcomes 30
  History of the ECERS 34
  Investigations Into the Validity of the ECERS-R 36
  Critical Gaps 48
    Messicks' Construct Validity Framework 51
  Psychometric Framework of This Study 53
  Research Questions 56
    Content Validity 56
    Substantive Validity 57
    Generalizability Validity 57
    Structural Validity 57
    External Validity 58
    Predictive Validity 58

Method 60
  Data 60
    Sample Characteristics and Sample Size 60
  Measures 62
    Child Assessment Data 62
    Measure of Classroom Quality 62
  Analytic Methods 64
    Analytic Summary of Rasch Methodology 64
    Wolfe & Smith's (2007) Validity Tests 68

Results 77
  Content and Substantive Validity Evidence 77
    Item Technical Quality 77
    Rating Scale Functioning 81
    Revised Item Technical Quality 84
    Summary of Revised Item Technical Quality 88
    Expected Item Difficulty Hierarchy 88
  Generalizability Validity Evidence 91
    Item Difficulty Invariance 92
    Reliability of Person Estimates 92

PREVIEW

    - Differential Item Functioning 94
    - Precision of Person Estimates 96
  - Structural Validity Evidence 100
    - Goodness-of-fit 100
    - Comparison of Rasch Model Subscale Correlations 101
    - Discrepant Case Analyses 102
    - Summary of Structural Validity Evidence 107
  - External 107
    - The Rasch Person Strata Indices 108
    - Descriptive Statistics and Overlap for Item and Classroom Quality Estimates 109
    - Examining the Wright Maps 110
    - Correlations with the Arnett Caregiver Interaction Scale 118
    - Summary of External Validity 118
  - Predictive Validity 119
    - Associations with Children's Reading Outcomes 119
    - Associations with Children's Math Outcomes 119
    - Summary of Predictive Validity Evidence 120
  - Key Takeaways for the Psychometric Properties of the ECERS-R 120

- Discussion 122
  - New Dimensions in the Dimensional Debate for the ECERS-R 124
  - The ECERS-R as a Risk Assessment 129
  - Implications for Users 132
  - Implications for Developers 134
  - Limitations and Future Directions 138
  - Summary 142

- References 143

- Appendix A 170

- Appendix B 188

PREVIEW

List of Tables

Table 1. Comparison of items on both forms of the ECERS 35
Table 2. Percentage of Quality Response for Each Category of Rating the Scale 63
Table 3. Items from the ECERS-R Corresponding to the Provisions for Learning/Teaching and Interactions Specification of the ECERS-R 65
Table 4. Items From the ECERS-R Corresponding to the Structural and Process Dimensions ... 66
Table 5. Adaptation of Wolfe & Smith’s (2007) Conceptualization of Rasch Validity Evidence for Messick's Validity Framework 69
Table 6. Variables Used in the Multiple Regression for All Models 74
Table 7. Rasch Outfit Statistics for Each Specification of the Measure 77
Table 8. Rasch Infit Statistics for Each Specification of the Measure 78
Table 9. Average Difficulty for Each Level of the Rating Scale for Each Model 82
Table 10. Average Difficulty for Each Level of the Rating Scale for Each Model With Collapsed Categories in Rating Scale 83
Table 11. Rasch Outfit Statistics for Specifications with Collapsed Categories 84
Table 12. Rasch Infit Statistics for Specifications with Collapsed Categories 85
Table 13. Dimensional Ordering of Mean Population Parameters in Logits 89
Table 14. Comparing the Rank Order of Item Difficulty Estimates Across all Model Specifications 90
Table 15. Item Correlations Between Pairs of Item Difficulty Estimates 92
Table 16. Rasch EAP Reliability and Spearman Brown Additional Items 94
Table 17. Notable Differential Item Functioning for All Model Specifications 96
Table 18. Model Fit Statistics for All Model Specifications 100
Table 19. Correlations Between Rasch Classroom Quality Estimates 102
Table 20. Percent of Discrepant Cases for Each Model Specification 103
Table 21. Rasch Person Strata Indices 108
Table 22. Average Classroom Quality Rasch Estimates for Each Dimension 109
Table 23. Percent of Overlap Between Item Difficulty Estimates and Classroom Quality Rasch Scores 110

Table A1. Descriptive Statistics for the ECERS-R Items 170
Table A2. Differential Item Functioning for Center Type for the 37-Item Unidimensional Model 175
Table A3. Differential Item Functioning for Teacher Education for the 37-Item Unidimensional Model 176
Table A4. Differential Item Functioning for Half Time Program Status for the 37-Item Unidimensional Model 178
Table A5. Differential Item Functioning for Center Type for the 16-Item Unidimensional Model 179
Table A6. Differential Item Functioning for Teacher Education for the 16-Item Unidimensional Model 180
Table A7. Differential Item Functioning for Half Time Status for the 16-Item Unidimensional Model 181
Table A8. Differential Item Functioning for Center Type for the Provisions for Learning/Teaching Interactions 2-Dimension Model 182

PREVIEW

Table A9. Differential Item Functioning for Teacher Education for the Provisions for Learning/Teaching Interactions 2-Dimension Model 183
Table A10. Differential Item Functioning Half Time Program Status for the Provisions for Learning/Teaching Interactions 2-Dimension Model 184
Table A11. Differential Item Functioning for Center Type for the Structural/Process 2-Dimension Within-Item Model 185
Table A12. Differential Item Functioning for Teacher Education for the Structural/Process 2-Dimension Within-Item Model 186
Table A13. Differential Item Functioning for Half Time Program Status for the Structural/Process 2-Dimension Within Item Model 187
Table B1. Multivariate Regression for Children's Reading Scores for the 37-Item Unidimensional Specification of the Measure 189
Table B2. Multivariate Regression for Children's Reading Scores for the 16-Item Unidimensional Specification of the Measure 190
Table B3. Multivariate Regression for Children's Reading Scores for the Provisions for Learning/Teaching Interactions 2-Dimension Specification of the Measure 191
Table B4. Multivariate Regression for Children's Reading Scores for the 2-Dimension Structural/Process Within-Item Specification of the Measure... 192
Table B5. Multivariate Regression for Children's Math Scores for the 37-Item Unidimensional Specification of the Measure 193
Table B6. Multivariate Regression for Children's Math Scores for the 16-Item Unidimensional Specification of the Measure 195
Table B7. Multivariate Regression for Children's Math Scores for the Provisions for Learning/Teaching Interactions 2-Dimension Specification of the Measure 196
Table B8. *Multivariate Regression for Children's Math Scores for the Structural/Process 2-Dimension Specification of the Measure* 197

PREVIEW

List of Figures

Figure 1. Sample logic model for QRIS functioning, which is adapted from McCawley (2001) and the W. K. Kellogg Foundation (2004) 21
Figure 2. Graphical depictions of the different ways to model multidimensionality using the proposed Rasch analytic procedures. 67
Figure 3. Precision of the Classroom Measure Estimates.......... 98
Figure 4. Discrepant Cases for each dimension of the two multidimensional specifications against scores on the 37-item unidimensional measure. 105
Figure 5. Discrepant Cases for each dimension of the two multidimensional specifications against scores on the 16-item unidimensional measure 106
Figure 6. Person item map for the 37-item unidimensional specification of the measure.......... 114
Figure 7. Person item map for the 16-item unidimensional specification of the measure.......... 115
Figure 8. Person item map for the Provisions for Learning/Teaching Interactions two-dimensional specification of the measure.. 116
Figure 9. Person item map for the Structural/Process two dimensional within-item dimensional specification of the measure.......... 117
Figure 10. Wright Map for the thresholds for the easiest item (i.e., Furnishings for Routine and Care) and the hardest item (i.e., Meals and Snacks) for the 37-item unidimensional specification of the measure.......... 171
Figure 11. Wright Map for the thresholds for the easiest item (i.e., Staff-Child Interactions) and the hardest item (i.e., Nature and Science) for the 16-item unidimensional specification of the measure. 172
Figure 12. Wright Map for the thresholds for the easiest item (i.e., Staff-Child Interactions) and the hardest item (i.e., Nature and Science) for the Provisions for Learning/Teaching and Interactions specification of the measure… 173
Figure 13. Wright Map for the thresholds for the easiest item (i.e., Staff-Child Interactions) and the hardest item (i.e., Nature and Science) for the Structural/Process specification of the measure. 174

PREVIEW

Introduction

Policymakers have increasingly focused on investments in education as a way for laying the foundation of America's 21st century economy. Societal factors like the increasing cost of childcare, increased maternal labor force participation (Sall, 2014), and a focus on educational equity for all children (Burkam, 2013), have coalesced with concerns about the preparedness of U.S. children to meet the demands of a fluid and technology-based economic landscape (Education & Workforce, 2008). In an effort to address these societal concerns, policymakers have moved to provide unrestricted access to preschool for every child, based most recently on evidence by Heckman (2006; 2010), which has highlighted the strong return on investment in preschool programming. Indeed, the role of high quality preschool classrooms has become a linchpin among solutions to improve our country's school readiness efforts (Yoshikawa, Weiland, Brooks-Gunn, Burchinal, Espinosa, Gormley & Zaslow, 2013).

PREVIEW

This emphasis on early childhood education (ECE) settings is pragmatic, as ECE is theorized to be one of the most important proximal influences on children's development (Bronfrenbrenner & Morris, 1998; Weisner, 2002). It is theorized that high-quality ECE settings offer a multitude of learning opportunities that have the potential to promote children's early educational successes in positive ways. These kinds of opportunities include, but are not limited to: exposure to positive peer interactions (NIHCD, 2001), early vocabulary, reading and mathematics concepts, rich instructional materials, and nurturing teachers who facilitate positive adult and peer interactions to optimize learning (Mashburn, Pianta, Hamre, Downer, Barbarin, Bryant & Howes, 2008). The empirical literature has continually highlighted small, positive direct effects for preschool exposure on children's later educational outcomes (Gormley, Gayer, Phillips, & Dawson, 2005; Hustedt, Barnett, Jung, & Goetze, 2009; Hustedt, Barnett, Jung, & Thomas, 2007; Weiland & Yoshikawa, 2013; Wong, Cook, Barnett, & Jung, 2008).

Despite research supporting the impact of preschool on children's development, the statistics surrounding preschool utilization in the U.S. show a broad unmet need. First, the U.S. lags behind many developed countries in the world in the rate with which children 3 to 4 years of age are enrolled in formal educational programs (i.e., preschool). The Organization for Economic Cooperation and Development (OECD), in their annual report titled *Education at a Glance* (2015), show that, across 31 developed countries, the average enrollment rate of 3- to 4-year olds in formal schooling is 81%. Among the list of 31 developed countries, the United States ranked in the bottom three in terms of the number of children enrolled in preschool, with only 59% of 4-year olds and 41% of 3-year olds enrolled in preschool. In order to mitigate these crucial gaps, the U.S. is currently investing 6.2 billion dollars into preschool programs. That figure is up by 10% ($573 million) from the 2013-2014 year (NIEER The State of Preschool Yearbook, 2015).

Increasingly, as tax payer dollars are being used to expand access to preschool, accountability efforts have been commensurate (Schultz, 2015). Most of these accountability efforts have focused on ensuring that preschools are meeting expectations for quality. For example, in an effort to increase current levels of quality, some states have set up incentives for programs to engage in continuous improvement efforts, typically by tying reimbursement for services to providers' scores for quality (Pianta, 2012). The success of these efforts is contingent on the ability to measure classroom quality in a way that is both accurate and comprehensive. As such, many states have co-opted extant observational measures of classroom quality into their accountability efforts – with the most utilized measure being the Early Childhood Environment Rating Scale-Revised (ECERS-R) (National Center on Childcare Quality Improvement, 2013).

PREVIEW

Policymakers have assumed measures like the ECERS-R are valid for these purposes, and this provides some assurance that the data gathered are relevant to outcomes that are the impetus for investments into preschool. However, measures like the ECERS-R were not designed for this purpose, and scholars are increasingly questioning the validity of observational measures like the ECERS-R (Goldstein & Flake, 2016; Gordon, 2013; 2015b; Votruba-Drzal & Miller, 2016).

Currently, there is a dearth of comprehensive and rigorous psychometric evidence to support the assumptions underlying the use of measures like the ECERS-R in accountability efforts (Charalambous, Blazar, McGinn, Kraft, Beisiegel, Humez & Lynch, 2012; Gordon, 2013; 2015b; Pianta, 2012). Further, studies that have investigated the validity of the ECERS-R have failed to situate the work within larger validity frameworks. As such, findings from research have led to equivocal conclusions, and psychometric information about the measure is difficult for policymakers to access and utilize. Situating psychometric analyses of the ECERS-R in larger validity frameworks provides a systematic way to document a range of evidence which can be used to support the use of the ECERS-R for policymaking. The goal of this dissertation is address these deficits. As such, a primary aim of these analyses is to provide both researchers and policymakers with comprehensive information about the psychometric properties of the ECERS-R that can be used to either support or disprove a range of claims about the validity of the ECERS-R for its varied uses in policy applications (Charalambous et al., 2012).

PREVIEW

The forthcoming sections will first situate the work in theoretical conceptualizations for classroom quality. An overview of relevant literature will then be summarized. This will build to critical gaps in the literature, which are followed by the research questions for the study. This is followed by a description of the methodology used to address research questions, where Wolfe & Smith's (2007) Rasch validity framework is used to situate and organize analyses in line with

Messick’s (1989) construct validity framework. Next, results are presented. In the final section conclusions and policy implications are discussed, which is followed by limitations and future directions for research.

PREVIEW

Guiding Theoretical Frameworks for Understanding Preschool Classroom Quality

Measures of classroom quality can be described theoretically using the bioecological model (Bronfrenbrenner & Morris, 1998). Within this theory, children's learning and development, teacher pedagogy and practices, and classroom environments can be examined as interrelated levels which are part of a larger developmental system. Each level of the developmental system has the potential to impact children's learning and development through varying degrees of proximity to a child’s lived-in experience. Specifically, this theory provides a framework for understanding how microsystems (e.g., family and school contexts), mesosystems (e.g., the interactions between various microsystems), exosystems (e.g., neighborhood and school districts, etc.) and macrosystems (e.g., economic and social policies of a nation) can impact children’s development and learning.

At the macrosystem level, local, state, and federal policies can influence preschool quality, and subsequently impact children's development. For example, education policy is a reflection of a society's ideologies and value systems (Ball, 2006). This can be seen in the educational reform efforts in the U.S. that have occurred over the last two decades, which have favored expanding access to high-quality preschool for all children in the U.S. The aggregate momentum of these efforts has been in the service of increasing the school readiness skills of children in the U.S. A natural consequence of these efforts at the national level, is that certain pedagogical environments and practices have become favored.

The exosystem level entails the environments that affect children's everyday experience – those systems that indirectly influence children, but are nonetheless distal from their day-to-day interactions. At this level, children’s families can play a critical role in shaping their development, primarily through their decisions about the appropriateness of a given childcare

PREVIEW

arrangement for their children. Meyers and Jordan (2006) use an accommodations framework to describe how variations in the quality of care children are exposed to is driven by parental decisions about childcare. This framework highlights the importance of contextualized decisions about childcare using three dimensions: parents' a priori preferences and tastes for quality, social networks as a source for information, and parents' perceptions of available supply and resources for obtaining care. As parents navigate decisions about childcare, they construct their beliefs and preferences for childcare through accommodating trade-offs between the conditions necessary for their continued participation in the labor market, and optimal care for their children. For most parents this is not a simple consumption choice, as childcare is an infrequent purchase in the lives of most people. Further, information about the quality of childcare options is often lacking, causing the cost incurred by parents when making decisions about childcare to come with considerable risk. As a consequence, parents turn to their social networks for signals to guide their decision-making. These social networks are valuable, because they function as heuristics that are filled with cultural information and values, which aids in limiting the range of acceptable/normative options for care that parents consider. However, this limited pool of acceptable childcare options is also impacted by the supply of childcare. Parents do not decide among similar options, as factors like family income, proximity of diverse center types, and hours of operation, etc. can limit the pool of available childcare options. In sum, parents navigate these contexts in making consumption choices about childcare arrangements, and the choices parents make about childcare can lead to variation in the quality of care children receive.

Primary caregivers (i.e., teachers) are also part of a child's exosystem. The characteristics of caregivers also indirectly determine the quality of preschool environments that children are exposed to on a regular basis. For example, childcare centers with more support for teachers'

PREVIEW

professional development might moderate the quality of their interactions with children in the classroom (Pianta, DeCoster, Cabell, Burchinal, Hamre, Downer, LoCasale-Crouch, Wilford & Howes, 2014; Yoshikawa, Leyva, Snow, Treviño, Barata, Weiland, Gomez, Moreno, Rolla, D'Sha, & Arbour, 2015). Further, administrative and other supports a center provides to a teacher might also indirectly influence classroom quality (Goelman & Guo, 1998; Wells, 2017). Other teacher characteristics, such as teacher education level, have shown associations with classroom quality (Burchinal, Cryer, Clifford & Howes, 2002; Howes, Whitebook & Phillips, 1992; NICHD ECCRN, 2002; Scarr, Eisenberg & Deater-Deckard, 1994). Further, teacher perception variables, such as their sense of self-efficacy, might also indirectly influence the quality of care children receive (Guo, Piasta, Justice & Kaderavek, 2010).

PREVIEW

The microsystem entails contexts with the moment-to-moment interactions between teachers and children. At this level, interactions between children and teachers becomes a proximal driver of children's development of school readiness skills. In fact, teacher-child interactions are hypothesized to be the "primary mechanism by which children learn in the classroom" (cf. in Curby, Rimm- Kaufman, & Cameron Ponitz, 2009, p. 913). For example, complex interactions between children and educational materials in preschool environments likely mediate their interaction with teachers (Kontos, Burchinal, Howes, Wisseh & Galinsky, 2002; Kontos & Keyes, 1999). Further, research has demonstrated that levels of positive interactions between children and preschool teachers are associated with higher levels of motivation and engagement to learn in children (Howes, Burchinal, Pianta, Bryant, Early, Clifford, Barbarin, 2008; Ladd, Birch, & Buhs, 2003; Pianta, Steinberg, & Rollins, 1995). Interactions between children and ECE teachers have also been shown to be dynamic, unfolding in complex ways over time, with children's engagement associated with later teacher emotional and organizational supports (Curby, Downer & Booren, 2014). This is a finding which indicates that children might also aide in helping shape the quality of the preschool environments they are exposed to through their active participation and engagement with teachers, who in turn reshape the environment to respond to the needs of the children in their classrooms.

Constructionist learning theories are also useful for understanding the importance of the type and quality of early childhood environments. Socio-cultural theories, like Vygotsky's (1979), contend that mediated interactions between a child, and either a teacher or a more competent peer, in the use of materials for learning, can help a child move to more advanced levels of cognitive functioning. Within Vygotsky's theoretical framework, high quality classrooms contain caregivers who are sensitive in their interactions with children. Teachers create a high-quality classroom through their scaffolded interactions with children, whereby teachers are attentive to the individual needs of children in their classroom, and structure their interactions with the explicit purpose of supporting children in developing increasingly complex knowledge and skills. A critical tool that teachers utilize in these interactions is complex language, which encourages children to reason through their experiences in a classroom. Caregivers in these classrooms are sensitive to the need for continuous monitoring of the current skill-level of children in their classrooms so that these interactions can be modified. Furthermore, constructivist theories of learning (Piaget, 1952) stress the importance of children's active exploration and manipulation of their environments. As part of this active exploration and manipulation, children play the role of "little scientists" who form cognitive schemes to understand their lived-in experiences, and refine these through testing new hypotheses. Using constructivist theories as a framework for understanding high quality preschool classrooms, classrooms filled with open-ended activities and manipulatives that are organized in such a way

PREVIEW

as to encourage autonomy and exploration are optimal (Harms, Clifford, & Cryer, 2005; Stipek & Byler, 2004).

In sum, notions of what constitutes preschool classroom quality are influenced by several theories of child development. Constructionist theories have coalesced to emphasize several core features of high quality classrooms. Specifically, these theories have stressed the sensitivity and warmth of caregivers, the availability of developmentally appropriate materials and activities in the classroom, the time allotted for unstructured engagement with those materials and environments, and the importance of sensitive caregivers who scaffold their interactions with children to support children moving through increasingly complex levels of cognitive development (Burchinal, Magnuson, Powell & Hong, 2015). Further, systems theories, like Bronfrenbrenner & Morris's (1998) bioecological model and Meyers and Jordan's (2006) accommodations framework, assert the important role of the interaction between various contextual factors in shaping the quality of preschool environments to which children are exposed.

PREVIEW

Literature Review

Current Policy Context for Expanding Access to High Quality Early Childhood Education in the U.S.

The importance of early childhood education within the landscape of education policy efforts in the United States has expanded considerably under the reauthorization of the Elementary and Secondary Education Act (ESEA), which was first signed into law by President Johnson in 1965. The ESEA was reauthorized and amended with the Every Student Succeeds Act (ESSA), which was signed into law by President Obama in 2015. ESSA shifted the discussion of education policy away from K-12 to a focus on P-12, and provided states, districts, and schools with several important policy levers to both expand access and improve the quality of early childhood education programs.

Primarily, ESSA provided states, LEAs, and SEAs with considerable flexibility in using Title I funds to support the learning and needs of children before they enter kindergarten, thereby creating a powerful lever that could be used to strategically expand access to preschool. For example, Title I, Part A, gave schools the ability to use all, or a portion, of their Title I funds to provide preschool for eligible students (i.e., low income students). Under ESSA SEAs are also allowed to expand the pool of students who have access to free or subsidized preschool, by allowing Title I funds to be used for any student in a school if at least 40% of students in that school are from low-income families. An LEA may reserve a portion of funds from its Title I allocation to operate a preschool program for eligible children in the LEA as a whole or in a portion of the LEA. Further, LEAs are authorized to use Title 1 funds to improve the quality or extend the number of days children spend in childcare programs. SEAs and LEAs were also given considerable flexibility in using Title I funds to improve the quality of care, by providing early childhood teachers with access to high quality professional development. Finally, Title IV,

PREVIEW

Part X of the ESEA allowed Title I money to be allocated to charter schools that provide access to preschool, provided these schools also provide either elementary and/or secondary education.

Further, in an effort to expand the quality of early childhood education programs, ESSA stressed the importance of alignment, collaboration, and coordination between early childhood education programs and the K-12 system. The law encouraged states and LEAs to be thoughtful about the consistency and connectedness of both programs and professional standards across contexts that serve the same grade-level of children (i.e., horizontal alignment). For example, the State plans that are required under ESEA for the use of Title I funds require states to coordinate efforts across programs providing preschool to children, which includes programs administered under other departments like Head Start. The Department of Education issued non-regulatory guidance which urged states to think about how the quality of early childhood programming could be improved through aligning different early childhood programs within a State. Further, states were also encouraged to think about how their early childhood programs vertically aligned with state K-12 systems. As a policy aim, ESSA encouraged states to adopt a strategic P–3 approach to early childhood programing. Through this approach, states were encouraged to either document or develop a framework for what children should know prior to entering kindergarten. Taken in sum, both means of alignment are designed to encourage activities that are likely to increase the quality of programs (U.S. Department of Education, 2016).

PREVIEW

Ahead of ESSA becoming law, congress authorized funds to catalyze and incentivize the process of vertically and horizontally aligning early childhood programs within states. These funds were made available for two competitive grant programs, specifically the Early Learning Challenge program (ELC), which was part of the Race to the Top (R2T) challenge (Race to the Top Act, 2011), and the Preschool Development and Expansion grants, which were authorized in

2014. Both of these grant programs influenced the alignment activities recommended to states by the Department of Education as part of their non-regulatory guidance for implementing ESSA (U.S. Department of Education 2016). Previously, the R2T ELC required states to design efforts to implement integrated systems to ensure that their state’s preschool programming was of high quality. Forty states applied for these grants, and 20 received awards. States were urged to adopt cross-sector Quality Rating and Improvement Systems (QRIS) to both, rate the quality of providers, and track children's progress as they moved from preschool into the K-12 pipeline. Further, states were encouraged to link children's progress to their teachers in order to build an evidence base for what was and was not effective in early childhood classrooms. The Preschool Development and Expansion grants that were authorized ahead of ESSA becoming law were designed specifically to expand access to high quality preschool for low-to moderate-income students. Thirty states applied for these grants, and 18 were awarded. Building off of the R2T ELC, the grant process required the awardees to have or develop a QRIS. However, the Preschool Development and Expansion grants went further, in that qualified providers who received support from these grants were required to participate in their state’s QRIS.

The QRIS National Learning Network, an organization that helps states implement QRISs, defines a QRIS as “an intentionally transparent definition of the progression of program quality from basic to excellent” (Schilder, Iruka, Dichter, & Mathias, 2015). A QRIS is a systemic approach to assess, improve, and communicate the level of quality of early-childhood programs. QRISs are composed of five common elements: (1) Program standards that are used to assign ratings to participating providers; (2) Supports for programs and practitioners, typically in the form of technical assistance, to help programs in their continuous improvement efforts; (3) Financial incentives for providers to participate in the QRIS and engage in continuous

PREVIEW

improvement efforts; (4) Quality accountability and monitoring processes that are used to set benchmarks to determine how providers are meeting or exceeding expectations; (5) Consumer education that is used to provide stakeholders with key information to aid in the decision-making process about which programs are good fits for the needs of their children. A logic model that outlines the basic functioning of QRISs can be found in Figure 1.

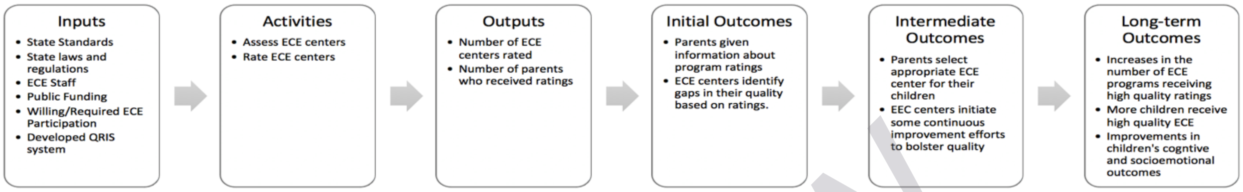
Inputs
State Standards
State laws and regulations
ECE Staff
Public Funding
Willing/Required ECE Participation
Developed QRIS system
Activities
Assess ECE centers
Rate ECE centers
Outputs
Number of ECE centers rated
Number of parents who received ratings
Initial Outcomes
Parents given information about program ratings
ECE centers identify gaps in their quality based on ratings.
Intermediate Outcomes
Parents select appropriate ECE center for their children
EEC centers initiate some continuous improvement efforts to bolster quality
Long-term Outcomes
Increases in the number of ECE programs receiving high quality ratings
More children receive high quality ECE
Improvements in children's cogntive and socioemotional outcomes

Figure 1. *Sample logic model for QRIS functioning, which is adapted from McCawley (2001) and the W. K. Kellogg Foundation (2004)*

The premise underlying QRISs is to provide parents with the information they need about the quality of early childhood programs, so they can "vote with their feet" as to which program best fits the needs of their child. The goal of providing power to parents in the form of specific information about quality of programs is that it should incentivize programs to engage in continuous improvement efforts to bolster their program quality, and as a result expand the number of high quality programs available to families within states. (Buettner & Andrews, 2009). As a result of the R2T ELC and the Preschool Development and Expansion grants, every state in America has or is creating a QRIS (Tout, Starr, Soli, Moodie, Kirby & Boller, 2010).

Both the R2T and the Preschool Development and Expansion grants gave states considerable control over defining their standards for quality used in QRISs, but required that states create systems for tracking quality across time. As states rushed to implement QRIS systems, policymakers raced to find effective indicators for program quality. As a result, many states rushed to adopt extant measures of early childhood classroom quality into their QRISs (Tout et al., 2010). The most commonly implemented measures have been the Classroom

PREVIEW

Assessment Scoring System (CLASS) and the Early Childhood Environment and Rating Scale-Revised (ECERS-R; NIEER The State of Preschool Yearbook, 2015), both of which are observational measures, with the CLASS emphasizing more of the teacher-interactional components of quality, and the ECERS-R emphasizing more the structural aspects of program quality. To date, the ECERS-R remains the most widely implemented measure of classroom environment implemented into these systems (NIEER The State of Preschool Yearbook, 2015), with 30 states adopting the ECERS-R into their programs of assessment and evaluation (Child Trends, 2014; Mashburn et al., 2008).

As the utilization of QRISs has increased rapidly, scholars have called into question whether QRISs function as intended. As a consequence, a complex picture of the efficacy of these systems has emerged. Specifically, links between measures of quality embedded in QRISs and child outcomes have been mixed. For example, Sabol and Pianta (2015) found no relations between QRIS measures of quality and child academic outcomes. Another study found no evidence between QRIS measures of quality and children's socioemotional outcomes (Hestenes et al., 2015). Increasingly, researchers are calling attention to the need to validate QRISs by examining the validity of components of QRISs, not limited to a thorough vetting of the psychometrics pertaining to ratings and associated measures of quality embedded within QRISs. Many of these measures of ECE quality were never designed to be used for the purpose of program accountability (Lahti, Elicker, Zellman & Fiene, 2014). Given the widespread use of the ECERS-R in QRISs, it is critical to ensure that rigorous and comprehensive information about the psychometric properties of this measure is available to stakeholders and policymakers.

PREVIEW

Summary of the Context for Investigating the Validity of Measures of Preschool Quality

Public policy pertaining to early childhood programing over the last decade has increasingly emphasized the role of high quality preschool in relation to child outcomes. High quality early childhood environments have been a central focus in these efforts, in part because of the well documented positive associations between these environments and child outcomes (Gormley, et al., 2005; Hustedt, et al., 2007; Hustedt, et al., 2009; Weiland & Yoshikawa, 2013; Wong, Cook, Barnett, & Jung, 2008). As such, policymakers are looking to leverage indicators of ECE quality to positively impact children's school readiness— school-entry cognitive (i.e., mathematics, language and reading), attentional, and socioemotional skills (Duncan, Dowsett, Claessens, Magnuson, Huston, Klebanov & Sexton, 2007). The passing of ESSA into law in 2015 marked a critical point in these policy efforts, as the law gave states levers to expand access to preschool to a larger number of children using Federal dollars. Further, many states are utilizing QRISs to provide parents with information about the quality of preschool options, and also using these as a tool to improve the quality of early childhood programming within states. However, scholars are increasingly highlighting potential issues with the use of QRISs for both formative and summative purposes (Hestenes et al., 2015; Sabol & Pianta, 2015; Le, et al., 2015). In doing so, an urgent need has emerged to investigate the validity of measures of quality embedded in these QRIS systems, as many of the existing instruments designed to measure quality in early childhood classrooms were not initially developed to be used for accountability purposes (Lahti, et al., 2014). Given that nearly every state in the U.S. is now implementing a QRIS system, and that most of these systems are utilizing extant ECE measures for quality, it is important to establish validity evidence for the myriad purposes for which measures like the ECERS-R are currently being used in the U.S.'s educational policy efforts.

PREVIEW